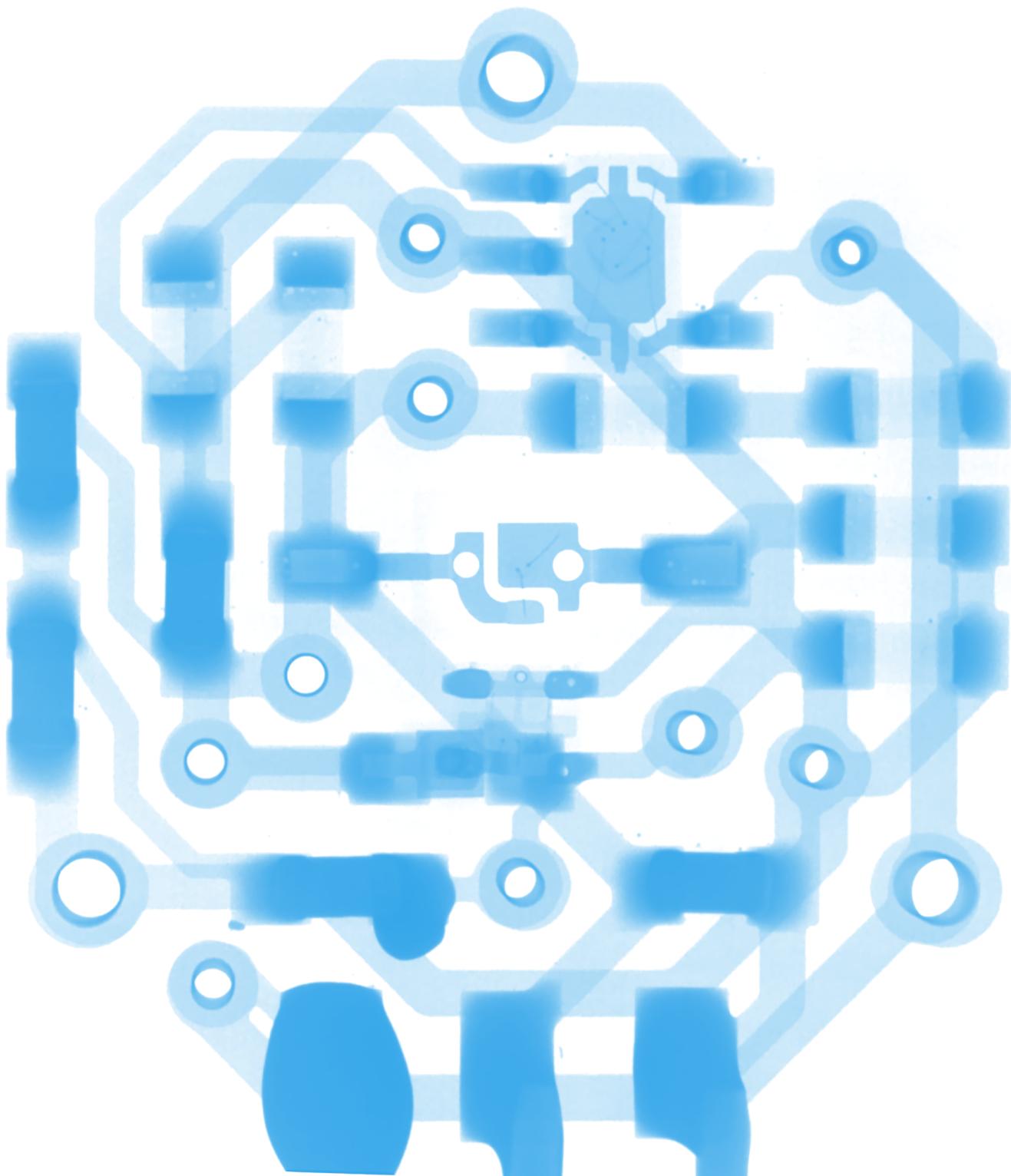


Empowering AI Leadership: AI C-Suite Toolkit

JANUARY 2022



Contents

Executive summary	4
1 Introduction to AI	6
1.1 What is AI?	7
1.2 What is the history of AI?	8
1.3 What are some examples of problems AI can solve?	10
1.4 Why have AI technologies exploded in popularity in recent years?	14
1.5 What are some positive and negative impacts of AI?	16
2 AI strategy	20
2.1 Introduction	21
2.2 What strategic options does AI create?	21
2.3 What is an AI strategy and why does it need to be aligned with your organization's larger corporate strategy?	22
2.4 What is the relationship between AI strategy and digital transformation and change management?	23
2.5 What is the relationship between AI strategy and data strategy?	23
2.6 How can you assess your organization's maturity on the AI journey?	24
2.7 How can you assess your competitor landscape when it comes to AI?	25
2.8 What steps are involved in designing an AI roadmap?	26
2.9 Who should be consulted in designing an AI roadmap?	29
2.10 What are the main reasons that AI initiatives fail to deliver business value and how can you mitigate these risks?	30
3 People and organization	32
3.1 How do you drive a culture of AI in your organization?	33
3.2 How do you evolve the organization towards implementing AI?	35
3.3 What is the role of executive sponsorship in ensuring success with AI?	36
3.4 What skills do you as an executive business leader need to lead AI initiatives successfully?	37
3.5 What opportunities, capabilities and risks of AI do business stakeholders need to understand to deliver AI initiatives successfully?	38
3.6 What kind of AI and data leader is the right fit for your organization?	39
3.7 Whom should the AI and data leader report to?	40
3.8 Who are the main stakeholders and what are the roles in an AI project team?	40

3.9 What are the most important skills to hire for in an AI team?	41
3.10 Should you hire, train or outsource these skills?	42
3.11 How do you hire and retain AI talent?	43
3.12 What is the most effective organizational design to embed AI and data-driven decision-making in your organization?	45
3.13 How can the AI team align with business, on the one hand, and engineering, on the other?	46
3.14 What are the key considerations for C-suite executives when building an AI change management programme?	46
4 Responsible AI	49
4.1 Introduction	50
4.2 Leading with ethical AI: A strategic imperative	51
4.3 Risks	56
4.4 Governance	59
4.5 Operationalization of responsible AI	64
5 Implementation of AI	74
5.1 Introduction	75
Stage 1 Design	75
Stage 2 Development	76
Stage 3 Deployment	83
Stage 4 Ongoing monitoring and feedback	84
5.2 Sample case studies of ethical AI	85
6 AI in sustainable development and industrial AI	86
6.1 AI, sustainability and the C-suite	87
6.2 Three lenses for AI in sustainable development	90
6.3 Taking action	93
6.4 Industrial AI	95
Appendix 1 Preparedness Assessment Tool	104
Appendix 2 Guidance Tool	108
Contributors	110
Endnotes	112

Disclaimer

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

© 2022 World Economic Forum. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, including photocopying and recording, or by any information storage and retrieval system.

Executive summary

A holistic, multifaceted approach to AI is needed.

“AI can enable opportunities and create new risks that were previously unthinkable.”

Any technology that has the potential to significantly transform businesses and society raises important challenges for organizations to adopt, but also potentially new risks that can prove difficult to foresee and manage. New strategic possibilities, processes, practices, skills, culture and other significant organizational changes required to *safely leverage* major technological innovations make it imperative for C-suite executives to stay close to these technological developments, understand their implications and skilfully lead and manage innovation and adoption while minimizing risks.

Artificial intelligence (AI) is arguably one of the historically most impactful technologies for business, the economy and society, considered to be a driving force behind the Fourth Industrial Revolution. According to some estimates, it is expected to create a total GDP uplift of about \$15.7 trillion globally by 2030.¹ AI can enable opportunities and perform tasks unthinkable until recently, such as large-scale visual inspections of physical assets or livestock; medical diagnoses with higher accuracy than the best experts; significant support for the creation of new discoveries and IP, for example for the development of new drugs; automated conversations with people anytime, anywhere, for example to enable efficient and effective customer service; or full automation of complex tasks, such as autonomous driving. On the flip side, if not managed carefully, it also creates new major risks unlike anything experienced with previous technologies. For instance, as discussed in this toolkit, AI can exacerbate inequalities and contribute to unfair treatment at scale, create new possibilities for malicious attacks and new unsafe scenarios, or contribute to misinformation and the creation of polarization that can pose threats to democracy and even to national security.

Given the potential of AI for value creation, disruption and destruction, it is imperative for executives to better understand “the art of the possible while managing the risks of what is possible” with this technology.² This requires a *multifaceted approach and holistic understanding of AI, spanning technical, organizational, regulatory, societal and also philosophical aspects, among others. That is the aim of this toolkit: to provide a one-stop place for the C-suite to identify and understand the multiple and complex issues*

that AI raises for their business and society.

This is necessary not only for the success and potentially survival of their own organizations, but also to ensure the responsible innovation and deployment of AI systems, taking into account the macro, short-term and long-term impact and potential second-order effects.

Building on previous World Economic Forum work to guide boards of directors on how to tackle AI,³ this document provides a practical set of tools that can help corporate executives understand AI’s impact on their roles, ask the right questions, understand the key trade-offs and make informed decisions on AI projects and implementations. While the Forum’s toolkit for boards of directors was geared towards executives that oversee the implementation of AI, this toolkit is targeted towards the C-suite executives who will operationalize the strategy. It will equip leaders with foundational knowledge of AI across several topics. Corporate executives are responsible for stewarding their companies through the current period of unprecedented technological change and its attendant societal impacts. A practical set of tools can empower them to understand the positive and negative aspects of using AI or machine learning in their businesses and oversee the responsible operationalization of AI in their corporations.

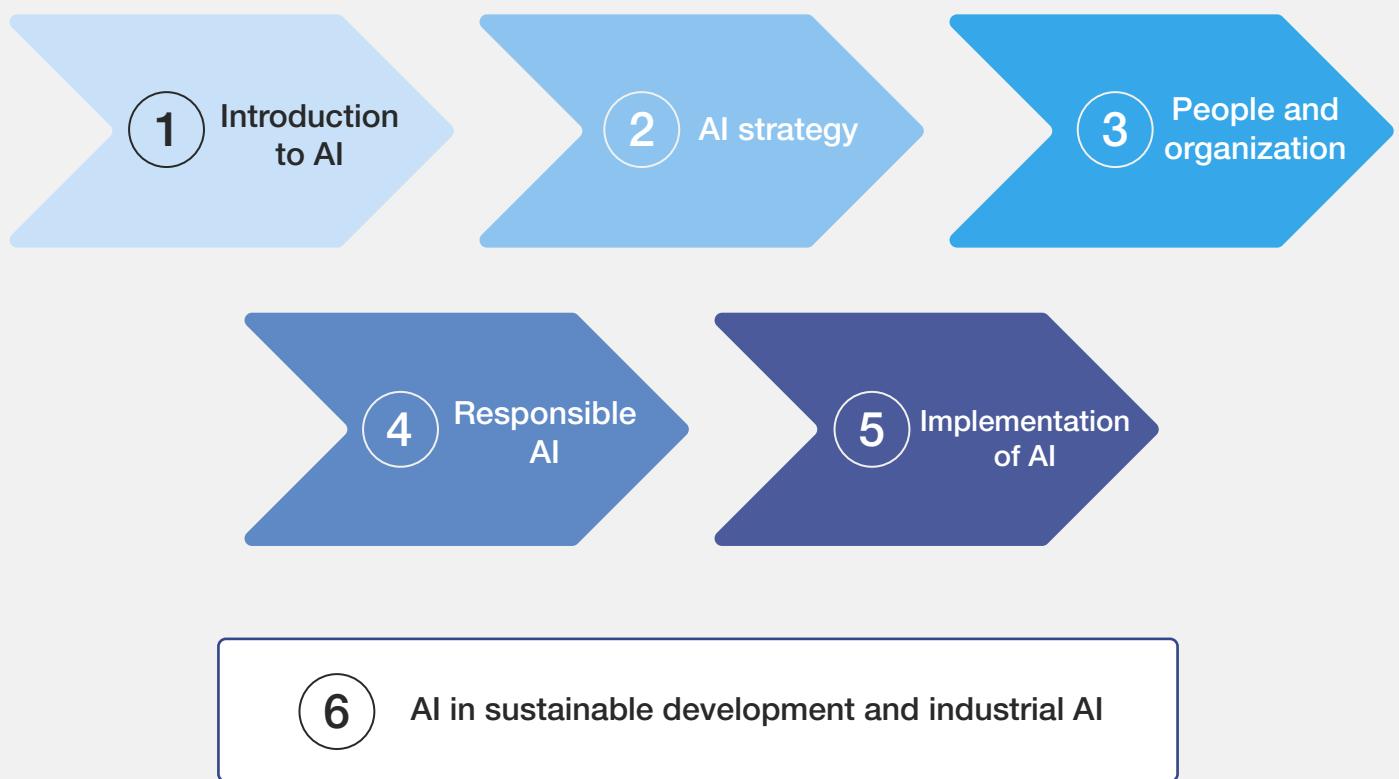
The toolkit is organized according to themes (Figure 1) covering the key aspects every C-suite needs to consider in five broad modules:

1. Introduction to AI
2. AI strategy
3. People and organization
4. Responsible AI
5. Implementation of AI

In addition, a sixth module includes two main contributions on specialized topics: AI in sustainable development and industrial AI.

They provide information on the specific topics as well as additional viewpoints on the toolkit’s five themes.

FIGURE 1 | AI C-suite toolkit modules



Source: World Economic Forum

Given the need for a multifaceted approach to AI, this toolkit is built in a modular way so the reader has a complete map of the issues and can use any of the toolkit modules without necessarily using the others. When necessary, links between the different components are also made. Moreover, as the understanding of AI and its implications continue to evolve, the idea was not to provide “complete answers” to important questions – which, given the changing nature of the field, would be impossible to achieve – but to focus instead on helping the C-suite identify the key questions they need to consider when thinking about AI.

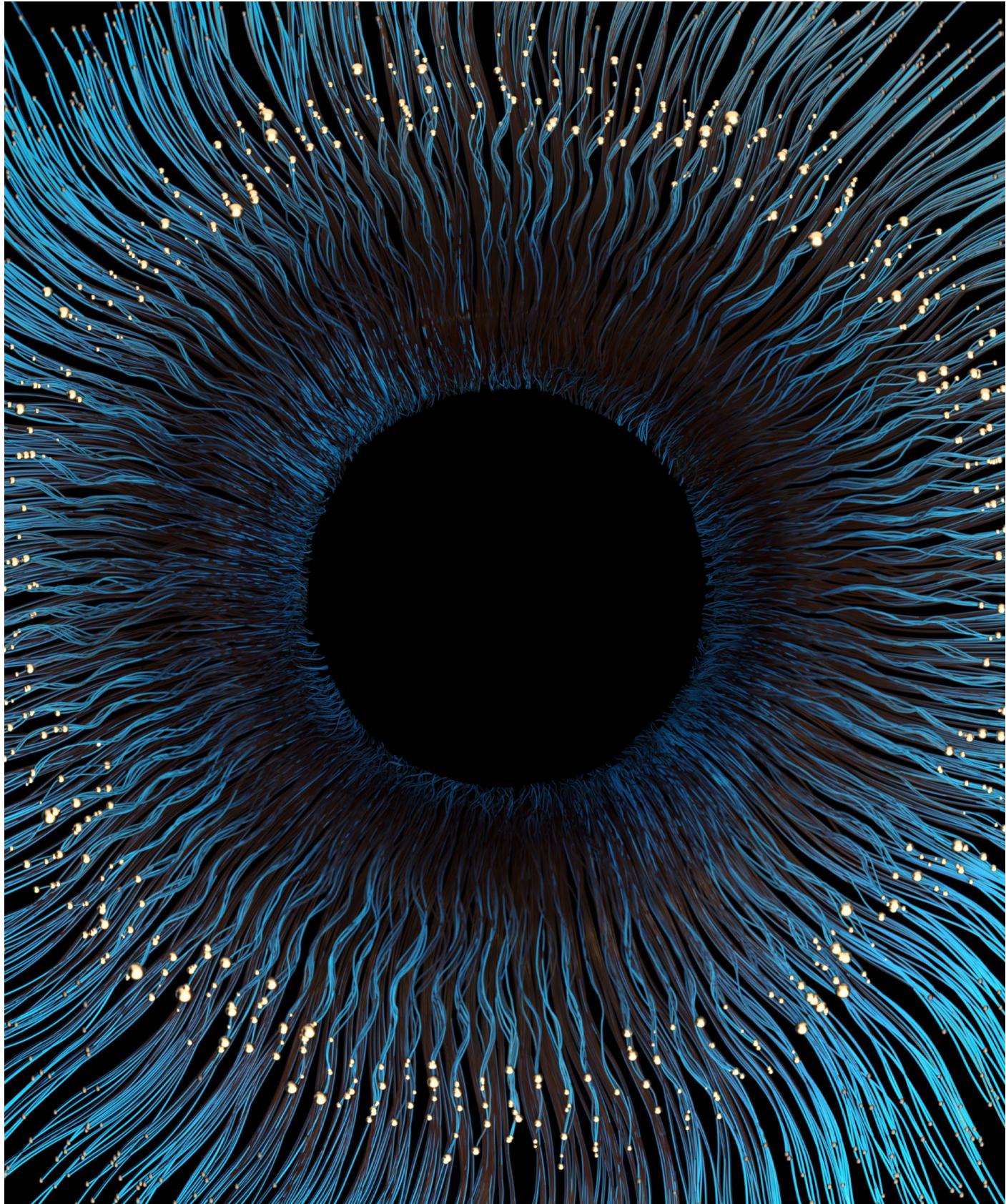
For complex topics such as the implications of AI for business and society, asking the right questions can be more important than identifying the possibly many “right” answers. Therefore, each module is organized as a set of questions and answers. The answers are by no means complete; they are meant to provide starting points for C-suite discussions. The C-suites might also consider alternative answers and approaches to the questions raised in the toolkit. However, C-suite executives are invited to consider all the questions raised in the toolkit. Doing so can help avoid important blind spots.

This toolkit is the result of collaboration between several AI experts and executive officers from diverse companies and industries to ensure it meets the needs of the targeted user community, in addition to input from relevant stakeholders to ensure it includes diverse perspectives. The contributing members brought in diverse experiences from across the AI life cycle, identifying opportunities of implementing AI and revealing the potential risks, from leading professional services and technology consulting firms, thought leaders and World Economic Centre for Fourth Industrial Revolution affiliate centres in various geographic locations.

Organizations across the ecosystem at various levels of AI maturity can benefit from the steps laid out in this toolkit to identify the opportunities of AI and implement it successfully, using relevant governance approaches that help mitigate any possible risks. The World Economic Forum calls upon organizations to pilot this toolkit and share their learnings of using it. Best practices and lessons from organizations implementing this toolkit will be incorporated in future iterations, which will further enhance this toolkit’s robustness.

Introduction to AI

Machine learning is at the core of AI, creating unprecedented possibilities and impact.





1.1 What is AI?

Machine learning has enabled the development of AI systems that can perform certain human activities often better than humans do.

Discussions about what artificial intelligence (AI) is have been taking place since the term was coined in the famous Dartmouth Summer Research Project on Artificial Intelligence, a workshop in 1956 considered to be the first event on AI as a field. Many definitions have been proposed over the years. Beyond the potential theoretical aspects of how to define both “intelligence” and “artificial intelligence”, there are important practical and business reasons to have a common definition. For example, a commonly accepted definition of AI allows consistent measurements of investments in these technologies or enables the development of regulations and the appropriate use of risk management practices. Thus various global organizations have worked on how to define AI in recent years. For example, a definition proposed by the World Economic Forum is “systems that act by sensing, interpreting data, learning, reasoning and deciding the best course of action”,⁴ while the OECD defines AI as a “machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments”.⁵ For example, this can be a system that recommends products online, predicts infrastructure failures used for preventive maintenance, or decides how to navigate an autonomous vehicle.

While there are many types of AI algorithms and AI systems, two that are widely used in business are those that employ mainly predetermined, and possibly human-defined, rules to make predictions, recommendations and decisions, and those that are learning these “rules” (which are in general mathematical functions) from data. Organizations have employed the former type of AI for years. For example, a credit card company that uses a collection of rules, often defined by experts, based on which transactions are classified as fraudulent is effectively using the first type of AI, if these rules are applied (semi)automatically using a machine. The second type of AI systems rely on what is called “machine learning”. Machine learning (ML) is at the core of most of today’s AI systems.

ML is a field at the intersection of computer science, mathematics and statistics. It focuses on the development of algorithms that can analyse data, typically large in volume and complex in structure, to identify any regularities that can be used, for example, to make predictions, recommendations or decisions. These regularities can sometimes be expressed in the form of simple rules, similar to the credit card example above, but the rules are automatically found from the data using an algorithm instead of being defined by an expert. However, for many applications of ML, simple rules are not enough for the AI systems

to make predictions or decisions accurately. Instead, today’s commonly used ML algorithms find complex mathematical functions based on data, which can be considered in a sense to be “generalizations” of rules, except that they are often not comprehensible by humans. For example, an algorithm may find a complex mathematical equation that uses the characteristics of a loan applicant as input and, after performing all calculations according to the equation the algorithm found, produces a credit score for the applicant that can then be used to decide loan approvals. These equations are identified by the ML algorithms themselves using past data which, for example in the case of credit, can be the data of past applicants, some of which have defaulted on loans and others that have not.

A special type of ML are so-called “deep learning” algorithms. These identify mathematical equations that are highly complex relative to other AI algorithms, often involving millions of calculations and parameters. Deep learning is typically used for some of the most complex problems, such as in computer vision, speech recognition or natural language processing.

ML has enabled the development of AI systems that can perform certain human activities often better than humans do. Of particular interest is the ability of AI systems to perform complex visual, speech and language tasks. There are specialized subfields of AI that focus on these problems, such as computer vision, natural language processing and speech recognition. Progress in recent years, largely driven by the availability of very large data sets (e.g. of images, text, multimedia) and the development of ML algorithms such as deep learning ones, has been significant to the extent that AI systems can “see”, “hear” or “speak”, “understand” and “write”, and enable new applications. For example, progress in computer vision allows the use of AI systems for face recognition in security checks, medical diagnoses based on medical images, or surveillance and visual inspection for preventive maintenance. Many of these specialized AI systems are currently available as relatively easy to reuse software packages. Indeed, the ability to replicate-tune-use-reuse such widely available modules enables the relative democratization of these AI capabilities at scale.

Machine-learning-based AI systems are sometimes also referred to as self-learning systems. Even after an AI system is deployed, the mathematical equations or rules the system uses to make predictions, recommendations or decisions can continue to change based on new data that is provided to the system (e.g. during usage). In

that sense, AI systems continuously change in terms of how they make these predictions, recommendations or decisions. Moreover, combining ML or other AI algorithms with additional technologies can lead to other types of AI systems. For example, ML combined with sensors or robots enables the creation of intelligent internet of things (IoT) and robotic systems.

The learning capabilities of AI systems is a key strength but also a source of potential new risks. It is important to understand the fragility of AI systems,⁶ for example their potential exposure to so-called “adversarial attacks” that can fool them into making

wrong decisions. A well-known example is the ability to fool a computer vision system into wrongly recognizing an object or face in an image by simply adding noise (a certain type of visual distortion) that is invisible to the human eye to that image. AI systems can be significantly affected by issues in the data used during their learning. For example, if that data captures people's past biases, the resulting AI systems can exacerbate these biases when deployed. Similarly, if the data used does not capture dynamics that may be happening in the future, the AI systems may fail to cope with novel situations and scenarios. It is critical to understand the potential weaknesses of AI systems and ensure their fair, robust, safe and stable use.



1.2 What is the history of AI?

AI combines mainly logic, mathematics, information science and computing. Its growth throughout the decades (Figure 2) was spurred by increasing processing power and data, as well as countless experiments and combinations of algorithms and mathematical innovations.

AI started as rule-based computing systems that tried and failed to achieve human-level performance on important tasks in the 1960s. By the 1970s, disillusion had set in and most funding stopped, starting the first “AI winter”. In the early 1980s, Japanese funding focused on expert systems; artificial intelligence was refocused on solving narrow tasks using predetermined or sometimes learned rules but with limited data and simpler algorithms, which eventually brought new products and business opportunities. Expert systems also had hard limits to their capabilities, and a brief spike of interest in AI died down by the late 80s, the second AI winter.

Progress continued with less publicity and the 1990s saw a revolution in AI: the so-called “bottom-up approach”, where “intelligence” was not preprogrammed but learned instead. ML has become the main approach to AI since then. Down the line, this approach proved more fruitful. Nevertheless, the old rules-based brute-force AI had a blaze of glory in the form of Deep Blue, which beat the reigning world chess champion Garry Kasparov in 1997.

Two main factors that spurred AI's growth in the late 1990s and 2000s were the explosion of computing power and data on the internet. In recent years, AI adoption has spread to almost all industries, including highly regulated fields such as healthcare.

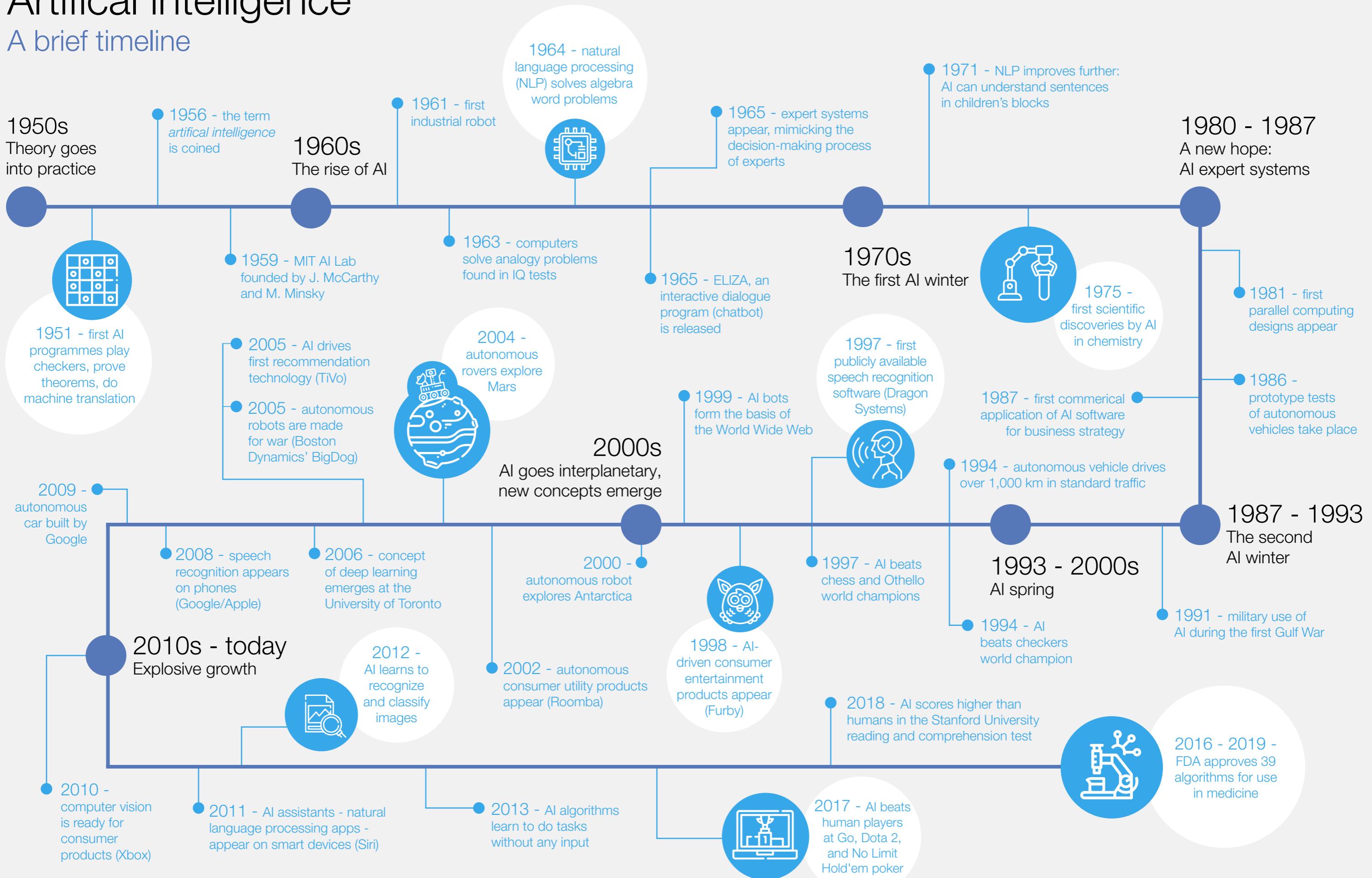
Major AI advances in the 2010s stemmed from deep learning, algorithms that develop highly non-linear models with very large numbers – billions and even trillions – of parameters, which are estimated (what is called “trained” or “learned”) from pre-labelled data sets.

Self-learning deep learning neural networks create a conundrum: their decision-making has no clear set of rules. Interestingly, much like our brains, these AI systems provide no real explanation for why exactly a decision has been made. Nevertheless, these systems often match or exceed human performance on various tasks, for example, image classification or crop fertilization planning.

Any venture that can be quantified is an area for AI. Increasingly, AI outperforms humans in more areas, but AI is not perfect. For example, it relies on algorithms that are often not transparent. Anyone applying AI should consider several key factors, such as that the data sets used for ML must be as precise, representative and clean as possible, scrutinized for biases and constantly audited to provide feedback for improving the algorithms.

Artificial intelligence

A brief timeline



Sources: Harvard University, IEEE, AIMultiple, Google, Quartz, The Medical Futurist, BBC, Thomson Reuters, Gartner⁷



1.3 What are some examples of problems AI can solve?

AI can perform certain human functions such as sensing and learning:

SENSE: AI enables machines to perceive the outside world, analysing images, sounds, speech, text and other data picked up through IoT sensors or fed directly into the system.

LEARN: AI enables machines to continuously optimize their performance by learning from previous results, successes and failures.

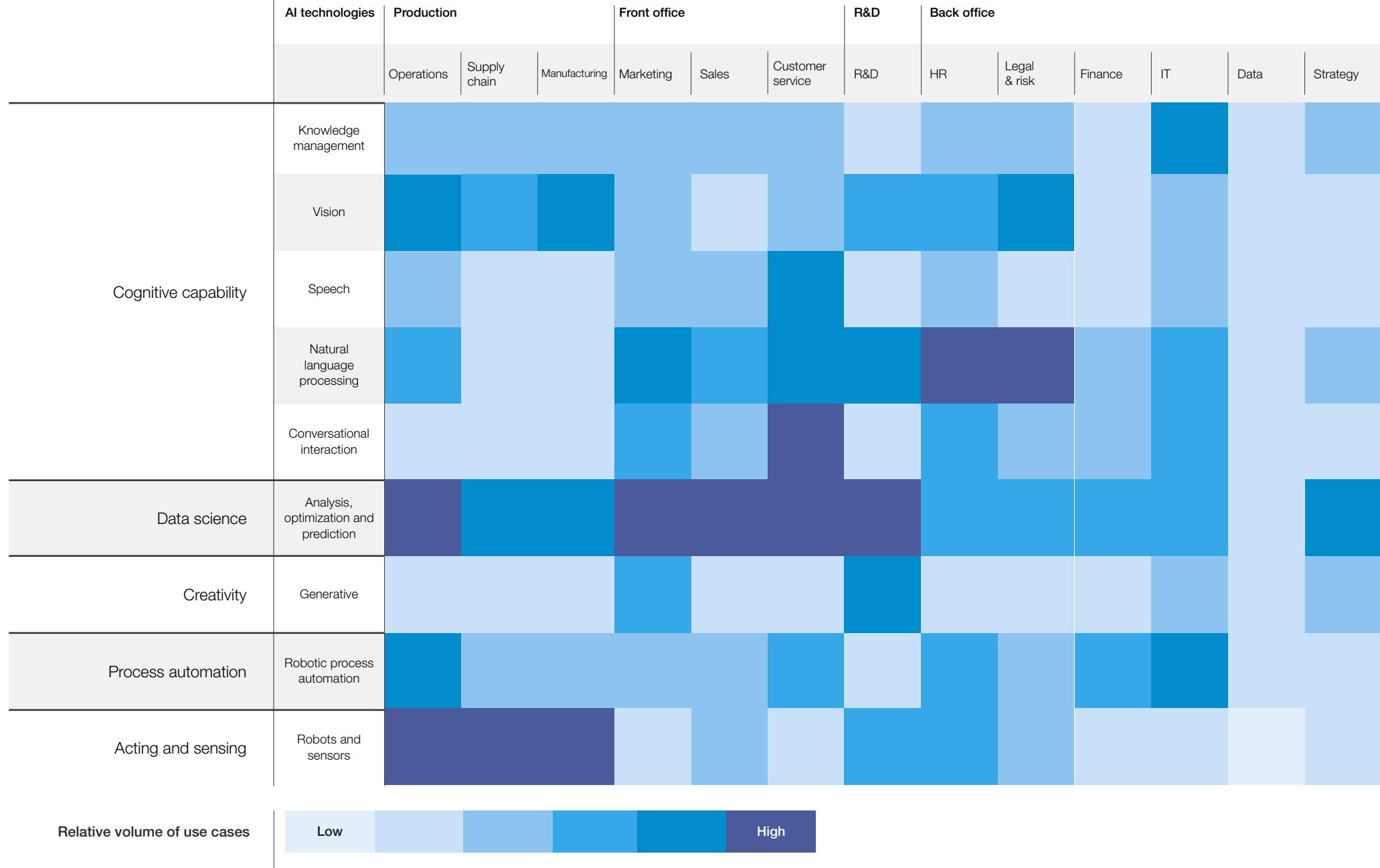
AI has slowly crept into the common. It now permeates people's lives, tracking steps, mapping directions and recommending content to watch, read and listen to. Search engines, recommendation systems and digital assistants mediate access to internet information. ML matchmaking algorithms utilize growing data stores of hyper-personal information to form new relationships around the globe. AI is used by some companies to determine whether someone should receive a new job or be approved for a loan. AI systems have defeated world champions at chess,⁸ Jeopardy,⁹ Go¹⁰ and online games like Dota 2¹¹ and ATARI 57.¹²

AI systems can perceive and act on the outside world. In the agricultural industry, image recognition is being used to monitor and raise livestock,¹³ accurately dole out pesticides to weeds while preserving crops¹⁴ and enhance farm-to-table supply chains. In the medical domain, AI is making use of vast biomedical databases¹⁵ to enhance drug discovery and disease recognition (particularly in radiology), and enabling synthetic clinical trials to test drugs before they reach humans.¹⁶ Synthetic Trials uses synthetic data created by neural networks known as generative adversarial networks. Synthetic data mirrors the statistical properties of real data,

but does not contain any personally identifiable information. Combined with the IoT, AI enables innovations such as digital twins: virtual copies of running machinery that capture real-time insight into the current state of a machine. Digital twins are useful for devices that need regular upkeep and can be monitored through smart sensors. They utilize sensor information to create precise simulations of current and future issues that may arise, and enable workers to make accurate repairs during a normal machine life cycle. Digital twins are highly useful for predictive maintenance in automotive, engineering and environmental projects.¹⁷

AI enables developers to reach larger audiences than ever before. Through natural language processing algorithms and extensive data repositories, companies can converse with millions of individuals in seconds, accessing detailed profiles of individual users' purchase histories, online shopping habits, user preferences and past conversations with virtual agents, or chatbots. Chatbots now serve as the first customer/citizen service representative that a user interacts with. Chatbots enable direct lines of communication to individual users, are capable of providing answers to frequently asked questions and can pass off to live human agents in complex situations. Major software tools such as Amazon Lex and Google Dialogflow have made chatbots simple to set up, often in low or no code environments. Regulations such as the General Data Protection Regulation (GDPR) require that chatbots and other virtual agents reveal themselves as such, to prevent confusion regarding whether or not someone is communicating with a computer. AI also enables personalized education and training at scale.¹⁸ Through massively open online courses¹⁹ and enhanced recommendation systems, learning is available on demand and tailored to individual needs.

FIGURE 3 | Common AI use cases by organizational function and technology type



Source: Best Practice AI research into use cases and case studies across functions and sectors

Examples of common use cases in sales (Figure 3) that use AI technologies for analysis, optimization and prediction include to:

- Predict optimal next actions for sales representatives (e.g. when to offer discounts)
- Generate actionable analytics from sales calls to support training and mentoring
- Personalize customer journeys across multiple sales channels
- Evaluate customer behaviour (e.g. credit risk) and likely future value
- Predict demand to support pricing decisions and supply chain optimization.

Similarly, in HR, multiple use cases use natural language processing technologies, including to:

- Automate screening processes in recruitment
- Support the personalization of staff onboarding and the provision of support services
- Provide targeted appropriate training courses
- Provide analytic and predictive support to staff compliance monitoring (e.g. spotting inappropriate behaviour such as bullying in internal emails)
- Assess employee sentiment to predict the risk of churn.

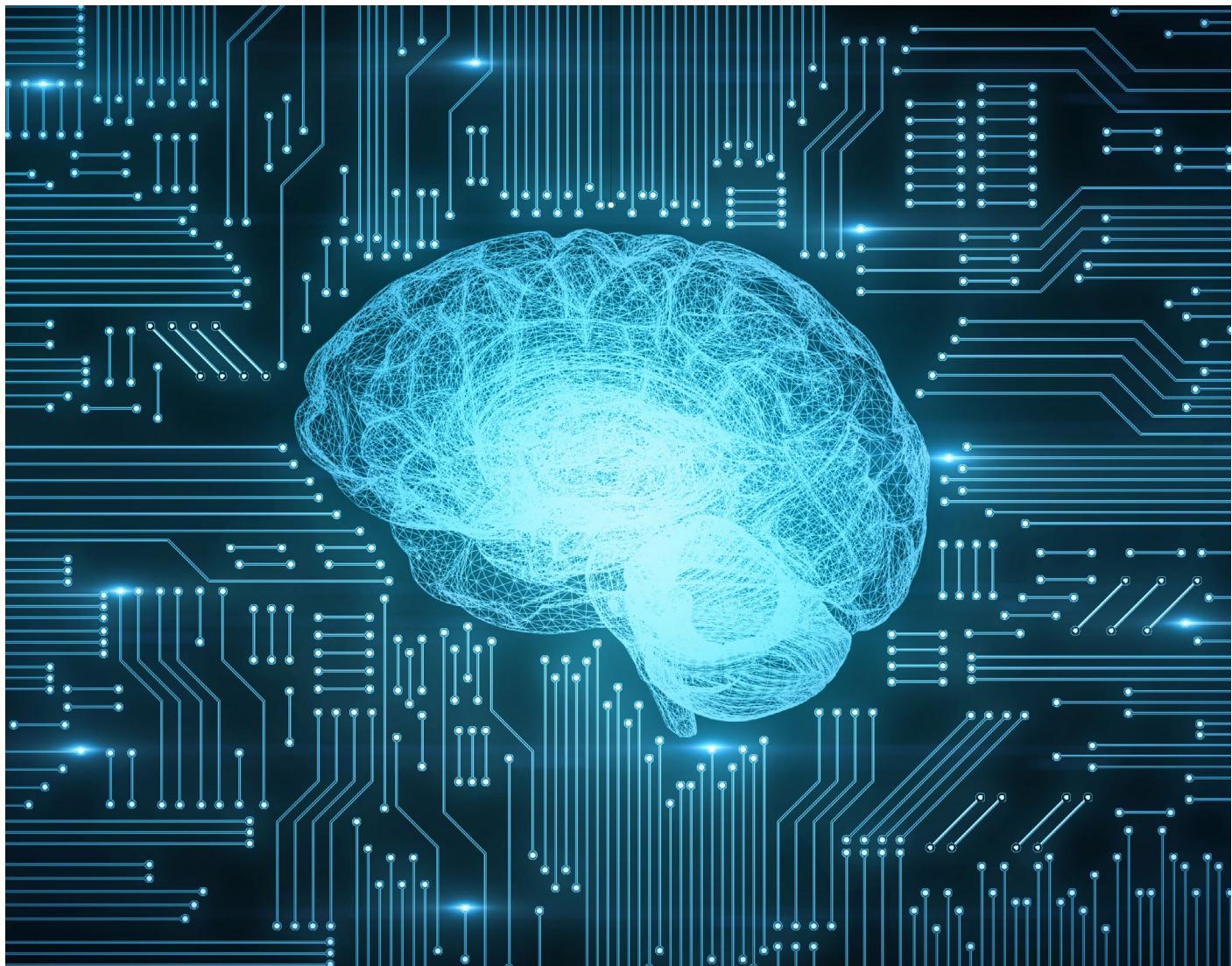
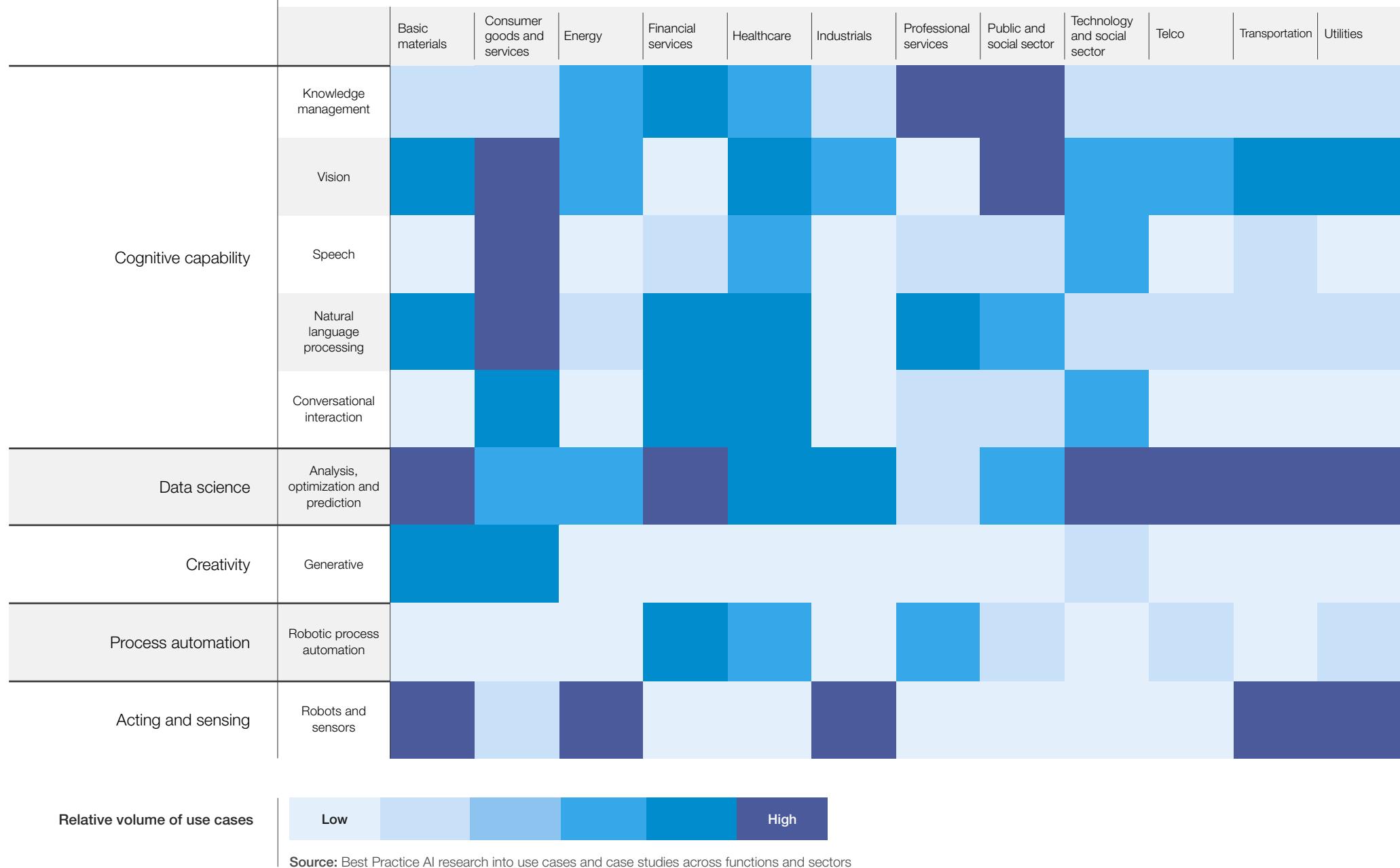


FIGURE 4 | Common AI use cases by industry and technology type



Source: Best Practice AI research into use cases and case studies across functions and sectors

Examples of common use cases involving vision technologies in the transportation sector (Figure 4) include to:

- Detect objects in drone piloting for delivery to remote areas and object detection
- Monitor rail track and road surfaces for signs of erosion that require maintenance intervention
- Analyse, predict and manage passenger traffic flows to minimize crowding (for crime, safety and COVID reasons)
- Monitor drivers for signs of fatigue (e.g. when driving heavy good vehicles over long distances)
- Accelerate check-in at airports through facial recognition.

Common use cases involving natural language processing in the consumer goods sector include to:

- Automate video captioning to allow hearing-impaired viewers to enjoy the media
- Support sophisticated chatbots capable of having higher quality interactions with consumers
- Offer translation tools for multilingual product marketing materials
- Automate analyses to check sales conversations for quality and compliance
- Create novel interfaces for products (e.g. toys that respond to audio commands).

Certain advanced AI use cases are still out of reach for many organizations that are grappling with putting the foundations for AI in place, including moving data out of silos and enabling wide-scale usage, having meaningful methods of data collection and protection, and optimizing towards one specific goal rather than trying to create a catch-all machine that can solve everything. And while AI is increasingly used in decision-making scenarios, doing so is risky because “from the data we choose to collect to the questions we ask, models are opinions embedded in mathematics”.²⁰ We must be wary of the ways in which AI can perpetuate

and amplify existing societal biases. Historical data is not sacred and AI can struggle to predict correct information given partial representations, as demonstrated by bias unearthed through the Gender Shades²¹ project. AI can enable deepfakes and disinformation²² to spread widely; while systems like GPT-3²³ are able to generate full media articles,²⁴ they have also been found to perpetuate known stereotypes.²⁵ The ability to solve problems using AI in an ethical, scalable, efficient way is an arduous journey, but the organizations that invest the time and resources to iterate until they get it right will be the winners in tomorrow’s economy.



1.4 Why have AI technologies exploded in popularity in recent years?

AI has been available in various forms for several decades. Why, then, is its usage within enterprises increasing dramatically now? Both supply and demand factors are driving the use of AI. The explosion of data has both supply and demand effects. New AI algorithms have been critical, too. Somewhat less important as a growth factor, but still quite important, has been advances in processing speed and power. Advances in software engineering and ML practices have led to the

development of easy-to-use AI tools and software packages, and AI frameworks are being made available for free in the open sources domain. Similarly, the Big Tech companies have made much of their software available in the form of APIs, which can be easily integrated into applications and have enabled the democratization of AI at scale. Finally, less important are certain socio-political factors. These are considered in the order of their importance to the growth of AI.



Data as a driver of AI – The explosion of data within businesses and societies is a key factor in the recent rise of AI. The rapid growth in data both makes AI possible and necessary. Data of multiple types has proliferated wildly in the past several decades. International Data Corporation (IDC), a market research firm, operates a Global DataSphere Project that attempts to measure the amount of data created and consumed in the world. For 2020, it estimates a total of 59 zettabytes will be created, copied and consumed worldwide.²⁶ Increasingly familiar are the large terabyte disk drives on personal computers; 59 zettabytes would fill such a drive every day for 161 million years.²⁷ IDC projects continued growth as well: “The amount of data created over the next three years will be more than the data created over the past 30 years, and the world will create more than three times the data over the next five years than it did in the previous five.”²⁸

“The amount of data created over the next three years will be more than the data created over the past 30 years.

The data comes from many sources: sensors of various types, which are the fastest-growing source; text, audio and video communications; entertainment; 14 billion mobile phones across the world, each of which leaves a data trail; internet activities; and enterprise applications in businesses. It is increasingly easier to identify the few human activities that do not generate data than those that do.

Data is an essential element for AI systems. Most of them – all those involving ML, in particular – are trained on data. Deep learning, one of the newest and most advanced forms of ML, usually requires a very large volume of data to train models. For example, GPT-3, a recently developed deep learning system to create text by machine from OpenAI, has 275 billion parameters or feature weights, and the training of those parameters was based on about 500 billion units of text. Just as the amount of available data is growing rapidly, the amount of data used to train AI models is growing at a similar pace.

Data is an important enough prerequisite for AI that any company planning to develop a ML system must think in advance about what data will be available to train the model. The data should be of large volume and high quality, and the data used for training should be “labelled”, i.e. the outcome to be predicted by the AI system should be present in the data. If a healthcare provider, for example, wants to use ML to forecast what factors predict the development of type 2 diabetes, it needs to have a training data set in which various predictor variables for diabetes are listed for each patient, and whether or not that patient developed diabetes as well. After training the model, it can be used to predict diabetes onset in patients for whom the outcome is not known.

Data is also a demand factor for AI. In other words, there is so much data available that AI is increasingly needed to make sense of it. To summarize, categorize and make predictions on data – in other words, to put it to valuable use – AI and analytics are required. As one example, cities all over the world are placing cameras on streets. It is certainly possible that all the resulting images will make citizens safer, but only if the images are viewed and interpreted. It is increasingly difficult to find enough humans to examine the images, so AI-based image recognition is taking over that role, at least as a first set of eyes.

The same thinking applies to medical images. Images from CAT scans, magnetic resonance imaging, ultrasound and nuclear medicine technologies have continued to grow in number despite increased concerns about cost and, with some technologies, radiation exposure. Even one imaging session may create dozens of images. One reason why many researchers are pursuing AI-assisted radiology is that it is becoming increasingly difficult for a human radiologist to devote significant attention to every image.

“Beyond increased processing power and data availability, new machine learning algorithms have been a primary driver of much of today’s AI.”

It is not just images that are becoming too voluminous to view. In every scientific or medical field, the number of publications is far too great for scientists or clinicians to read. AI that could synthesize all that knowledge and present it at the appropriate time would be extremely valuable in multiple different fields.

In short, it is nearly impossible for AI to exist without data and for data's value to be optimized without AI. This synergistic relationship has certainly fuelled the growth of AI and will only continue to do so in the future.

Processing power and the cloud – The other key supply factor explaining why AI is feasible and popular now is increased processing power, on the enterprise premise, but especially in the cloud. The relentless advance of Moore's Law has made possible the rapid and relatively inexpensive development of AI models and production deployment in virtually any size organization. In the cloud, massive amounts of computing power are available on demand. The cost to process an AI application today is a small fraction of what it would have cost a decade ago, if it were even possible.

Now also microprocessors are tailored to run particular AI application types. Graphics processing units, or GPUs, were found (somewhat through luck) to run deep learning applications extremely quickly. Today a variety of microprocessor companies manufacture GPUs, and some are even optimized for particular types of deep learning code. Several cloud processing vendors offer GPU processing in the cloud. Without these processing capabilities, AI in its current form could not be as efficient or effective.

New algorithm types – Another factor fuelling AI's growth is a new set of algorithm types. Already mentioned is deep learning, a complex form of

neural networks that has come to maturity in the last decade or so. Within that family of algorithms also lies reinforcement learning, a technology for improving machine models over time, and generative adversarial learning, an approach to letting deep learning or other models challenge and learn from each other. GPT-3, the data-intensive system for generating text, relies on another new class of deep learning related model called a “transformer”, a type well suited to sequential data like text. These new algorithm types have been a primary driver behind such AI use cases as image, video and voice recognition, speech recognition and natural language processing, and natural language generation.

The progress outside of ML has not been as rapid; older AI technologies like rule engines and semantic natural language systems have not progressed much over the past several years. However, combinations of AI technologies are certainly advancing the field. Many companies, for example, are combining robotic process automation (RPA) with ML. And RPA itself is a combination of rule-based approaches with other types of computing, including graphical user interfaces and presentation-layer integration.

Socio-political factors – Socio-political factors that are somewhat responsible for AI's recent rise include the growth of technology capabilities in China, the relative absence of data privacy standards in the United States, the massive acceptance of digital business offerings in many countries around the world, and spending by government bodies (for example, the Defense Advanced Research Projects Agency (DARPA) in the US, and several government institutions in China) on AI research and development. However, these factors pale in comparison to other “why now?” drivers as an explanation for why AI has seen rapid growth and uptake over the last decade.



1.5

What are some positive and negative impacts of AI?

AI is already transforming the world, “from moment-to-moment interactions people have with devices in their homes to large-scale decisions made by global organizations that affect millions of people”.²⁹ A McKinsey report estimated that deep learning alone, a subset of AI, could enable up to \$6 trillion in value annually.³⁰ It is not just about speed to market and value-add. Global leaders have the power to ensure their organization builds and runs ethical AI that helps make the world a better, not a worse, place.

While more attention is focused on the negative impacts of AI, it is important to focus on the positive impacts as well. They provide motivation for research, development and investment. And if anticipated positive outcomes are not being achieved in an AI project, it is important to review the project's design and implementation and to refocus it.



81 %
of executives
are optimistic
about AI.

Positive impacts

Many of AI's positive impacts are still potential, but some sectors are further along than others in achieving benefits from the technology. In a 2021 survey of technology executives in large businesses by NewVantage Partners, for example, positive impacts were strongly suggested. Almost all (96%) of the respondents indicated their organization has achieved successful outcomes from big data and AI – up more than 25% from the 2020 survey. Most (92%) reported that the pace of big data/AI investment in their organizations was accelerating – up 40% from the 2020 survey. Overall, 81% were optimistic about the future of big data and AI in their firms.³¹

One general benefit of AI, at least in principle, is productivity. Technologies that can think and act with a high degree of intelligence and autonomy offer plenty of potential productivity gains. Little productivity growth has occurred in the US, Europe or other advanced economies over the past decade. In the US, productivity growth has not reached 3% since the mid-1970s and, over the last decade, it has grown at an annual rate of 1.2%. Over the last two years, it has grown at only about half that dismal rate, and Europe's annual productivity growth is even less at about 0.5% of late. Yet productivity is needed to grow economies and increase the standard of living.

The definition of increased productivity, of course, is that the same number of workers can do more work, or fewer workers can do the same amount of work. While greater productivity could lead to job loss, thus far it appears that, except in the case of physical and software robotics, AI is not generally replacing workers but freeing them up to perform more complex tasks. In addition, AI-enabled productivity could counteract the negative impact of ageing workforces and declining labour force participation. It could also lead to higher levels of leisure, although it does not seem to have done so yet.

Most AI applications are specifically aimed at knowledge work processes within organizations, and these have been particularly important and problematic in terms of productivity. Peter Drucker argued as far back as 1959 (when he coined the term "knowledge worker") that knowledge work productivity is critical for economic success in advanced economies. Drucker also pointed out that the professions with the highest percentage of knowledge workers – healthcare, education and professional services – have some of the worst productivity improvement rates. Witness, for example, the very high cost of healthcare and university education in many countries. If some of these tasks could be performed by smart machines – and they can – it would be a positive outcome for many.

Other general benefits of AI in business include better, more data-based decision-making, more

targeted marketing offers and advertising, efficient sales processes based on "propensity to buy" models, optimized supply chains and better-managed human resources. In digital marketing, ML is already used to target particular publishers and individual customers with digital ads. Chatbots and intelligent agents offer the potential for better customer service, available at all times. The same technologies are also being used within organizations to interface with employees. Some of these benefits have already been achieved to a substantial degree, such as better credit decisions and the identification of money laundering in banking.

AI is also of benefit to manufacturing companies. Some, of course, are embedding more intelligence in their products. AI-enhanced products are becoming available in vehicles, telecommunications and computing devices, industrial machinery, home appliances, and many others. Automobile and truck manufacturers are rapidly trying to make their vehicles more autonomous, which could transform mobility and the structure and operation of many cities. AI is also being used to transform various aspects of the manufacturing process itself. Robots continually become more collaborative and flexible. They increasingly have the intelligence that only AI can provide. And as they start to communicate and learn from each other, their intelligence will mushroom. In addition, AI can make industrial machines more efficient and reliable by diagnosing faults and predicting the need for maintenance, ultimately reducing or eliminating unplanned downtime in that machine.

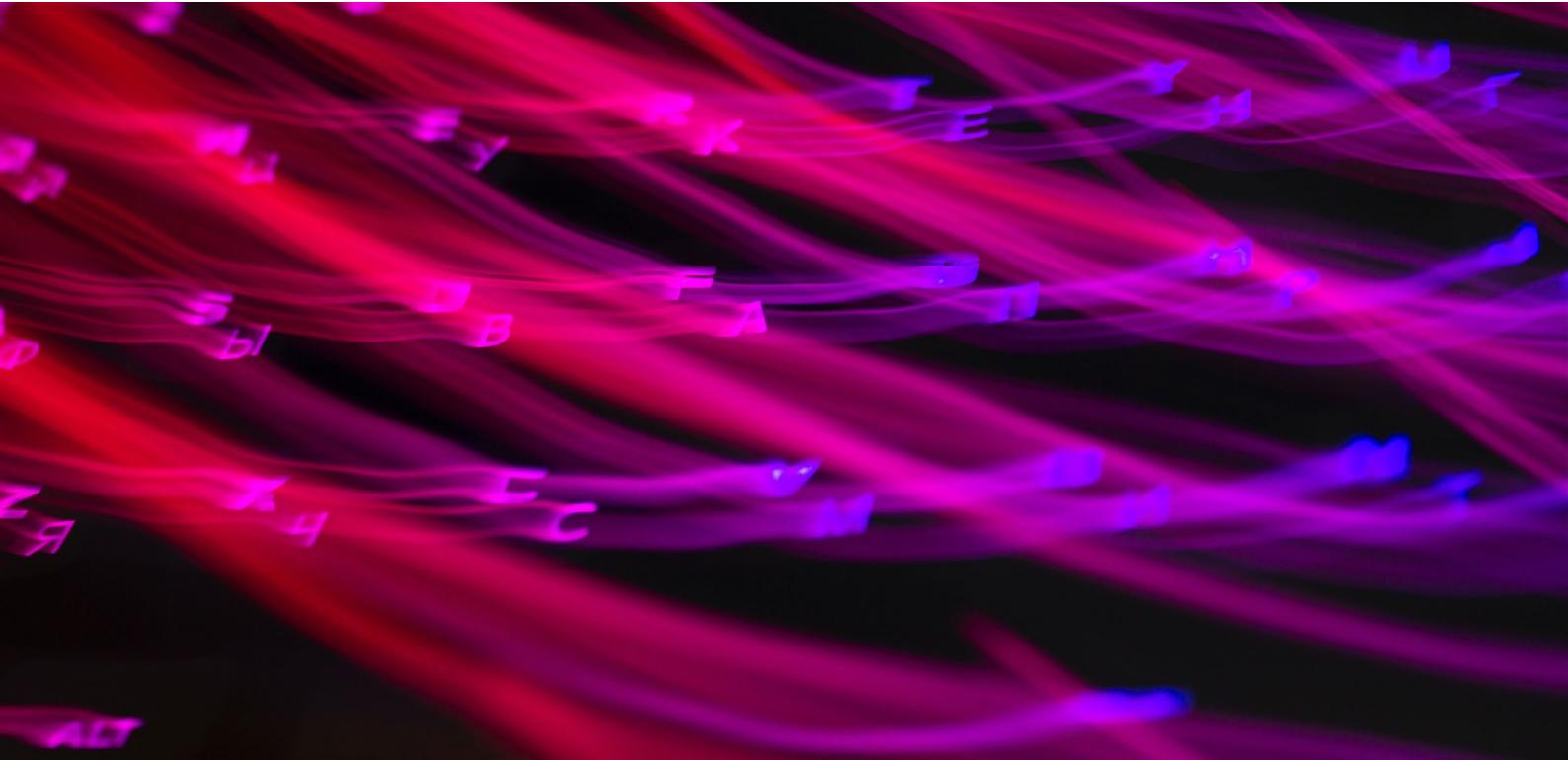
As suggested earlier, AI holds many potential benefits for healthcare, although they have not yet been fully implemented in clinical settings. A key use case is patient diagnosis and treatment based on applications like precision medicine and radiological image recognition, both based on forms of ML. AI applications in research labs have achieved or exceeded human capabilities for diagnosis and treatment recommendations for many different diseases, including cancer, diabetes and sepsis. AI can also be used as a tool to increase patient engagement in care, using customized "nudges" to change behaviour.

AI technologies could also make prescription drugs cheaper and bring them to market faster. While early enthusiasm about AI for drug development has faded somewhat, several start-ups in the pharma industry have R&D processes based heavily on ML on genomic and proteomic data. AI can also guide "structural biology" to design new drugs and vaccines, and has already been used to successfully predict the structures of protein folding. Additionally, pharma firms are using ML to target the physicians and patients who are likely to achieve better outcomes with their drugs.

Most AI applications are specifically aimed at knowledge work processes, which have been particularly important in terms of productivity.

However, positive outcomes from AI in healthcare are still often restricted to research labs rather than clinical practice. Regulatory factors, implementation issues and slow adoption by clinicians are some of the reasons the technology has remained experimental. AI has not yet made treatment more effective or lowered its cost. It is also fair to say that AI has had only a limited role in drug development and administration. But over the next decade, it is likely that AI will play a much more substantial role in improving human health.

Beyond business and healthcare, the potential benefits of AI related to daily life are many. AI could make energy consumption more efficient, as it has done for some cloud vendors' giant data centres.³² It could make government services more efficient and effective. AI could provide recommendations for better decisions in many aspects of life, including how to invest, when and how to retire, deciding where to live, what job to take, what school to attend, and so forth. Some have argued that AI will become as valuable – and as invisible – as electricity, enabling many aspects of human life to become easier and more pleasant.



Negative impacts

As with many of the benefits of AI, several – but not all – of the negative impacts are potential rather than present in practice thus far. It is important, however, to understand the potential negative impacts to prevent them when possible. Further information on the types of risks AI can raise is available in [module 4](#). “Responsible AI”.

One of the most frequently discussed negative impacts of AI is job loss. Already mentioned is the possibility that AI will lead to substantial job loss through automation of human work tasks. This is a large and complex issue that is often addressed by various researchers, the media and the World Economic Forum.

The concern about large-scale automation probably reached its peak in 2015, when several books acquainted readers with the idea that AI was improving rapidly, humans were not, and technology was likely to replace many workers across a variety of industries and business

functions. This logic and some further investigation informed a series of predictions about how likely AI is to take over human jobs. The method employed in these predictions is usually to break down jobs into their constituent tasks, and then to assess how likely future AI will be able to perform that task. If a majority or sizeable minority of the tasks can be performed by a smart machine, the job is classified as “automatable”. These analyses speculated that up to half of all jobs could be automated.

Other researchers have emphasized that “whether a job was technically automatable was only one of several factors determining whether it would be automated over a certain timeframe. Other factors include the costs to automate, the scarcity and costs of the human workers”³³ who might do the activity, positive and negative factors in the job performance of humans vs machines, and regulatory and social acceptance issues. Including these factors has resulted in reductions in predicted job loss from near 50% to below 10%.

As AI is used to enhance or automate decisions, issues of accountability that are difficult to resolve become critical.

While some job loss is certainly possible, thus far “augmentation” – smart humans working alongside smart machines – has been more common than large-scale automation. AI tends to support or automate tasks, not entire jobs. Almost every job consists of a variety of tasks. While the mix of automatable vs non-automatable tasks varies across job types, relatively few jobs have so many structured and quantitative tasks that they can be fully automated. Surveys also suggest that eliminating jobs through automation is a low priority for most managers. Researchers studying particular jobs have also found that job loss because of technology has been a relatively slow process. Finally, some researchers have posited that AI will create many new jobs and tasks.

Algorithmic bias is another potential downside of AI. The use of ML AI in classification or prediction tasks often comes with the risk of bias introduced in training data that can put certain groups at a disadvantage. This has already been observed in some situations, including algorithms that are used to score job applicants and appear to be sexist or racist, facial recognition systems that inaccurately identify people of colour, or AI-based sentencing recommendations to judges that produce racist outcomes. Another form of bias comes when ML algorithms behind social media propagate fake news or enable the targeting of individuals for political campaigns. Organizations implementing ML need to perform small-scale experiments and simulations before implementing algorithms; regularly evaluate data sets used for training; and involve human reviewers of algorithms and outcomes. In politically and socially sensitive domains like judicial sentencing, firms may find it necessary to publish their algorithms to preclude accusations of bias. It should be pointed out, however, that many human decisions are subject to a wide variety of cognitive biases, as researchers in behavioural economics have pointed out. Some degree of bias in AI-based decisions may still be better than decisions made by humans.

Lack of transparency is another negative impact of certain forms of AI. It is often argued that the decision outcomes of some ML algorithms, such as deep learning, cannot be easily explained due to their complexity and the vast amount of feature

layers involved in their production. This could lead to unexplainable personnel evaluations, student test results or sentencing decisions, all of which have already occurred somewhere. Organizations need to respond to regulators’ calls for explainability by either avoiding the use of “black box” AI applications for important decisions or pairing black boxes with so-called “explainable AI” algorithms that can provide explanations for the outputs of these algorithms. Being open about the data that is used and explaining how the model works in non-technical terms can also be necessary to ensure customers’ trust and to avoid potential dysfunctions triggered by the lack of transparency. However, AI explainability may itself come with unnecessary caveats³⁴ that need to be considered depending on the use case. The goal is to use the best performing AI algorithms while ensuring trust and user acceptance.

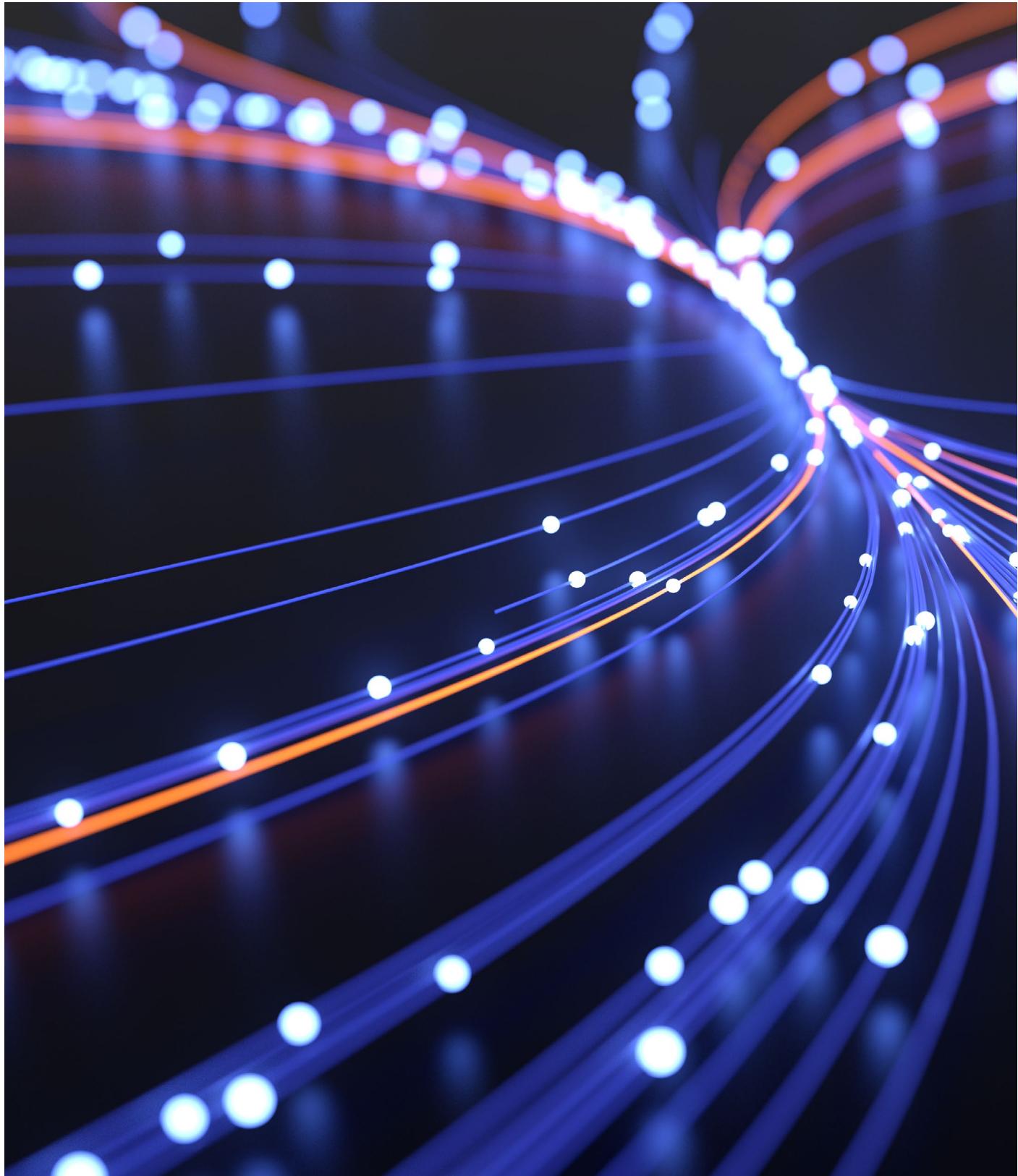
As AI is used to enhance or even automate decision-making, issues of accountability arise in decisions with poor outcomes. Who is responsible in the case of a traffic accident with a driverless car? Who should be blamed for approving parole to a criminal who eventually commits another crime? Who bears responsibility for a large financial loss in algorithmic trading? These are only a few of the cases where decision accountability is difficult to resolve. Managers and political leaders will need to proactively focus on the accountability for AI-based decisions before potential harm occurs. They also need to carefully consider the stakeholders of AI systems and outcomes (AI developers and designers, business users, customers, institutions) and clarify responsibility and legal liability upfront.

Finally, AI systems may be implicated in the loss of individual privacy. AI’s need to process increasingly large amounts of data, much of it involving consumers, may conflict with people’s rights to privacy and autonomy. Organizations need to ensure that their data practices comply with the relevant regulatory frameworks on the use of personal data and avoid possible privacy violations. Clear descriptions of what data is being used, permissions for using personal data, anonymization approaches and transparent algorithms are means to reduce threats to personal data privacy. Overall, openness about data and how it is managed and used is necessary to ensure customers’ trust.

2

AI strategy

A successful AI strategy clearly defines a vision for the organization and how that vision breaks down into specific, measurable goals.



2.1 Introduction

The most important contribution executives can make to setting up a successful AI strategy is first to clearly define a vision for their organization and how that vision breaks down into specific, measurable goals for each department or team in the organization, and identify how AI can support the vision and corporate strategy. This is because with AI, possibly more than with any other function or technology, it can be tempting to be swept up in all the hype – buzzwords like personalization, sentiment analysis, reinforcement learning and neural networks, to name just a few – forgetting to consider whether these tools and techniques truly have the potential to help deliver on top priorities. Having a defined business strategy provides the

data and AI talent with a clear understanding of what initiatives they should prioritize – those that support their organization's overarching goals.

This module focuses on the AI strategy, how it links to C-suite executives' vision for their business and the data strategy and change management programmes that support it, how they can assess their organization's data maturity and benchmark against their competitors, how to design an AI roadmap, prioritizing use cases according to their projected impact and costs, and whom to involve in this process. It also outlines the most important business risks they are likely to face as their organization increases AI.

2.2 What strategic options does AI create?

AI shifts the cost base from people-dependent variable costs to software economics of fixed cost that can be scaled relatively cheaply.

Artificial intelligence will become an increasingly important part of a company's strategy. AI is an enabler that has the potential to affect functions of a business across all sectors (Figures 3 and 4). Typically, companies approach AI from one of two angles: 1) a way to automate certain key business processes or tasks; or 2) more ambitiously, a platform to deliver a firm's data-driven transformation. The direction of travel for both is likely to be similar, but the steps to get there are very different.

Separating AI strategy from the broader strategic context is usually best avoided. The three key points, among others, to understand about AI when considering its potential role in your organization's strategy are:

1. It is important to understand the potential economic impact of AI. AI tools look for and find patterns in data enabling predictions to be made that then can enable decision-making, whether automatically by AI or by people supported by the AI predictions. Many common tasks, such as medical triage, advertising targeting and quality control on a production line, can be framed as making a prediction. This potentially delivers competitive advantage because it shifts the cost base from people-dependent variable costs to one closer to the software economics of fixed cost that can be scaled relatively cheaply. This changes the way that firms compete in a market; for example, competitive markets with low variable costs tend to lead to lower prices. This in turn will potentially drive a significant increase in demand for a service.
2. Successful AI needs to be part of a broader digital transformation. Moreover, the organization needs to be in a position to benefit from its transformational potential on firm economics. The configuration of data infrastructure, clear

management goals, operational alignment and appropriate business model required for successful AI implementation means that it needs to be part of a broader organizational plan. This is rarely a simple task, especially when a complex and often misunderstood technology is at the heart of the change. Competitive advantage in AI is as much captured from having the capabilities to do AI as it is from actually doing AI.

3. AI is different from most technologies deployed in terms of the wider impact across stakeholders. Multiple internal teams will be needed to implement it and potential staff concerns will need to be addressed but external stakeholders will also need to be engaged. These may be customers worrying about data privacy, suppliers will need to understand new ways of working, the media excited by the opportunity to explain the future, or regulators keen to monitor the approach. Investment in communication, and education, will help maximize the chances of success. And this cuts in more than one direction: the organization may well need to better understand how it is being impacted by AI deployment elsewhere. One example is how to manage search optimization strategy to maximize an organization's digital profile on search engines like Google.

AI also expands opportunities for automation with sensing capabilities to see, hear, read and analyse at mass scale. For example, it can "read" legal documents to find potential contractual issues. It can "view" closed-circuit TV cameras in real time to identify suspicious behaviour around a facilities. It can find anomalies in internet traffic patterns to identify potential cybersecurity breaches. It can "watch" interviews of candidates to determine behavioural characteristics. These new AI capabilities also come with associated risks around privacy and trust.

2.3 | What is an AI strategy and why does it need to be aligned with your organization's larger corporate strategy?

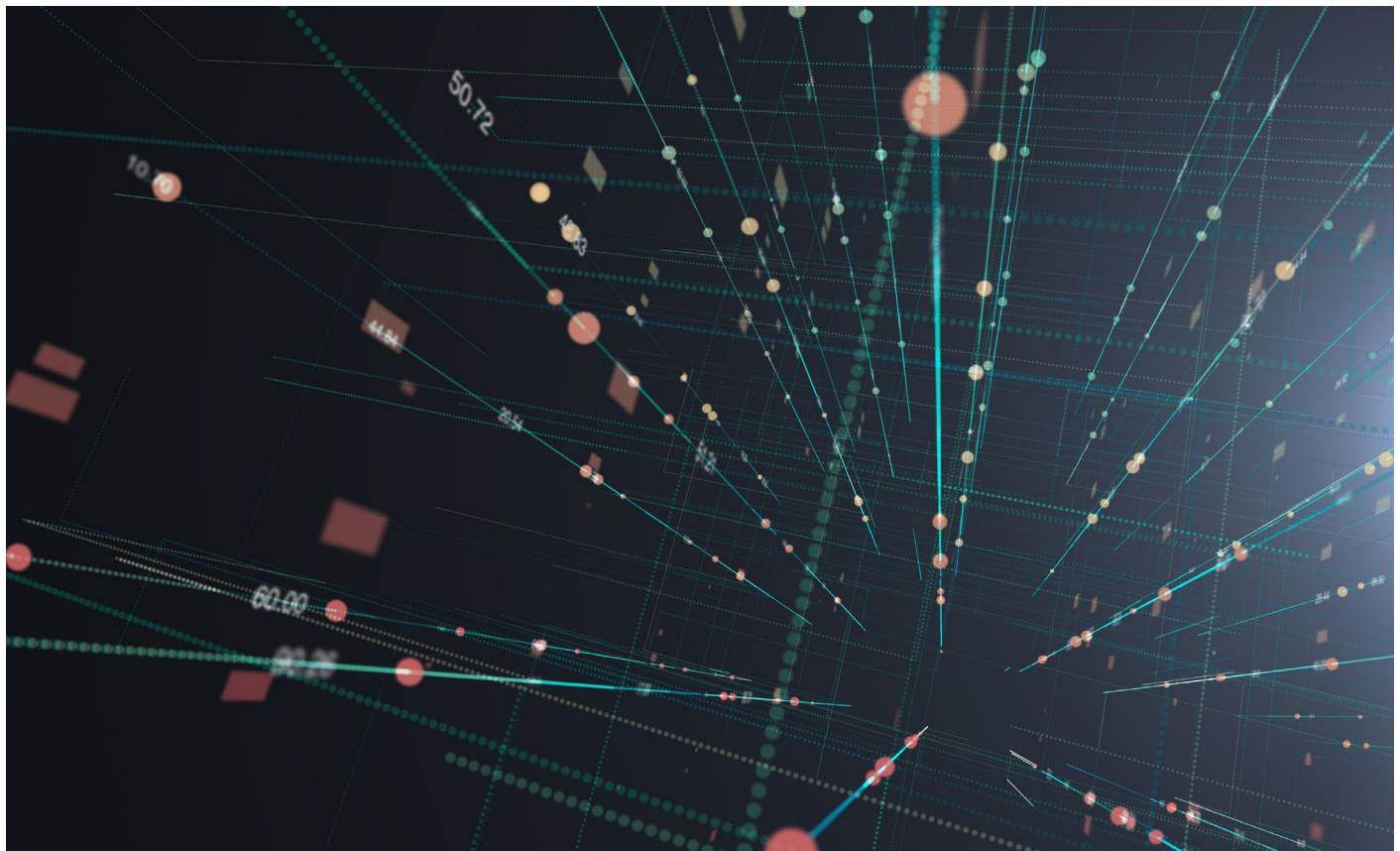
An AI strategy is simply the way in which AI supports your organization's larger corporate strategy. It needs to define the role that data and AI will play in the enterprise (in case it does play a role) and enable key data and analytics functions throughout the organization to deliver on potential use cases. This includes an understanding of the organization's current state, target state and a prioritized roadmap of the steps to implement potential AI applications. Both incremental and, when possible, more disruptive AI enabled business process/product/practices improvements and innovations should be considered in the organization's AI portfolio. Organizational leadership must understand the value and benefits but also the risks AI can enable, and align key stakeholders on the AI vision. Furthermore, it is important for the organization to allocate the appropriate resources to these activities to ensure they are realized and actioned.

In addition to an enterprise AI strategy, independent lines of businesses are also encouraged to develop a roadmap on how they wish to use and embed AI within their business.

Articulating a clear vision for the company and how that vision breaks down into concrete goals – key performance indicators, and objective and key

results – for different business units is one of the most important contributions executive leaders can make to set up their organizations for success with AI. Having a strategic vision with clear supporting targets is the foundation for success with AI for three main reasons:

1. It minimizes the chances of building data capabilities for the sake of it. AI has so many buzzwords and "shiny new toys" attached to it that it is easy to go down rabbit holes chasing the latest trends. Having a clear strategy ensures that the company makes AI investments that have the power to move the metrics that matter.
2. AI initiatives generally require significant cross-functional collaboration to be effective. Clear business goals unify cross-functional teams by providing them with common outcomes to work towards.
3. AI teams tend to receive a lot of inbound requests from business teams, many of which may be of low value to the company and underutilize the team's skill sets. Servicing them all is a quick way for an AI team to become a bloated cost centre. A clear vision will help them prioritize requests and initiatives effectively.



2.4 What is the relationship between AI strategy and digital transformation and change management?

AI-driven change often requires hard trade-off decisions about people and products, about building the future and optimizing the present business.

With the introduction of new personnel and business functions supporting the development of AI comes the requirement to help the current workforce adapt to the changes through a robust change management/transformation function. A change management function is intended to create business transformation programmes and help the organization adapt to new technologies and processes. To maintain a strong change management function, the organization must follow and measure organization change management standards across projects and ensure all stakeholders are involved, informed and empowered to partake in the transformations. Change management often works to:

1. Understand what changes are required from the current practices and culture
2. Identify who (the roles that) will be impacted
3. Design programmes that help resources to understand the changes needed to adopt the new procedures
4. Communicate the value (e.g. "data/AI evangelization") as well as change process

5. Allow those impacted to provide feedback on the roll-out of the changes
6. Track the success or failure of changes over time with key performance metrics.

A successful change management programme or function encourages a stronger innovative culture, a willingness to experiment with new technologies and improved knowledge management.

The level of organizational change that AI requires often involves hard trade-off decisions. These range from decisions about people to product lines. Running through this will be the trade-off between building the future and optimizing on the present business. Decisions on what to invest in, what to sell and how to consider customers and marketing will be hard. Key to getting more of these decisions right than wrong (a 100% success rate is unlikely and is not a helpful standard to hold anyone to – it will minimize the willingness to take risks) is a sense of direction: a strategic vision for your firm in the age of AI.

2.5 What is the relationship between AI strategy and data strategy?

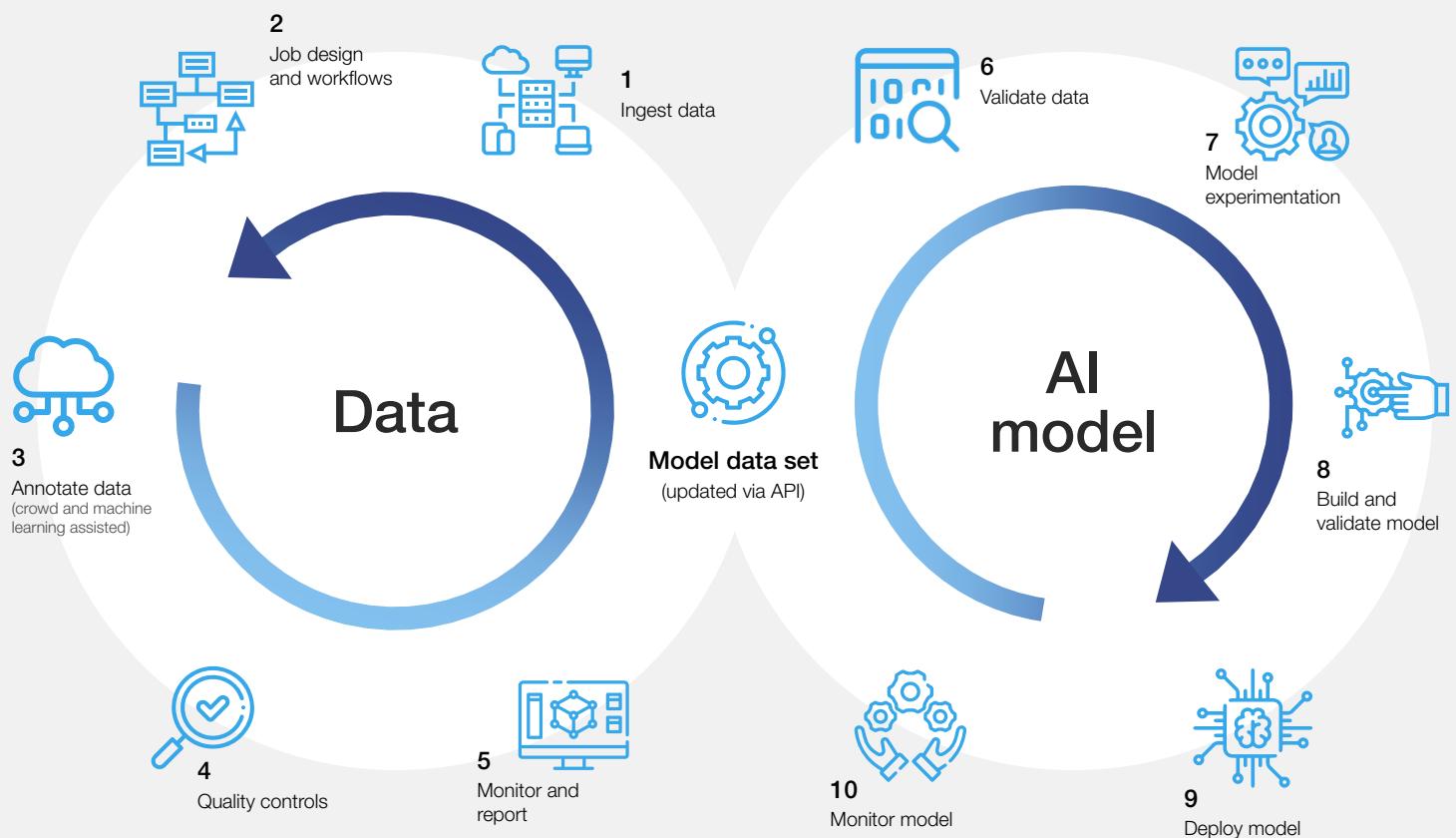
Data strategy – origination, acquisition, management, operations – is a key component in building an AI strategy. Assessing internal data availability is often a good place to ground this discussion. Understanding what data the business requires, and what data it generates, can provide very specific AI use cases but, with imagination, can also suggest new, potentially more entrepreneurial ones.

A strategic approach must be taken to data, both because its availability will define the scope of what is possible with AI, but it may be that investment in data acquisition could open up new opportunities.

Similarly, reliance on a third party could create risk dependencies while some internal data might offer new entrepreneurial opportunities, whether through the core business or a new venture arm. Investment in data cleansing, upgrade, curation, protection and legal oversight will be key, especially as this becomes an area of rapidly increased oversight (the EU's GDPR is likely to be the harbinger of more regulation to come).

AI model work is a double cycle, one that focuses on data and one that focuses on the model itself (Figure 5). Both are equally important and must be factored into the AI strategy:

FIGURE 5 | Data collection and AI model life cycle



Source: Appen, “Technology, Single Data Collection and Annotation Pipeline”, 2021, <https://appen.com/industries/technology> (accessed 11 November 2021)

2.6 How can you assess your organization's maturity on the AI journey?

Organizations need to walk before they run when it comes to AI. It is often the case that building the foundational capabilities for AI requires greater investment than rolling out AI use cases. The first step in a corporate AI journey is to understand the gap between your current capabilities and where you need to be. This means looking at questions such as:

1. Do you have a clear business plan for the use of AI?
2. Do you have a person assigned to lead AI in the organization?
3. How many AI prototype and production projects have you completed? How many different functions (e.g. HR, finance, marketing) have production AI projects?
4. Do you have the necessary technical, data science and management skills for AI?

5. How many data scientists, AI engineers, AI product and project managers do you have?

6. Do you know how to identify and govern the risks of AI with internal controls?

7. Do you have strong technology platforms and strong digital foundations for data access and AI model-building?

Organizations should typically start small as they develop their AI plans, building up capacity and knowledge. One option is to acquire expertise and skills, whether through outsourcing or acquisition. As always there is a trade-off between growing the organic skills that will be better positioned to exploit internal data and relevant use cases for long-term strategic advantage, and sharing what might well be mission-critical data and insight with a third party.

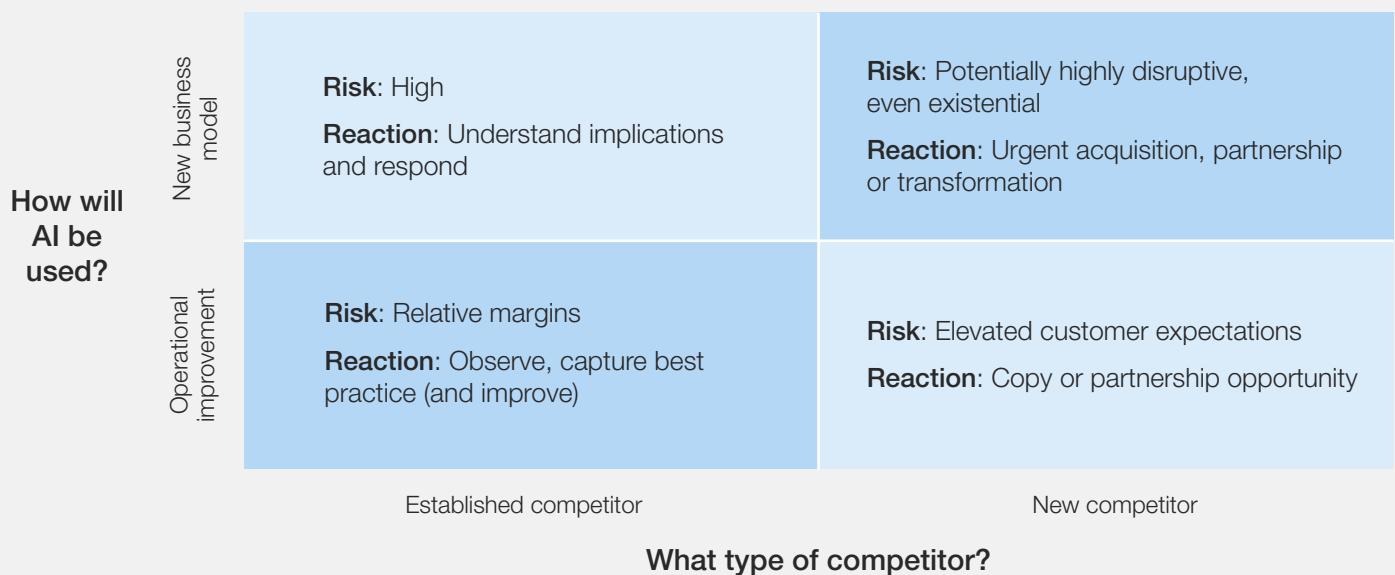
2.7 | How can you assess your competitor landscape when it comes to AI?

Typically, two questions should be considered:

1. Will the strategic impact of AI improve operations (e.g. demand prediction) or will it create a transformative new business model (e.g. on-demand pricing)?
2. Who is deploying AI? Is it one of your traditional peer competitors or is it a new player, whether a start-up or a new market entrant from an adjacent sector or geography?

FIGURE 6

Assessing the competitor landscape when considering AI



Source: Best Practice AI research

Assessing the competitor landscape (Figure 6) is especially complex in the field of AI. Both tactical operational tools and new business model development options exist, from both established players and new market entrants. Meanwhile, lessons can be learned from different industries that may be copied by key market players.

Research by McKinsey into the motivation of senior executives suggests that the most compelling rationale to get an AI strategy up and running is understanding what the competition is doing. The fear of falling behind is a very real one, and matters both in terms of direct business competition but also in perception terms with key stakeholders from investors to the media. Moreover, waiting, watching and only then responding is rarely the route to building leading-edge competitive advantage.

The use of AI can have increasing returns, so companies need to ensure that they are not left behind as a result of slow adoption compared to their peers. In addition, building AI competitive capacity can be a slow burn, requiring consistent investment in building human and data resources, leading to sudden breakthroughs. By the time a

competitor understands the situation in the field, the game may well have already changed.

However, given the range of techniques and use cases applicable across a firm, watching competitors is a useful tool for the active AI executive team. But unlike much competitive monitoring, this should not be limited to the immediate industry. Ideas and tools can be imported across industrial and functional boundaries and, similarly, cooperation and access to best practice may be easier with players in these areas.

They also need to be aware of companies in historically non-competitive but adjacent sectors that might be positioned with a nimble, lower cost and disruptive AI-first market proposition. For example, in China, Alipay has moved from providing e-commerce escrow services to being a leading financial services player with economics many orders more attractive than peers. The opportunity to leverage skills, data and consumer advantage across boundaries is a hugely disruptive threat that potentially comes from both established data platform giants as well as rapidly scaling start-ups that place AI at the heart of their propositions.

2.8 | What steps are involved in designing an AI roadmap?

FIGURE 7

Steps to design an AI roadmap



Source: World Economic Forum

Business leaders need to be a key part of the internal AI strategy debate (Figure 7) to help fashion what use case opportunities are key, and to help calculate the potential return on investment (ROI).

1 Articulate your vision for AI

To guide AI projects in your organization from conceptualization to delivery, it is important to have a clear business vision and to use it to establish a vision for AI. Translating your organization's vision to a vision for AI will help guide the many decisions that will have to be made. A vision statement communicates the big picture to any stakeholder

in the organization, articulating the value you hope to drive with AI. To build a vision statement for AI, a good understanding of the organization's current strategy, its business issues and priorities, and the broader industry trends, is paramount. These considerations should all feed into a vision statement that captures the organization's hopes for AI.



2

Define the business objectives that you would like AI to help you attain

With your vision in mind, think about the most pressing needs or urgent problems facing your organization. There are myriad business use cases for AI, so starting with your organization's business objectives is critical to ensure you'll invest in the right solution. Business objectives usually fall into one of the following categories:

- Improving customer experience (e.g. the personalization of services or faster service delivery)

- Improving employee experience (e.g. increasing productivity or on-the-job satisfaction)
- Optimizing resource allocation (e.g. driving efficiency in resource planning or distribution)
- Managing risk (e.g. improving disaster planning or detecting cyber-risk).

3

Identify potential use cases for AI that can help you attain these business objectives

Once you have defined business objectives, you can begin to look for AI use cases that will help your organization achieve those objectives. Good use cases for AI often involve tasks or activities that:

- Are repeated frequently and often involve large numbers of employees and costs, such as answering customer service queries

- Have fairly clear inputs and outputs that focus on prediction and optimization, such as predicting customer churn or forecasting supply chain demand
- Involve data sets that can be digitized and structured, such as reviewing unstructured CVs and resumes to help automate the screening of candidates.

4

Articulate and quantify the impact/business value of each potential use case

The business value side is often an earnings before interest and tax (EBIT) measure: directly, such as specific revenue or cost savings targets, or indirectly, such as Net Promoter Scores or customer satisfaction scores, or marketing efficiency. It is important to note that measures tied to revenue growth are often the most effective key performance indicators for changing hearts and minds in the organization. Cost savings can be equally powerful from an EBIT perspective but tend not to generate the same level of excitement that

is at the heart of cultural change, which may be an important part of the benefits that your organization is trying to achieve with a particular use case.

Other, non-financial metrics can also be used as proxies for measuring cultural change and the quality of decision-making across the organization. These include adoption and utilization rates of AI tools, internal customer satisfaction ratings of the AI team's outputs, and how quickly the AI team is able to launch new products and improve existing ones.

5 Articulate and quantify the ease of implementation and costs of each potential use case

- Identify the required skills, internally and externally, along with the support of appropriate stakeholders
- Assess and estimate the cost of the appropriate data capabilities to support the programme
- Assess and estimate the cost of the appropriate technology capabilities to support the programme.

It is important to note that calculating the ROI may prove less reliable than traditional IT analysis due to the risk of AI experiments failing but experienced operators can provide useful guidance. It is relatively straightforward to identify AI use cases but the cost side can often be high if the skills and technology and data infrastructure are not in place. The cost of the infrastructure investment needs to be amortized across many AI use cases.

6 Prioritize use cases according to ROI horizons and investment strategy

The uncertainty in AI projects is higher than in traditional IT and software projects in terms of both predicting benefits and costs.

The traditional method for prioritizing use cases is to plot impact (business value) against feasibility and costs and to focus on short-term investments that deliver value quickly at a low cost, while simultaneously investing in projects that are riskier and more expensive but also have the potential to yield more significant gains over time. However, thinking about investments in this way is overly simplistic in the context of AI use cases. Firstly, human beings in general tend to overestimate the benefits of a task they plan to undertake while underestimating the time and resources it will take to accomplish that task – what economists Daniel Kahneman and Amos Tversky call the “planning fallacy”. Secondly, the uncertainty in any AI project is even higher than in traditional IT or software projects, meaning that it is next to impossible to accurately predict, before a project commences, the benefits it will provide and the costs it will incur along the way. Thirdly, the reason so many AI projects fail is generally not technical, but organizational. It can, indeed, be challenging to achieve the buy-in and commitment from stakeholders across various departments that are necessary to delivering AI projects successfully.

So what does this complexity mean for prioritization?

- Make calculations about impact and costs, but be aware of the planning fallacy while you do. Where possible, try to make your predictions more reliable by using prototypes (also called “minimum viable products”) that you A/B test (using a randomized experiment with two variants), and/or by conducting surveys and user research along the way. These can help you establish realistic confidence intervals for your estimates that will establish more realistic expectations

and help you catch unprofitable projects earlier on. More often than not, an AI solution will not make a more manual process that much more accurate or reliable, so it is important to test often against your baseline and continually check your assumptions along the way.

- Bear in mind your stakeholder map. Are the people whose collaboration and buy-in are required really ready for this? Are they genuinely enthusiastic and excited, and will they put in the effort required to deliver results? Or are they simply nodding along because they have been told to, but possess little natural interest in the project? Prioritize those use cases for which stakeholders demonstrate true excitement. If those projects are successful, these same stakeholders can become far more powerful ambassadors for AI in your organization than the executive leadership team or data leader alone ever could. Peer recommendations often create a “pull” effect for AI solutions – a kind of grassroots demand – that is often absent when AI is merely “pushed down” from the executive level.
- Think not just about impact and cost but also time to value. How long will a project likely take to deliver? It is important to be able to have a roadmap that delivers results incrementally over time – smaller projects that can deliver quick wins at steady intervals – while simultaneously working on riskier but also more scalable projects. The latter are usually foundational investments in systems and processes around data quality and infrastructure, whose benefits are much harder to quantify but are necessary for doing AI at scale and moving beyond the proof of concept stage.

7

Incorporate the appropriate governance mechanisms for managing AI risks

Creating a set of ethical design principles that fits the needs of your organization should be part of the development of your AI vision and strategy (see [section 4.4](#) “Governance”). Ethical design is most effective at minimizing risk if it is embedded from the beginning of an organization’s AI journey instead of at the end. Ethical design principles should flow from your organization’s existing values and should also be influenced by best practices in the AI ethics field, as well as by the regulatory environment.

Consider the five areas of ethical focus outlined by IBM³⁵ as an example to your thinking on incorporating ethical design into your AI strategy:

1. Accountability: “Every person involved in the creation of AI at any step is accountable for considering the system’s impact in the world, as are the companies invested in its development”.³⁶
2. Value alignment: “Designers and developers collaborate with each other in order to ensure consideration of existing values. Care is required

to ensure sensitivity to a wide range of cultural norms and values”.³⁷

3. Explainability: “Transparency and data governance policies will ensure people understand how an AI system came to a given conclusion or recommendation”.³⁸
4. Fairness: “As humans are inherently vulnerable to biases, and are responsible for building AI, there are chances for human bias to be embedded in the systems we create. It is the role of a responsible team to minimize algorithmic bias through ongoing research and data collection which is representative of a diverse population”.³⁹
5. User data rights: “It is your team’s responsibility to keep users empowered with control over their interactions”.⁴⁰ Complying with applicable legislation, such as the EU GDPR, is an important part of ensuring your users’ power over the use and access of their data is preserved.

2.9 Who should be consulted in designing an AI roadmap?

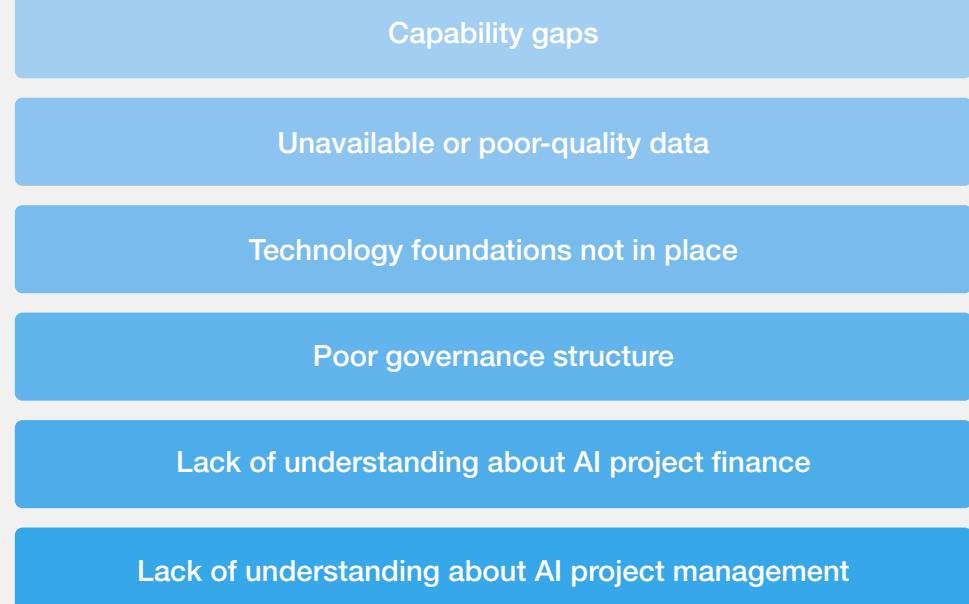
A successful AI strategy requires a mix of people in the room: senior leaders but also technical, data and domain experts. “Translators” that can bridge the languages of AI technology and the business are often the lynchpin. Critically, senior management and business line owners need to be able to provide support and guidance. Key functions, including

Risk, HR and IT, will be key stakeholders. It is also critical to ensure that the mix of people not only include some insight on those who might be affected (e.g. patient representation for a medical application) but also that not everyone in the room is of a similar profile. The latter is important to minimize the bias and ethical risks that may emerge.

2.10 | What are the main reasons that AI initiatives fail to deliver business value and how can you mitigate these risks?

FIGURE 8

Reasons AI initiatives fail



Source: World Economic Forum

The principal reasons AI initiatives fail to deliver business value (Figure 8) are the following:

Capability gaps

Capability gaps may emerge at various levels. Immediately, they might pertain to the data science capabilities inside, or available to, your organization. However, they will likely also include data engineering and operational management. Key processes may not be adequately mapped or understood, and over-hasty deployment may sideline the very internal domain experts required to ensure high-quality deployment (for example through eliminating middle management posts). More widely, a lack of education about the role of AI in an organization can lead to negative impacts on employee morale and even attempts to sabotage roll-out.

Unavailable or poor-quality data

Data is the key fuel for AI. Data can be volatile and with its increasing regulation, such as the European GDPR regulation that governs the corporate use of personal data, having a clear understanding of the data lineage, quality, statistical representativeness and use cases is necessary. Data accessibility is often a challenge for AI projects when the data is stored in multiple, disparate legacy data siloes with little linkage between the data sets limiting a big data picture. Extensive resources, often including manual labour, need to be invested in labelling data sets so they can

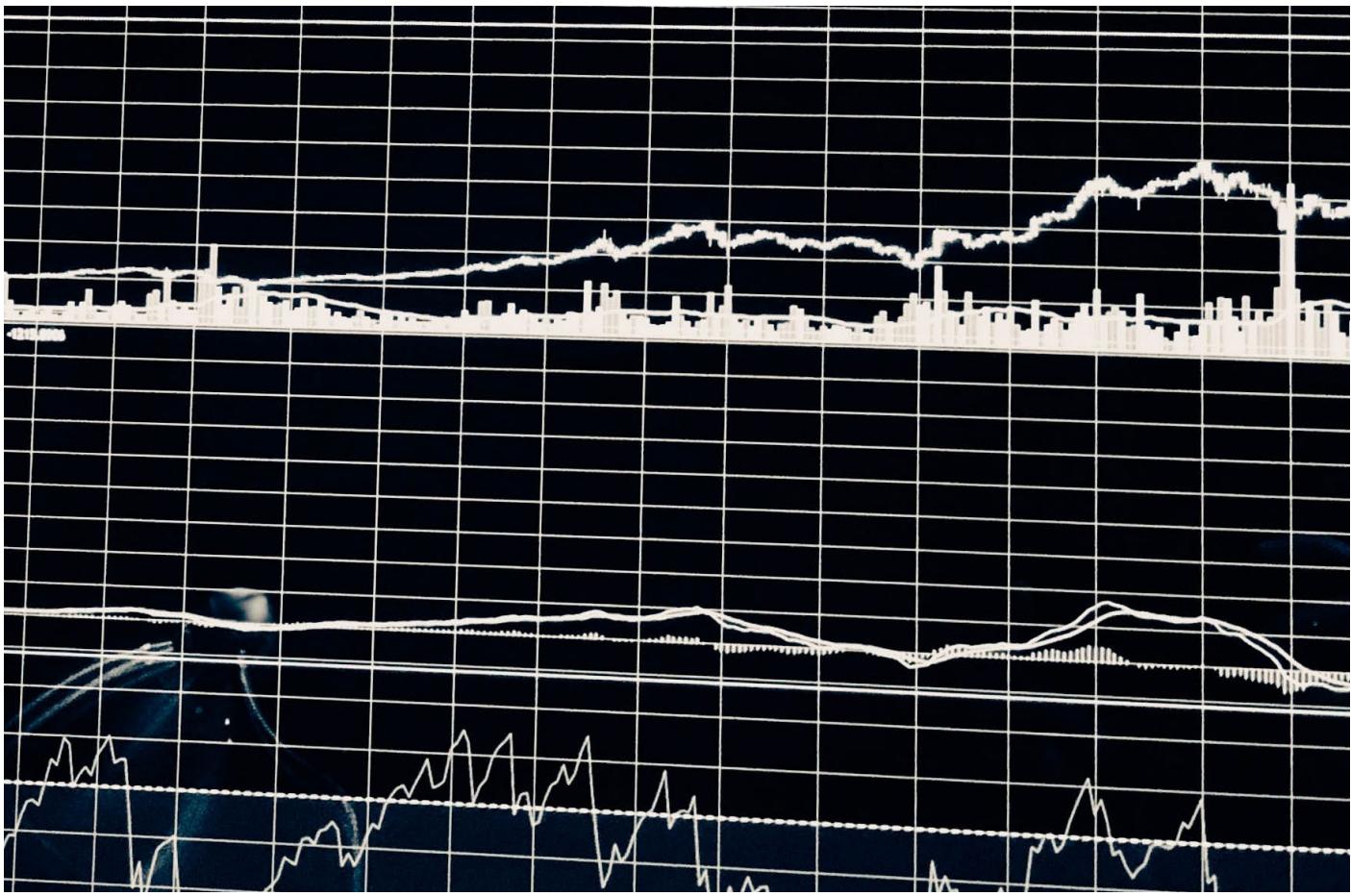
be used to train the AI models. It is not uncommon that 80% of the time needed to create an AI system is devoted to data cleansing and manipulation.

Technology foundations not in place

AI systems are dependent on having the necessary technical foundations. While many technology vendors are incorporating AI capabilities into their existing enterprise software offerings, ambitious companies are also digitizing and architecting their technology platforms, processes and tools to enable the rapid roll-out of AI use cases across functions. Typical foundations include data storage, access and cleansing capabilities, such as data lakes and data pipelines; the selection of ML and AI algorithmic frameworks; and access to best-in-class AI as a service (AlaaS) solutions. Not all organizations have these foundations in place and may be trying to run before they can walk.

Poor governance structure

Ensuring that a clear governance process with appropriate controls is in place from the beginning is critically important but often overlooked. This needs to cover both traditional project oversight roles (budgeting, stage-gating, delivery support, internal stakeholder management and critical questioning) but also a wider set of ethical and regulatory controls (the management of bias, data protection, “ethical issues”, external stakeholder



management) that will ultimately need board oversight. Organizations handle these challenges differently, although it is probably fair to say that many have yet to fully engage with the latter set of issues in an appropriate fashion. High-profile failings include Google's early attempts at an Ethics Board.

Lack of understanding about AI project finance

AI projects need ring-fenced funding to support capacity-building but also aligned incentives across the wider organization, and this is not always the case. For example, AI will typically piggyback on digital transformation investments in data lakes, cloud computing or process engineering. Removing funding from these to fund a "sexy" new AI initiative may well prove counterproductive. There is also usually relatively high uncertainty in AI projects at this stage as the scale of outcomes can be hard to guarantee. Helping the finance department understand this dynamic can be key to a project's corporate survival.

Lack of understanding about AI project management

AI projects often require skills in project management that may not exist in the organization. Internal project management capability may need to be upskilled to deliver AI projects. A strong focus on data, ethics and external stakeholder conversations may be novel. Bringing in teams from across the organization will be key. The use of traditional agile development and project management methodologies is appropriate, and iterative experimentation is critical. AI projects often resemble R&D rather than IT projects. The risks to performance and predictable delivery

timescales of AI projects is higher than traditional software development projects. AI is generally less deterministic and more experimental in nature – more research in R&D than development. Getting a customer chatbot or a manufacturing production line computer vision quality check system to an acceptable level of performance and robustness can be challenging and unpredictable, for example. Similarly, many AI projects have fallen foul of the challenges of data accessibility, quality and manipulation, resulting in AI projects overrunning by factors of 2-3 times in cost and time. They also need ongoing maintenance and support, as the risk of model "drift" or a changing environment (for example, changing consumer behaviour patterns driven by an epidemic) can undermine an algorithm's effectiveness.

Understanding the internal risk appetite will help define the best approach, especially in the context of what use cases to pursue and how to execute. For example, the ambition to reimagine the business as an AI-first organization (where the business proposition is redesigned around an AI-optimized workflow, enabling a radical rethinking of internal economics and organizational design) versus the roll-out of a new customer-facing chatbot is very different, despite both being relevant AI strategic options. However, the risk appetite will likely need to be at a higher level than for traditional projects. The AI journey is a learning journey, and learning is all about taking risks. Managers must have room to fail, as long as they can course-correct and learn. Shutting down all risk will likely lead to the firm falling behind in the competitive race.

People and organization

An organization's culture is critical to develop AI capabilities and embed AI in the environment.



By now, many companies have articulated an AI strategy and made significant investments in talent and tooling to make that strategy come to life (for more information on designing an AI strategy, prioritizing use cases and looking out for the business risks, see [module 2](#). “AI strategy”). However, for most, the return on these investments may continue to disappoint. Gartner predicts that, “through 2022, only 20% of analytic insights will deliver business outcomes”.⁴¹ A *Harvard Business Review* article states that the reason for this failure is not a lack of adequate resources but rather a lack of alignment and understanding up, down and across the organization.⁴² The problem of AI is therefore not primarily a technical, but a *human* one as well.

This module highlights some best practices to achieve this human alignment. This includes driving an AI culture in the organization; gaining awareness of the role of executive sponsorship and the skills these executives need to lead AI initiatives successfully; composing AI project teams and the skills that business stakeholders ideally possess; finding the right kind of data leader for your organization, including that person's ideal reporting line and strategies for aligning with their peers on the senior leadership team; and identifying the roles typically found in a data and AI function, the most important skills to hire for in a data and AI team, and the most effective organizational models to deploy these resources effectively across the organization.

3.1 | How do you drive a culture of AI in your organization?

Fostering a data-driven culture

In the early stages of business expansion, strategic or intuitive decisions and past experiences can be relied on to guide business direction without access to clear data. However, as a company scales up, transitioning to a data-driven culture in which evidence propels all decision-making is critical.⁴³

One of the key roadblocks to developing such a data-driven environment can be an organization's culture. Teams typically need to transition from more traditional hierarchical chains of command to a data-driven culture, an environment in which data and evidence are transparently used to make business decisions as opposed to more opaque top-down management. A data-driven culture therefore requires employees at all levels to be enabled and empowered, meaning not only the teams building and analysing the organization's data models, but also front-end employees who are involved in generating data.

To assess an entity's level of “data culture”, it is important to consider the following:

- **Data access:** How widely/effectively is data generated, collected and exchanged/shared (i.e. data supply) and to what extent is this data available to all employees?
- **Data use:** How widely is data used for analytics, AI and ML to drive business decisions?⁴⁴

Primarily, companies can seek to build a culture in which insights derived from data drive all business decisions by collating data from diverse departments and streams to create a single source of “truth”. Increasing data accessibility and streamlining the process help to create a central and holistic understanding of a company's consumers, supply chain, revenue streams, among others and, ultimately, foster a data-driven culture. When data is fragmented and kept in silos, not only is the time spent searching for data typically higher but the effort required to clean and manipulate data is frequently duplicated across different teams. Forming central data hubs can also help diminish the risk of “stale” data or low-quality data as data is continuously reviewed and updated.

It is equally essential to focus on increasing data literacy to transition towards a data-driven culture. Understanding what the data is communicating can be the biggest challenge for most companies when trying to tease out broader or strategic inferences from discrete data streams. Therefore, developing visualizations, dashboards and regular data reporting mechanisms across teams to help interpret data is necessary for a smooth transition to a data-driven organization.

Although fostering a data-driven culture is the essential first step in a company's overall digital transformation, it is critical for this to be followed by experimentation.

Fostering an experimental culture

In recent times, tech giants, including Amazon, Facebook, Google and Microsoft, as well as market leaders in traditional industries, such as FedEx and H&M, have demonstrated that online, live experiments can be critical not only in increasing the effectiveness of marketing but also in strengthening the overall innovation process.

Experimentation is indeed one of the key success factors distinguishing the competitiveness of market leaders. Within an organization's culture, greater emphasis can be placed on either only conducting what it knows will be successful experiments or not conducting experiments at all, due to a lack of risk taking. This can often discourage teams from trying new solutions and innovative fixes, and inversely drive them to focus on more traditional directions that avoid testing ideas that have the possibility of failure.⁴⁵

Typically, experimentation is used to identify the most effective touchpoints for consumers, design options and product recommendations. A/B testing, for example, is a common method for experimentation, whether it be for a new website design or marketing messaging.⁴⁶ When Amazon launched a new version of its Air Patriots game, a game for mobile devices, user retention dropped by 70% causing revenue to fall by 30%.⁴⁷ The development team compared the new launch to the older version and discovered that the game's difficulty had increased by approximately 10%.⁴⁸ The team interpreted this evidence to infer that the level of difficulty of a game can seriously affect the user retention and thus the revenue – a relationship

that had not previously been identified. To check this hypothesis, the development team tested four new levels of difficulty (along with a control version). The experiment proved that the easiest version of the game was the most popular. The team then launched a new version of the game that resulted in 20% increased revenue.⁴⁹ This is a strong example of how an accident was not just fixed but used as an opportunity for greater experimentation, as a team, in the light of new evidence.

Many experiments, however, can be conducted more quickly than the Amazon example and result in a more direct and immediate result. After an experiment involving a small headline change on Microsoft's Bing platform, revenue increased by 12%. As the head of the Analysis & Experimentation team at Microsoft insisted, companies should adopt an "experiment with everything" culture.⁵⁰

It is important to note that experimentation, similar to any form of R&D and innovation will not always lead to a direct gain in revenue. According to one review, "At Booking.com, only about 10% of experiments generate positive results ... a modification that attempts to improve something (sales, repeat usage, click-through rates, or the time users spend on the site, for example), performs better among randomly assigned users than 'A', the control, which is the status quo. ... But when you conduct a large volume of experiments, a low success rate still translates into a significant number of successes, which, in turn, diminishes the financial costs of the failures",⁵¹ and the perceived fear of failure.

Instilling an ethics culture

The importance of fostering a culture that emphasizes a data-driven strategy and experimentation is widely discussed and acknowledged, with many companies already demonstrating its direct impact on profitability. However, the significance of balancing these goals with practices that instil ethical AI cannot be overestimated. AI technologies clearly have considerable value and their potential impact for businesses across the board is immense. However, these technologies also present complex and unique challenges for companies and countries due to rising concerns around bias, stereotyping and discrimination.

In this context, UNESCO has initiated a global public consultation with "24 renowned specialists with multidisciplinary expertise on the ethics of artificial intelligence"⁵² in addition to "civil society

organizations, decision-makers, the general public, intergovernmental and non-governmental organizations, media representatives, the private sector, the scientific community and all other interested stakeholders" to form a Recommendation on the Ethics of AI. The recommendation is planned for submission to Member States for adoption in November 2021 and will be the "first global normative instrument to address the developments and applications of AI".⁵³ The purpose of the draft will be "to ensure that real-world biases are not replicated online, and offer concrete policy actions anchored in universal values and principles. It will also mandate UNESCO to analyze the level of advancement of each country in the field of AI in order to assist them in the implementation phase".⁵⁴

3.2 | How do you evolve the organization towards implementing AI?

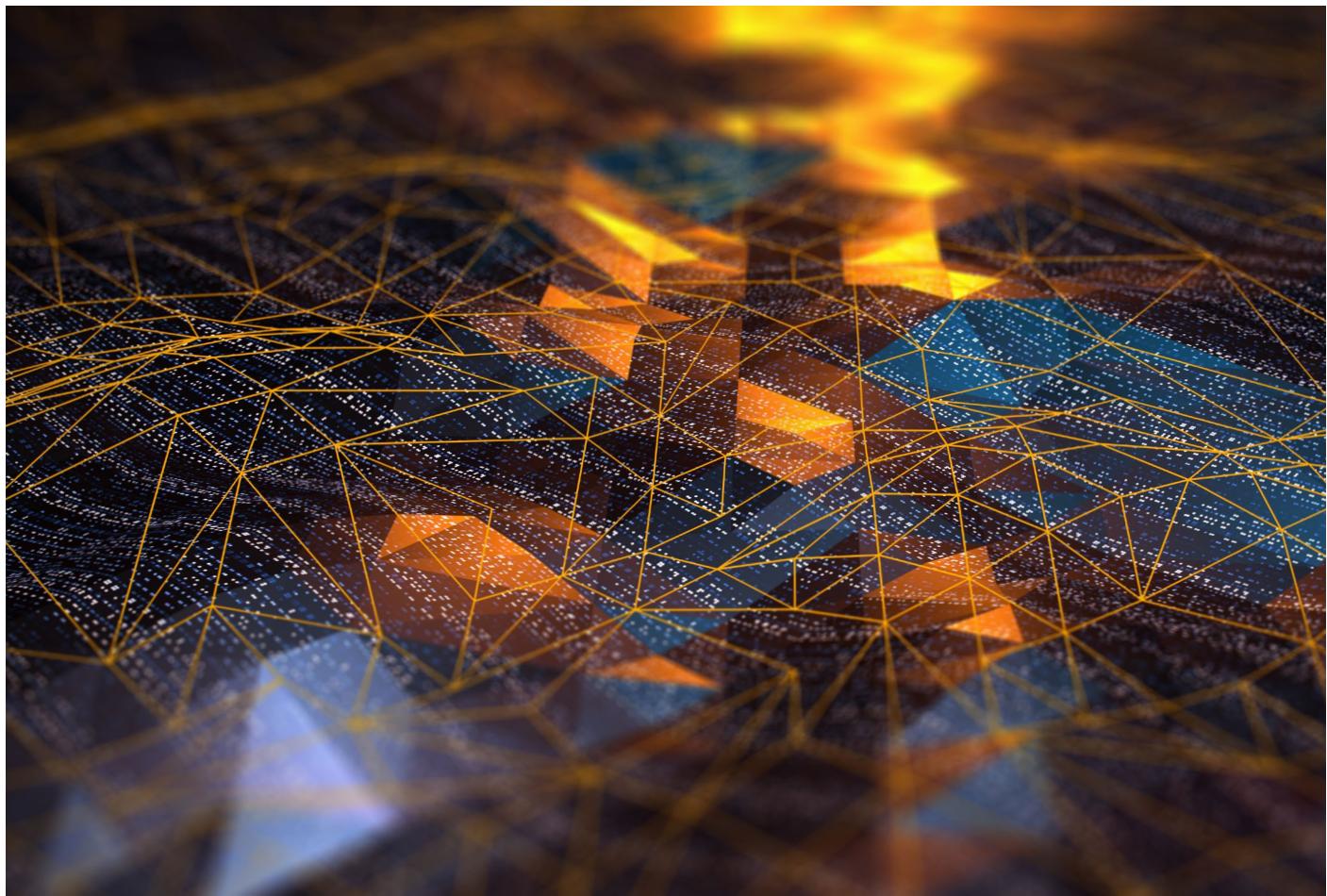
As an organization begins to take the relevant steps towards AI implementation, many aspects of the organizational structure need to be taken into consideration, including diversifying the team, hiring experts and ensuring that employees are aware of their new roles and responsibilities in the face of an AI-driven organization.

The transition into an AI-driven workplace is typically best achieved by following a phased approach. Such an approach includes, for example, the following:

- Identifying the stakeholder groups that will both benefit and be affected by AI adoption
- Providing stakeholders with a clear understanding of why AI adoption is necessary and how it will be of value to their role and the larger organizational direction. This awareness should not only be provided for employees utilizing AI on a daily basis but for other members of the organization. For full integration to be successful, all members of an organization (including the strategy, finance, procurement and operations department) need to understand the benefits of AI for their roles and the organization even if they are not directly affected by it

- Providing stakeholders with training and upskilling programmes to ensure that they can adopt the new applications
- Identifying the new roles that need to be filled, particularly by AI talent
- Adapting strategies and projects to use the new technology, and monitoring the outputs to mitigate risk and potential bias.

It is also key to note the need of multiple meetings with members of staff who may not see the value in AI adoption. The buy-in of staff and C-suite officials is key to the success of any emerging technology adoption. As part of the World Economic Forum's pilot of the AI Procurement Guidelines, the Centre for the Fourth Industrial Revolution United Arab Emirates worked closely with the Dubai Electricity and Water Authority (DEWA) to trial new procurement guidelines for acquiring AI. To ensure the new process was rolled out successfully, multiple meetings were held between the newly formed AI team (within DEWA) and the procurement and senior officials. Without their buy-in, the programme would not have worked.



Identifying roles

As an organization evolves with (or without) the adoption of technology, the focus on developing the team and the organizational structure becomes even more necessary. Many traditional “tasks”, such as data input and inventory, may become automated (by 2025, 85 million jobs may be displaced by a shift in the division of labour between humans and machines, while 97 million new roles may emerge that are more adapted to the new division of labour between humans, machines and algorithms⁵⁵⁾) and new jobs will appear. This means a focus on multidisciplinary skills will become even more prominent.

As mentioned in the data-driven culture section, this means that, according to research on the future of work, “the integration of early artificial intelligence tools is also causing organizations to become more collaborative and team-oriented, as opposed to the traditional top-down hierarchical structures”.⁵⁶ For example, teams such as strategy and IT, that may not have interacted directly previously, will have to become more collaborative as the strategic direction of the organization becomes more driven

by the IT department. In addition, new jobs will need to be created as traditional IT infrastructure developers and cybersecurity departments typically do not have the skills needed to develop or monitor AI algorithms. The development of these new collaborative teams will also require on-the-job training. For many AI-driven organizations, on-the-job training has become key to the success of an employee’s development. As technology adapts and changes, employees are no longer “experts” in their field longer than a few years. Employers cannot continuously replace these employees with new “experts” but instead must find a way to retrain and upskill them to adapt to the emerging technology (potentially using AI to tailor training programmes to the employees’ needs).

A good team structure, however, cannot be found quickly, especially if the organization is large. This is why it continues to be key to “scale” AI, instead of implementing it in bulk to begin with. It may be better “to start with a small team first, and let them evolve and scale up, rather than try to introduce the whole company all at once”.⁵⁷

3.3 What is the role of executive sponsorship in ensuring success with AI?

Executive sponsorship must be formally established and is critical in catalysing the vision of AI as AI adoption is not a siloed practice.

It is worth highlighting the importance of executive sponsorship and advocacy for AI within the organization. It is paramount that executive sponsorship be formally established in the form of an executive who is tasked with overseeing the company's AI and data agenda, such as a Chief Data Officer (CDO) or a Chief Analytics Officer (CAO), to ensure there are resources (e.g. funding, people, etc.), “protection”, organization-wide coordination and buy-in throughout the organization. Sometimes, depending on an organization's agenda, different leaders are asked to work closely on data (e.g. if the organization is planning or undergoing a transformation, the CDO/ CAO and the Chief Transformation Officer may tackle data issues together). Sponsorship and advocacy improve internal innovation, collaboration, leadership and branding.⁵⁸ Furthermore, as the adoption of new technologies is not a siloed practice but rather impacts various lines of business and stakeholders, it is important for organizational leaders to align employees to a strategic vision and encourage the adoption of new technologies.

Awareness and buy-in must exist broadly at the executive table beyond an identified project. Sponsors and the organization should be actively prepared to expand their AI practices based on their enterprise strategy. A study conducted by

McKinsey & Co indicated that organizations that continuously implement successful AI rate executive sponsorship twice as high as organizations that lag in AI adoption.⁵⁹ AI implementation is often driven by executive leadership, and members of the organization rely on this sponsorship for investments and strategic vision.

Operationalizing AI in the organization will require the involvement and commitment of many teams. As a leader, you will have a specific role to play in catalysing the vision for AI, identifying the opportunities and ensuring that the right people come to the table to achieve that vision.

Leaders should:

- Drive adoption and enthusiasm with other executives and members of the board by focusing on the big picture and value proposition and overcoming inertia, scepticism, fear of change and cost considerations
- Enact an effective communications plan that addresses the AI vision and goals for internal stakeholders, such as the business, operations and IT, and external stakeholders, like customers and vendors

- Secure the talent pipeline to plug gaps in AI teams and overcome hurdles in AI adoption; a skills gap in functional or business teams might delay AI implementation and integration.

In general, senior leadership needs to show active support, to set the direction, provide assistance

where required and, ultimately, own the outcome. As potential AI risk factors emerge (see [module 4](#), “Responsible AI”), it will be challenging for senior leaders to profess ignorance on the algorithm’s operations. Leadership understanding may prove one of the greatest economic motors in your company’s future.



3.4 What skills do you as an executive business leader need to lead AI initiatives successfully?

At a high level, the key skill is the ability to understand the art of the possible with AI, while identifying the main risks. To achieve that, one of the most important skills that executive leaders can develop to help set up their AI function for success is the ability to ask critical questions of the AI team. Ensuring that you are asking these questions sets your organization up for success because it helps the AI team focus on strategically important, commercially valuable problems rather than going down rabbit holes, which is surprisingly easy to do in an environment that is saturated with buzzwords and can be overhyped. So much of the

value in AI resides in choosing the right problems to solve and, while the AI team itself should also be skilled in challenging stakeholders and ensuring that a problem is set up correctly at the outset, business leaders possess a wealth of knowledge around the commercial levers of value in an organization that data and AI professionals will struggle to match.

A good AI team can use these prompts to refine these questions and discern the feasibility of producing an answer so the executive can then determine which opportunities are the most valuable to pursue.

 **The key skill is the ability to understand the art of the possible with AI while identifying the main risks it creates.**

A critical question is:

- Strategic: If it is not tied to the 3-4 big strategic bets of a company, it is only nice-to-have.
- Well-bounded: You should know what is in the scope of the question and what is not. For instance, if your organization is a large TV company, the question “What shows should I air this year?” would be far less effective than “Should I do more or fewer science fiction shows this year?” The more specific the question, the higher the likelihood that it can be broken down in such a way that it can be solved technically using AI.
- Actionable: A question that AI can answer is only of value if the answer is actionable. If it is not, it is not fundamentally worth the investment, no matter how interesting the information that emerges may be. For instance, the question about science fiction shows is only valuable if, after offering an answer using AI, a programming team can actually make changes to that year’s list of shows.
- Grounded in past findings: It may seem obvious, but it is important not to repeat work. Has a similar question already been answered in your organization? Are you about to attempt to reinvent the wheel when, for instance, the insights and analytics team has already done a lot of the data preparation and analytical groundwork to answer this question? It is important to complete that organizational research before making a new AI investment.
- Feasible: Feasibility is of course necessary but, given the complexity of AI, it may be difficult to assess the feasibility of an AI use case. For example, questions need to be answered about the availability and quality of necessary data;

whether similar use cases have been delivered in the past, either internally or externally; the maturity of the algorithms in terms of solving the particular technical aspects of the use case; and the potential costs as well as possible ways to de-risk the project, for example based on incremental proof of concept stages.

Executives that ask critical questions set themselves and their organizations up for a virtuous cycle, whereby the insights gleaned from data surface new strategic opportunities or challenge existing ones, which they can then use to revise or enrich their strategy. Executive leadership teams that are able to leverage data and AI effectively are those that can articulate a big-picture vision for their organization and stay open to reframing it according to what the data says. There is nothing fundamentally new about this: executive teams that welcome counterviews and believe that healthy debate is fundamental to creating a robust strategy have always been more likely to take optimal decisions over those that cherry-pick information in support of predetermined conclusions.

It also helpful if business leaders have some basic data literacy skills as well as an understanding of the AI life cycle – how an AI project goes from strategy and ideation to implementation – as well as the caveats along the way, as this will help establish more realistic expectations about what AI can and cannot accomplish, in which time frames, and the kinds of costs and feasibility challenges that are likely to surface along the way. Some technical understanding of AI is also useful, but not always necessary. For example, it is likely unnecessary to understand the different kinds of AI models that a data science team will explore to answer a certain question, although it may be relevant to ask how these different models could impact the investment case for a particular use case (see [section 2.8](#) “What steps are involved in designing an AI roadmap?”).

3.5

What opportunities, capabilities and risks of AI do business stakeholders need to understand to deliver AI initiatives successfully?

Business stakeholders can draw upon many of the skills required for digital transformation projects to deliver successful AI initiatives. However, there are unique opportunities, capabilities and risks of AI that staff and decision-makers need to understand to ensure the successful implementation of AI projects. These include:

- **Strategic opportunities and challenges.** AI can be used to create new and disruptive business models as well as drive revenue, cost and customer service improvements along with

deeper and more actionable insights. However, AI also presents unique strategic challenges. New competition can come from businesses in different sectors that build AI-first companies with radically improved economics. Similarly, direct competitors that invest heavily in AI can gain increasing returns from the technology, leaving laggards to AI adoption struggling to catch up.

- **Legal and ethical risk management.** AI projects can shine a spotlight on and amplify historic practices in an organization that could

lead to ethical, legal, financial and reputational risks. For example, automated hiring might reveal a historic hiring bias against certain protected groups, such as age or gender. Similarly, in some countries, such as the United Kingdom, regulation requires that fully automated decisions with legal or significant impact on individuals be explainable and that individuals be allowed to have a human review the algorithmic decision. Staff need to be educated to ask questions about whether projects have been assessed and remediated for AI ethical and legal risks.

- **Technical and data capability requirements.** Many organizations learn the hard way that getting the capabilities in place to do AI is much harder than doing the actual AI. The technical and data foundations required to implement an AI project are often more than 80% of the required effort. Staff need to be educated on what it really takes to build a successful AI project. They need to be comfortable asking questions about whether the organization has

the appropriate data accessible, quality assured and labelled; whether the underlying technology supports modern AI technologies; or whether the company knows how to scale AI from a proof of concept to production.

- **Project predictability.** AI projects have less predictable cost, timing and results than traditional software engineering projects, such as implementing a customer relationship management system or a mobile application. That is because AI projects look more like R&D projects, and space for experimentation and learning needs to be created for the projects to succeed. Moreover, AI projects may prove technically infeasible. AI projects might work well in 80% of cases, such as answering customer queries through automated chatbots, but it is increasingly difficult, if not unfeasible, for the last 20%. This is why it is important for companies to do proof of concepts and pilots in advance of committing to the timescales, costs and expected performance of AI solutions rolled out across an enterprise.

3.6 What kind of AI and data leader is the right fit for your organization?

The most effective data leaders all possess certain essential skills, in particular:

1. The ability to communicate with non-technical stakeholders at all levels of the organization; a strong storyteller and presenter who can simplify complex technical concepts and explain them in commercial terms
2. Sufficient technical understanding to manage their teams and systems effectively and grasp the art of the possible when considering the potential uses of data and AI in the organization
3. The ability to influence people at the top, bottom and across the organization
4. Basic business skills, such as an understanding of the budget cycle and how to present investment proposals.

How much weight an organization should give to each of these skills, however, depends on its strategic priorities and key business questions,

organizational culture, maturity on the data journey, industry and size. The best C-suite leaders will think meaningfully about what kind of profile they require, finding the right balance of skills according to the problems they need to solve. Will the data leader have a big piece of change management work to carry out at a large organization? A politically savvy influencer with industry experience is probably a good choice; that person can then hire a strong technical right hand if they lack advanced skills in this area. Will the data leader need to be very hands-on? Someone with a background in computer science or statistics – albeit one who is still a good communicator – is likely a better fit. A good rule of thumb for finding the right balance between soft skills and technical skills is how ambiguous the data leader's brief is. The greater the ambiguity, the more that leader will need to be able to think strategically, communicate and collaborate with heterogeneous stakeholders, and prove highly adaptable. Once value has been proven and it is a question of industrializing an existing process as efficiently as possible, someone with strong technical experience is often the best person for the job.



3.7 | Whom should the AI and data leader report to?

Having a C-suite executive who leads the AI transformation can be necessary for leading data and AI organizations. However, who the CDO/CAO reports to depends on the organization's business requirements. For example:

- If AI will be used primarily to reduce costs, the CDO/CAO would typically report to the Chief Financial Officer or the Chief Operating Officer.
- If data and AI will be used primarily to improve customer experience, the CDO/CAO would typically report to the Chief Marketing Officer.

- If the role of AI will fundamentally change the business model, the CDO/CAO would more likely report to the CEO directly.

These reporting lines establish and reflect the focus of data and AI for the business. Ideally, the data leader's manager will possess both strategic responsibilities for the business and have an appetite for quantitative thinking and curiosity for new technologies. It is this combination of responsibilities and skills that gives the data leader the best jumping-off point for building data and AI solutions that are both commercially viable and technically executable.

3.8 | Who are the main stakeholders and what are the roles in an AI project team?

There are two sets of stakeholders for building and using AI:

- Those required to deliver and manage AI projects
- Those whose trust is required to ensure the successful adoption of AI.

In the first instance, AI project and operational teams need to be built. As with many digital

projects, it is important to balance technology and business skills with domain knowledge. Broadly speaking, these roles include:

- Data engineers: responsible for creating the nuts and bolts of data infrastructure and pipelines
- Business intelligence, insight and analytics professionals: responsible for ad hoc reporting and dashboards

- Data scientists: responsible for building predictive algorithms and ensuring the recommendations of the data team are statistically robust
- ML engineers: responsible for scaling data science models and putting them into production
- Data product managers: responsible for coordinating all the above resources as well as aligning with business stakeholders and engineering to deliver value
- Data governance specialists: responsible for managing and documenting data assets and ensuring they are ethical and compliant.

Finally, depending on the desirable outcome, increasingly it is critical to have experienced product managers in the team to ensure the successful delivery of solutions that satisfies the end users' needs.

In the second instance, at the heart of successful AI deployment is trust, and the key to building trust is a transparent conversation with everyone who might be impacted. The capacity for issues to emerge – both real and through mischief created by a misunderstanding of AI – requires in-depth communication. For example, should issues emerge regarding bias in the data, having identified and started a conversation with representatives of the impacted group will translate into being better prepared to manage any emerging controversy. Key stakeholders that could be impacted and whose trust is vital to its successful adoption include employees, customers and consumers, suppliers and partners, and even wider society.

Against all of this, it is important for teams to be diverse and inclusive. All too often, algorithms are developed that are biased against ethnic groups as not enough representative voices were around the table to spot potential issues.

3.9 What are the most important skills to hire for in an AI team?

“Hiring talent that has both specialized expertise in AI and business acumen, and that can drive organizational change, is something to aspire to.”

Organizations must attract the required talent to ideate, design, develop and implement AI. This is not an easy task. In a 2018 study conducted by Deloitte, one-third of interviewed organizations (respondents) indicated a “lack of AI/cognitive skills” to be a primary concern, showcasing not only the importance of talent, but also the scarcity of AI resources.⁶⁰ As organizations seek talent such as AI researchers, AI software developers, data scientists and change management/transformation experts, they are developing partnerships with academic institutions and training academies to attract external members to the organization and develop skills within the current workforce.

AI is as much a science as an art, however, and thus the talent required for its successful application includes both technical and business leaders. As such, the organization must cultivate talent that is not only able to understand the technology but also able to identify practical use cases, mitigate potential risks, successfully implement projects, coordinate with multiple stakeholders, communicate value, manage products, and manage risks and drive organizational change, among others. The skill set needed is therefore quite broad, far beyond technical skills.

For instance, two of the most valuable skills in a data and AI team are pragmatism and a commitment to operational excellence. The world of AI is awash in buzzwords and hot new trends that may tempt data scientists and ML engineers to go down rabbit holes in a desire to satisfy their intellectual curiosity rather than because a project or technique stands to add business value. The data scientists most successful in business contexts are those who think about how to scale their algorithms in production from the

outset and those who genuinely care about whether their solutions get adopted by users. Similarly, the data and ML engineers who add the most business value are often those who care about time to value – those who are passionate about building faster, more efficient data pipelines and systems and are excited at the prospect of being invited to the table to discuss how to create business value rather than being relegated to a support function. Teams also need to include experienced product managers to ensure that the team receives continuous feedback from the end customers/users; the focus remains on business value creation; the end result successfully satisfies the needs of the internal or external customers; and AI projects do not just stop when a pilot is built but when a product is delivered successfully and business value is created.

And what about technical skills themselves? Might it be preferable to hire someone who is technically brilliant but has a lesser grasp of business strategy and communication? It can be, when business value has already been proven and the company is scaling or optimizing something that already works. A simple rule of thumb is that the greater the ambiguity of a situation or project, the more important it is to hire people with commercial and strategic understanding who have a good grasp of the technology and methods necessary to prototype solutions. In cases of less ambiguity, it can be an excellent time to hire a PhD with highly specialized technical skills who does not have to navigate high levels of organizational and process complexity to deliver results. Of course, hiring talent that has both specialized expertise in AI and business acumen and people skills is always something to aspire for. With the development of data and business literacy, this is becoming increasingly possible.

3.10 Should you hire, train or outsource these skills?

FIGURE 9

Before a detailed discussion on whether to hire, train or outsource individuals with specialist skills, such as data scientists or AI engineers, the business needs to understand whether it wants to buy or to build. The two key questions to answer are:

- What is the strategic value of the use of AI?
- Do suppliers offer specialized solutions for the AI use case?

Hiring, training or outsourcing specialist skills



Source: Best Practice AI research

In most cases, organizations will outsource AI skills in the form of buying off-the-shelf products that support common AI use cases, such as CV/resume scanning, customer service chatbots or marketing analytics. If the use case is not really an opportunity to create competitive differentiation but to provide operational improvement, it makes sense to outsource to suppliers that have invested and developed increasingly specialized AI software across hundreds of firms.

In other instances where the use case is not widely supported by suppliers or it is of key competitive advantage, it will make sense to build the application. This is especially true if the company believes that it has unique assets, such as differentiated data sets. There might be time to retrain existing staff, but it is still important to hire lead engineers, data scientists and researchers who have relevant experience. How far the build can take place purely with internal resources (hired or trained) will be a function of internal resource availability and access to a specialized external

resource, for example at a consultancy. An external firm with specialist knowledge or prior experience with relevant use cases or data sources may well provide considerable value and can often act as cultural accelerants to overall AI adoption in firms.

Similarly, the decision to hire, train or outsource (Figure 9) could also change over time. Many skills initially outsourced for reasons of speed to market or lack of internal technical capability will see capabilities brought in-house later if the use case is an area of strategic importance.

In all situations, companies should make sure that AI projects are led by an internal resource who is able to link technology and the commercial business.

Organizations should think about a programme to train all employees in the opportunities, challenges and risks of AI, especially the ethical and legal risks. Many companies have run training courses, often online, and some are even considering it at new employee induction.

3.11 How do you hire and retain AI talent?

Currently, AI talent is considered to be at a shortage on a global level. In the US, various labour markets have indicators that demonstrate a tight market for AI talent. This is evident in job site statistics when the demand for workers largely exceeds the supply of AI talent. Another indicator are the current high private-sector salaries for AI experts. Kaggle, a data science platform, published a user survey stating that the median salary for US respondents was \$110,000, whereas “top AI talent in corporate labs made several times that amount, even in entry-level jobs”.⁶¹ This illustrates the need for this type of talent but also the difficulty in acquiring this talent as an organization if the budget does not allow.

Moreover, a survey by the World Economic Forum in 2018 showed that “skill gaps” ranked among the top two barriers to technology adoption among firms, most of which reported a desire to invest in AI.⁶²

It is important to understand:

- Where to find AI talent
- How to retain AI talent
- Why it is important to build diversity and inclusion in the AI team.

3.11.1 Where to find AI talent

AI talent today is more interested in innovation and problem-solving. Monetary incentives can help to hire the workers but may not be enough to retain them.

AI teams are built of multidisciplinary experts and, increasingly, these skills are being found beyond the technology industry as well. A study looking into Asia-Pacific Economic Cooperation (APEC) economies, where AI talent is considered to be high, found that AI is becoming more prevalent across traditionally non-tech industries.⁶³ Apart from tech skills in AI, soft skills remain important to navigate. Programming skills are important, but many AI professionals have non-engineering backgrounds and bring critical soft skills to the workplace. Gender diversity numbers have been reported in various studies; they have not improved much for AI and show that the share of women in the AI workforce is still only 20-25%, which indicates interventions would be needed to avoid perpetuating the gender gap. Outside of the tech industry, AI professionals are heavily concentrated in education, manufacturing and finance.

A few points should be considered when hiring AI talent:

- Identify what role you are trying to fill and what expertise you are looking for: this will depend on the level of experience required and organizations should be able to adjust their search strategy.

- Understand where to find the skills you require: talent is scattered in different places and organizations should be able to identify where to look. For example, junior ML developers can be found in universities and hackathons while senior AI researchers might be found more easily through networking at an academic conference.
- Analyse candidates’ backgrounds and skills and do not limit yourself to resumes: it is important to look at candidates’ skills and experience, but it is equally important to look into their passion. Recruiters have found that as regards academic backgrounds, fluency in “mathematics and statistics often provide candidates with the ability to identify algorithms and optimize them for a specific problem or outcome. Personality-wise, curiosity and creative traits tend to work well with the abstract and unclear problems often facing AI and ML experts, and an ability to learn quickly pairs well with the dynamic environments of AI and ML”.⁶⁴
- Offer opportunities beyond salaries: to attract talent, organizations should keep in mind that salaries are not the only incentives for AI talent. “Today’s AI talent is full of curious minds who prioritize intellectual challenges and having an impact over monetary incentives”.⁶⁵

3.11.2 How to retain AI talent

Incentive structure

It is undeniable that attracting talent can be a challenging exercise for recruiters, but perhaps more challenging and equally important is retaining that talent. One-fifth (20%) of top executives reported “retention as their main concern”,⁶⁶ despite it often being over-looked in the hiring

process. Having monetary incentives could be successful in the short term to hire talent but it has lower success rates in the longer run. AI talent today is more interested in innovation and problem-solving. Organizations need to provide the “right combination of opportunities to contribute, experiment and improve”, making workers feel

"valued and responsible for their work, making them less likely to abandon a project".⁶⁷ This is also important for the organization, given that innovation is built from experimentation. To retain tech talent, it is important for them to realize their individual worth to the team. A few things that could help to accomplish this are:⁶⁸

- Include talent (AI or not) in important decisions
- Offer collaboration over supervision
- Work with them "to develop actionable and effective professional development plans, which create a setting that supports people, and their long-term goals".

3.11.3

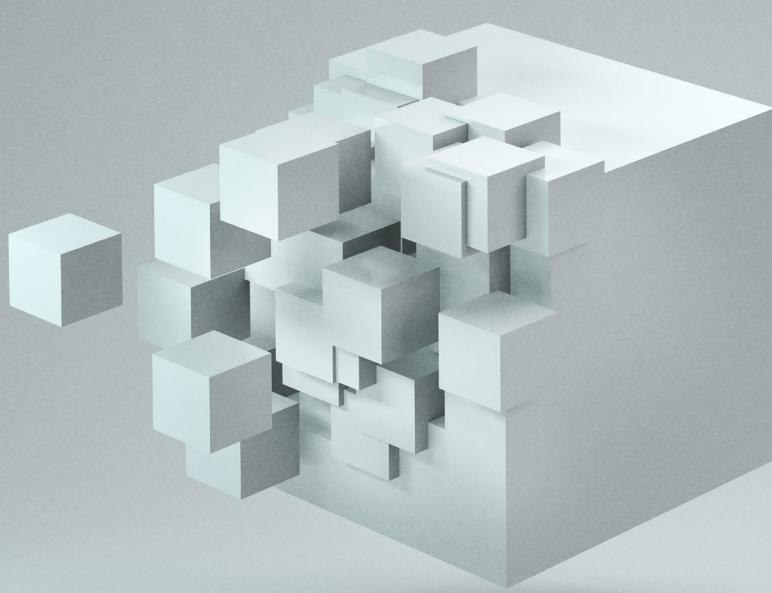
Why it is important to build diversity and inclusion in the AI team

Diversity and inclusion are two interconnected concepts that are not interchangeable.⁶⁹ Diversity is defined as the "representation or the make-up of an entity", while inclusion is about "how well the contributions, presence and perspectives of different groups of people are valued and integrated into an environment",⁷⁰ regardless of their differences. For example, a workplace with different groups of people where only the perspectives of certain groups are considered may be diverse, but it is not inclusive. An environment defined as both diverse and inclusive is "one that makes everyone, regardless of who they are or what they do for the business, feel equally involved in and supported in all areas of the workplace".⁷¹

Within an organization that looks to adopt AI, the diversity of the team (from a background and skills perspective) is key. The notion of "Western AI" has been criticized, with experts arguing that

AI is primarily developed in the West. This allows for larger bias and potential outputs that only take certain segments into account. For an organization to utilize AI properly and to benefit from its outputs, diversity within the team is key and yields several opportunities, including a larger data set variation.

As suggested in an article published by the World Economic Forum, "Diversity should be incorporated at every stage of the AI process, from design and deployment to questioning its impacts on decision-making. Research has shown that more diverse teams are more efficient at problem solving, regardless of cumulative IQ. Explicit attention to inclusivity in design, application and evaluation of the effects of AI-enabled decision-making will not only minimize inadvertent discriminatory effects, but can also lead to its design and application as a driving force for greater social, economic and political inclusion."⁷²



What is the most effective organizational design to embed AI and data-driven decision-making in your organization?

It is not enough to have the required talent; it is also important to design the appropriate means to organize the structure and roles of the organization's data and AI workforce. Leveraging data and AI requires organization-wide data-driven decision-making practices and culture, easy access to data and analytics by everyone, "data first" in decision-making processes, and broad involvement in continuous data and AI enabled business improvements and innovations.

While many organizations choose to outsource AI development to third parties, those with the highest AI ROI typically manage it in-house, as revealed in a 2020 industry survey by ESI ThoughtLab. Overperformers were found to be more successful at developing and acquiring talent, excel at training, have larger AI teams and know how to leverage their C-suite.⁷³

AI talent can be leveraged through a spectrum of models. The spectrum falls between:

1. A centralized model, where a single, enterprise-wide AI function controls AI resourcing, funding, assets, talent and AI strategy overall. Lines of business engage the centralized function for their data and AI needs and data governance requirements.
2. A fully segmented model, where data and AI activities are conducted entirely within business units. Each business unit is in charge of hiring the required talent and ensuring governance standards are maintained.

In practice, most organizations decide to adopt a model that falls in between the two extremes. Three example structures are:

- “Star”: This structure is best-suited for smaller companies just beginning their AI journey. It features a centralized AI team, usually the same as a Centre of Excellence, which manages AI efforts for the whole organization.
- “Matrix”: As a company matures, this structure may help scale and address a variety of business problems. It features cross-functional AI teams, with software engineers, data

engineers, data scientists, quality assurance specialists, DevOps engineers, product managers and designers, aligned around business investment areas. This structure can be versatile but can also grow complex quickly.

- “Fully-embedded”: This structure is best for the largest and most mature organizations. Each department has its own AI team with minimal reporting to a central unit. This type of team structure requires significant numbers of AI specialists.

The right organizational structure also depends on the firm's AI maturity, size, specialized skill sets, business knowledge and required focus by various lines of business and overall by the enterprise. Designing an organization to deliver AI is directly connected to how advanced the organization is in adopting and implementing AI. This decision also depends on internal AI resources availability and the overall size and structure of the technology functions.

Best practice typically involves a central cross-functional and multidisciplinary expert team working across the organization. A central enterprise function sets standards and common frameworks (e.g. leading practices, common enterprise tools, the oversight of enterprise-critical data). This central team typically needs a governance, education, ethics and oversight role, but whether it performs a delivery or deployment role will depend on the company's operating model. Business line data groups leverage these centralized enterprise assets by using them and extending them to meet business line's data needs.

The teams that own AI in an organization must have a comprehensive understanding of not just the technical skills but also of the business and project management. Despite the model being used by the organization, the AI function would be transversal and work with IT, and possibly external partners. One constant of success is a culture of working across the organization to ensure business and stakeholder buy-in. Good AI teams will be collaborative, focused on small wins, and have open communication and knowledge sharing. High performing AI teams will need top-line organizational goals, and reduced focus on functional goals, to enable them to concentrate on delivering ROI.

● The right organizational structure depends on the firm's AI maturity. One constant of success is a culture of working across the organization to ensure buy-in.

3.13 How can the AI team align with business, on the one hand, and engineering, on the other?

This alignment with business, on the one hand, and engineering, on the other, is one of the key aspects of effective organizational design when it comes to data and AI. In most organizations today, at least some of the analytical talent, such as analysts and insights professionals, will sit in business teams. Meanwhile, data engineering and ML engineers, who are responsible for scaling data scientists' models, often sit in the engineering organization. No matter the set-up, it is crucial for the data science team, whose members usually report directly to the data leader, to work seamlessly with their counterparts in both business and engineering. If that disconnection is too great and analysts, data scientists and data engineers are having trouble communicating, integrating their work with another and working on different sets of tools that integrate poorly with one another, it may be preferable, for a time, for analysts and data engineers to all be brought under the purview of the data leader, until they can establish strong communication lines, a streamlined way of working and a set of common standards and tools. However, a trade-off to centralization is often that

business teams are less likely to be an integral part of the ideation and creation process of AI models, which in turn makes them less likely to adopt them once they are launched. Similarly, the tools that a central data team creates are less likely to integrate smoothly with the rest of the organization's tech stack when its data and ML engineers are overly separated from the organization's larger engineering team. The most effective way of pushing these resources back into business and engineering teams without losing scalability and common working standards is to retain a dotted line to the data leader, who can remain involved in establishing roadmaps, hiring processes, the vision and career development.

Ideally, an AI team that is able to align closely with business and engineering operates as a flexible structure that can adapt to the changing needs of the organization as it matures on its data journey. It can be more centralized when it is necessary to prioritize common methodologies and vision, and more segmented when business alignment and tech integration are the main concerns.

3.14 What are the key considerations for C-suite executives when building an AI change management programme?

Change management is about identifying and managing stakeholders that can affect and be affected by the adoption of AI. While some may have little influence, others may be barriers to change or be greatly impacted by transformation. A poorly designed and implemented change management programme can change the goals of AI transformation from providing organizations with opportunities for competitive advantage to significant negative consequences, especially in the early days of the transformation process where resistance is high and clarity and unified messaging are key.

While there is some controversy within C-suites on the need to implement a formal change management system within organizations, there is a correlation that change management leads to better change outcomes.⁷⁴ As a result, a change management programme is key to the successful implementation and adoption of AI within an organization. While a

plan is often a step-by-step definition of what needs to take place in the short term, a programme looks at the long-term management of change throughout the AI journey, pushing small incremental changes while continuing to communicate what and how change will impact stakeholders as often as needed.

As regards the implementation of AI, change success is related to the extent to which an organization is ready for AI and the extent of which AI is adopted by various stakeholders with minimal disruption to business activities and/or revenue generation. In other words, **managing change in AI transformation is about identifying challenges in AI readiness and finding ways to overcome these challenges. The best C-suite leaders will think about their level of AI readiness and who their critical stakeholders are, and the challenges they may encounter particularly for employees and customers.**



AI readiness: Organization⁷⁵

The first steps in a change management programme involve an assessment of where the organization stands when it comes to AI readiness. According to research in this area, “An informed decision regarding an organization’s readiness increases the probability of successful AI adoption”,⁷⁶ from strategic alignment, available resources and the quality of underlying data and data flows, to the level of knowledge about AI across the organization and an organizational cultural assessment focused on attitudes. As an example, different perceptions among employees can either hinder or foster AI readiness and a change management programme would need to include an awareness of these attitudes along with initiatives to understand employee interests and fears and act upon them especially from executive leaders and managers. Managing attitudes towards AI is a balance of what the organization is and the journey of AI transformation, and having regular conversations about that.

Employees and AI adoption: Job automation and the future of work

A central part of success in implementing AI is aligning people with the change being made and, as mentioned in *The Future of Jobs Report 2020*,⁷⁷ the combination of a rising need for employment and parallel displacement of jobs because of the automation and adoption of AI raises job fears in employees. As already mentioned in the toolkit, trust and understanding are key to employee buy-in, and part of the change management programme is to have a clear vision for AI with employees at its heart because, essentially, job automation and AI are permanently changing the landscape of jobs and are associated with perceived risks to long-term job security.

Furthermore, Human Resource departments must have the right mindset regarding the implementation of AI. AI should work hand-in-hand with people to maximize their efficiency and performances overall. According to Jeff Schwartz of Deloitte Consulting, “It’s not about dividing work between people and AI – it’s not an either/or proposition – but about how to create new value in new ways.”⁷⁸ He offers the example of one of the earliest implementations of AI, the ATMs from the 1970s. “ATMs didn’t eliminate the jobs of bank tellers. We’ve actually seen an increase in the number of bank tellers and in bank branches since ATMs came on the scene.”⁷⁹ HR professionals must work towards reconfiguring a system that helps people and AI do things together better. Schwartz affirms that, “AI is good at performing tasks in large scale, working with speed and analysing massive amounts of data. People, on the other hand, are good with empathy, with communication, with flexibility in problem-solving.”⁸⁰

Primary research published in 2020⁸¹ is useful to help employees adapt to AI integration:

1. Be clear about your AI strategy and value.

It is essential that employees be receptive and adapt to the changes brought about by AI so that it can positively affect them. HR departments should listen to employees who may have concerns or fears on how systems will change.

2. Explain how AI can benefit their experience.

According to Anne Fulton, CEO of career experience platform Fuel50, “If deployed correctly, AI can serve as a journey accelerator for employees”. Repetitive tasks performed by AI will allow employees to do more empowering work and have a more rewarding experience on the job.

• An effective leader needs to recognize employees' concerns with AI and highlight the benefits of AI in a way that gives them an understanding of their own jobs and future.

3. Illustrate how AI can go beyond the elimination of manual processes to help employees. AI's future benefits will provide support in various ways, including distinguishing a change in an employee's state of mind and notifying a manager to suggest a longer lunch break. This is just one example of how AI can add value.

4. Develop key metrics to improve as a consequence of implementing the AI. Examples include the availability of upskilling training, and employee engagement and retention. Keeping track of progress on employee adaptation to AI integration on a monthly basis helps companies recognize their improved experience as a result of AI.

AI is complex; if change does not happen fairly, i.e. in an ethical manner, it will not be sustainable. In uncertain economic times, managing change is even more important and more challenging and an effective leader needs to recognize these concerns and highlight the benefits of AI in a way that gives employees an understanding of their own jobs and future.

Communicating the AI intent to employees and gaining buy-in

As mentioned, effective C-suite leaders know that articulating a clear vision for AI is critical to success. While this vision is used to align employees, it also helps align suppliers and customers to the transformation the organization is embarking upon. A vision statement communicates the big picture to stakeholders within and external to the organization. This vision statement is meant to articulate the value the company hopes to achieve with AI and is a combination of the organization's current strategy and industry trends.

Part of sharing a clear vision is telling a compelling story that inspires a path forward for an organization. McKinsey & Company⁸² has built an example for organizations to use during transformational change. From an AI perspective, this compelling story could include:

- A celebration of success that communicates what has been going well and the respective leadership position that an organization currently has
- Recognition of the need to adopt AI for competitiveness, efficiency or other transformational goals
- A clear roadmap for transformation, outlining the goals that the organization hopes to achieve and the associated benefits of the change from profitability to new capabilities
- A description of the initiatives to achieve these goals and how people would directly benefit from these initiatives, whether in education, behaviours, cultural improvement, etc.

- An understanding of the fears and challenges that employees may face and an explanation of how the organization recognizes them and how it will continuously communicate incremental changes and open the door to suggestions and improvements from employees.

Communicating the AI intent to the customers

AI is about a new operating model and, as part of AI readiness, customers should be made aware of the changes they should expect to see as a result of the transformation and adoption of AI. Customers are critical to success and acceptance to using AI-integrated offerings,⁸³ as seen during the COVID-19 pandemic and the associated digital transformation of e-commerce and other businesses.⁸⁴

The key challenges regarding customer adoption of AI-enabled offerings are AI's complexity and the perceived lack of transparency and concerns about privacy and data protection.⁸⁵ Customers who are not informed of the changes and who lack the knowledge may become frustrated, resulting in possible business disruption and long-term negative impacts on profitability. As a result, the best C-suite leaders know how to prepare customers for change through a clear vision of what will be implemented, how it will be used and what the associated benefits from personalization will be, and through automated and continuous customer assistance and service. Leaders, however, must also recognize the negative perceptions and implications of some AI-enabled processes as regards privacy, and implement processes that safeguard privacy and continuously communicate with customers and that are similar to the processes for employees, providing empathy and understanding.

Creating scalable practices requires starting with small, quick wins and building momentum. When redesigning and restructuring the organization, any form of digital transformation should happen in phases. Results based on World Economic Forum surveys for the blockchain and AI portfolios of the Centre for the Fourth Industrial Revolution United Arab Emirates showed that the public and private sectors found that when a pilot was initially tested within an organization and was successful, the scaling and adoption rate improved. This was primarily due to budget and change management concerns. Implementing new projects with emerging technology is costly and requires new resources to be put in place or old resources to be redistributed. Once the pilots have ended, as in the experimentation phase mentioned in this module, "those early adopters can encourage people in other departments to accept the technology with open minds and a positive outlook. For example, if the IT team is enthusiastic about an AI-powered virtual assistant, executives in the C-suite might be more receptive to the idea of allowing the organization to pilot the product."⁸⁶

Responsible AI

The responsible AI imperative is to move from principles to practice.



4.1 Introduction

“

That AI will have a major impact on society is no longer in question. Current debate turns instead on how far this impact will be positive or negative, for whom, in which ways, in which places, and on what timescale.

Floridi, Luciano, et al., “AI4People – An Ethical Framework for a Good AI Society”, *Minds & Machines*, vol. 28, 2018

4.1.1 How do you think about this new space to ensure your organization leverages AI to create value in a sustainable and ethical way?

The ability of AI to make decisions that influence its environment, which includes decisions affecting the lives of individuals or large groups of people, raises new risks for business and society in addition to ethical dilemmas that must be resolved. Beyond new risks stemming from the decision-making capabilities of AI, other issues can also be magnified due to the typically large scale of AI deployment. For example, as AI is embedded in physical products, such as automotive or manufacturing machinery, the concern for safety needs to be expanded. The close link between AI and data, the latter being necessary for the former, makes privacy considerations critical. Business

and society need to consider a number of ethical trade-offs, for example between privacy and safety or between short-term profits and fairness. In addition, the impact of AI on society can take a long time to materialize and may be difficult to foresee. For example, online recommender systems for products, content or news have existed since the mid-90s, but what is recent is the realization of the potential impact of these technologies on, for instance, how the political, socio-economic or even health-related views of people (e.g. attitudes to vaccination) may be formed, how views might be polarized, how disinformation is spread or how “echo chambers” are created.

4.1.2 Why do you need to manage the ethical implications of the AI revolution?

AI is a high-reward, high-risk new paradigm that requires revisiting our human and business values while mastering the rapid expansion of machine intelligence that will doubtlessly define the 21st century. Understanding and managing AI's potential risks and resolving its related ethical dilemmas while considering the relevant trade-offs are critical not only because they are the right things to do or because regulators will increasingly demand them, but also because business stakeholders – employees, customers, investors – will increasingly expect them. AI risks and ethical missteps can lead not only to regulatory fines but to reputational risks and the loss of revenues and markets. The new generation of customers and citizens is much more socially and environmentally aware⁸⁷ and digital- and data-savvy, demanding truly trustworthy organizations throughout the public, private and social sectors. In addition to the new norm of ethical supply chains and consumption, organizations are expected to demonstrate a higher degree of transparency in their decision-making

and actions, greater fairness in the way they treat their direct and indirect stakeholders, and more responsibility in and accountability for everything they do, including the way they develop and use technologies like AI. The public scrutiny of consumers puts pressure on regulators and lawmakers to act faster, raising the bar, and their response is strong.⁸⁸ It is a mistake to think the alternative is corporate leadership vs consumer preferences; democratic opinion and will are also factors. This is what ultimately confers a social licence on corporate activity.

It is therefore necessary to find a balance between fostering AI development, adoption and use while limiting and avoiding the possible underuse, misuse and abuse of these applications.⁸⁹ Since most of the risks associated with AI⁹⁰ have ethical implications,⁹¹ clear guidance for practitioners and decision-makers as well as the whole of society is required to illustrate how to deal with ethics in the context of AI and to suggest recommended actions and practice.

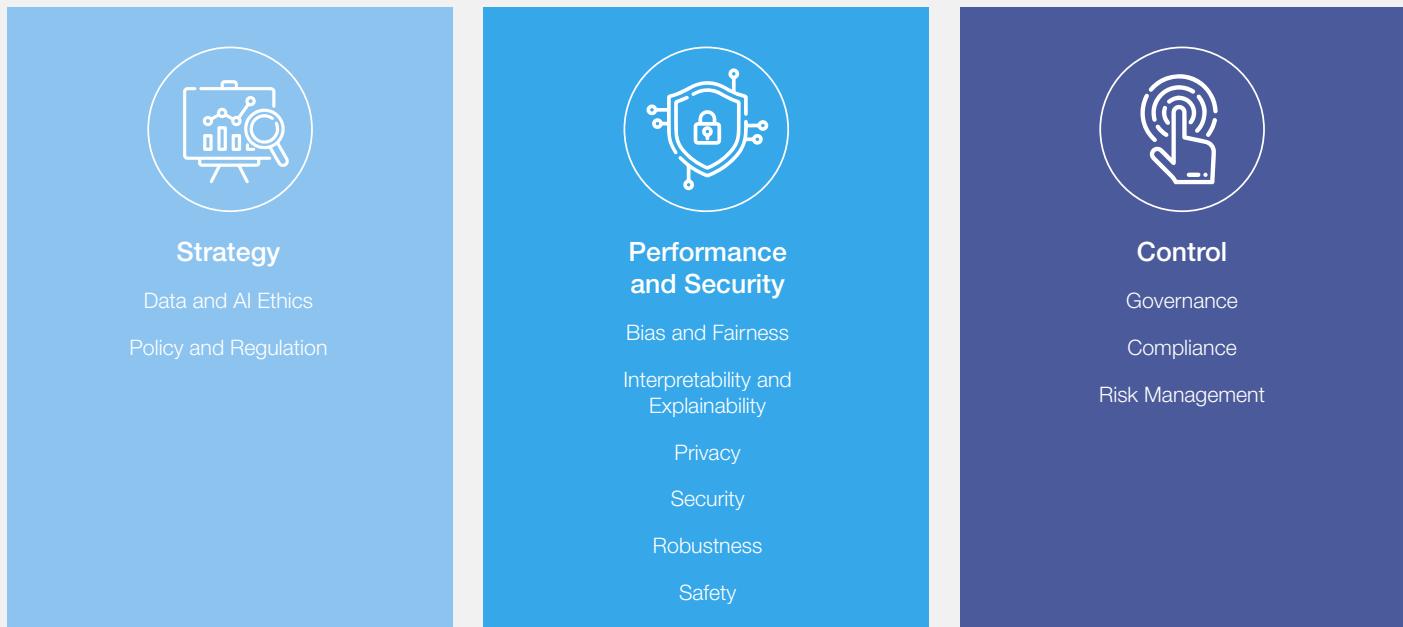
4.1.3 | What is responsible AI?

With AI poised to be profoundly transformational, ultimately, the vision of AI ethics needs to be of a good life for individuals and societies with AI, in terms of quality of life, preservation of the planet, human autonomy and the freedom necessary for a democratic and thriving society. But achieving this vision requires a practical and systemic approach, consisting of experts, tools,

methodologies and frameworks. **Responsible AI** as a multidisciplinary area of expertise seeks to elaborate how to translate the vision of ethical AI into ethical decisions and actions aimed at guiding the development and use of AI (e.g. Figure 10) in terms of embedding the right values across the AI life cycle while remaining in control and monitoring the performance of the AI applications.

FIGURE 10

PwC Responsible AI Toolkit



Source: Based on PwC's Responsible AI Toolkit in PwC, "PwC's Responsible AI", <https://www.pwc.com/gx/en/issues/data-and-analytics/artificial-intelligence/what-is-responsible-ai.html> (accessed 11 November 2021)

4.2 | Leading with ethical AI: A strategic imperative

4.2.1 | What is the ethics of AI?

Today's leaders need a moral compass and tools to help them navigate the complexity of the emerging moral dilemmas posed by powerful technologies like AI. That compass is the ethics of AI (Figure 11), the discipline concerned with applying ethical thinking (what is morally permissible, desirable and required) to all practical concerns raised by the design, development, implementation and use of AI.⁹² AI ethics is a golden path to realizing the benefits of AI (the Good) and mitigating the risks (the Ugly). It is the great maximizer-in-balance of AI. But ethics is not only about what can go wrong, about the

harms and risks it can trigger, but is perhaps most importantly about how to realize the potential of humans and society with the help of technology.

For the ethics of AI to deliver on that imperative, it has to follow the enterprise's pathway of AI development and adoption: think strategically, long term and holistically, and also act operationally, robustly and in an agile manner. As regards AI, ethics needs to touch all elements of its life cycle, from data to algorithm, infrastructure, interfaces and integration to governance and people.

FIGURE 11 | What ethics of AI is and is not

What the ethics of AI is	What the ethics of AI is not
<ul style="list-style-type: none"> – Doing good work that produces good and responsible AI – Maximizing benefits and eliminating or minimizing harms – Fairly distributing benefits and burdens – Understanding the impact and implication of AI systems – Challenging the status quo and checking the exercise of power – Bringing and including diverse perspectives – Navigating dilemmas and trade-offs 	<ul style="list-style-type: none"> – Only compliance – A set of fixed rules to follow – A purely negative frame, more than risks – A one-off event (“set it and forget it”) – A technical “bugfix”

Source: PwC adapted from Markkula Center for Applied Ethics, “Overview of Ethics in Tech Practice”, 22 June 2018, <https://www.scu.edu/ethics-in-technology-practice/overview-of-ethics-in-tech-practice>, and “Ethics in Tech Practice”, <https://www.scu.edu/ethics-in-technology-practice/sample-workshop-slides> (accessed 17 November 2021)

But what happens if the ethics are ignored or if insufficient attention is paid to it? The ethical failures of AI have been widely discussed in the media in recent years,⁹³ related to biased recruitment tools, racist bots and predictive policing algorithms, and mass surveillance. And while there is no consensus on the remediation approaches for future use cases,

the examples demonstrate the potential harm to individuals, organizations and societies⁹⁴ that must be considered in developing and using AI at all levels. And just as crisis response, cybersecurity and data privacy have moved from business units to boardrooms, the importance of responsible AI should be elevated to the top echelon of organizations.

4.2.2

What can AI ethics leverage from other fields of applied ethics?

At the core of a good business lies a strong ethical foundation. The value stance of the organization can drive its success. Good business is an ethical business. Successful businesses have understood that they must consider the dynamic nature of ethics in their strategic planning as well as in their ordinary operations. Both business ethics and bioethics can help inform AI ethics (Figure 12). Business ethics, as a study of situations, activities as well as decisions and issues of right and wrong, have had a profound transformative impact on businesses large and small. And with environmental, social and governance (ESG) issues, risks and opportunities now put at the core of business ethics, they are finally being recognized as “critically important to organization’s strategy, resilience and long-term viability – whether

it is a business, a non-profit, an educational institution, government entity or even a nation.”⁹⁵

Elsewhere, applied ethics have also had a profound positive impact, for instance in medicine where the bioethics movement arose and expanded as a result of advances that triggered major ethical dilemmas.⁹⁶ The significant questions about the ends and purposes of the life sciences and healthcare, cloning, pain and suffering, and rights and responsibilities are being addressed in a multidisciplinary manner. According to a historical overview, “philosophically trained ‘bioethicists’ could uniquely contribute to ethics in medicine. The growing momentum of the consumer rights movement lent its particular, rights-oriented contours to bioethics and the demands it placed on physicians.”⁹⁷

FIGURE 12 | What AI ethics can leverage from business ethics and bioethics



Business ethics – Various operational components, like processes, structures, role responsibilities, policies, governance and culture, for the AI ethics management ecosystem



Bioethics – The capacity to address and solve the big moral dilemmas resulting from scientific progress in a multidisciplinary approach (partnership between philosophers and physicians)

Source: World Economic Forum

4.2.3

What are the various perspectives on AI ethics today?

Ethics and the human rights law: Taking into consideration potential individual and societal harm with the possibly pervasive and disruptive nature of AI, it is necessary to tie the ethics of AI to specific human rights (Figure 13) to limit regulatory ambiguity, but also for the development of human-centric AI for the common good.⁹⁹ This approach aligns with the European Commission's Ethics Guidelines for

risks and unintended consequences to human and non-human life on earth and societies.⁹⁸ Leaders must not only manage this disruption but also the ethical and moral implications that emerge from it.

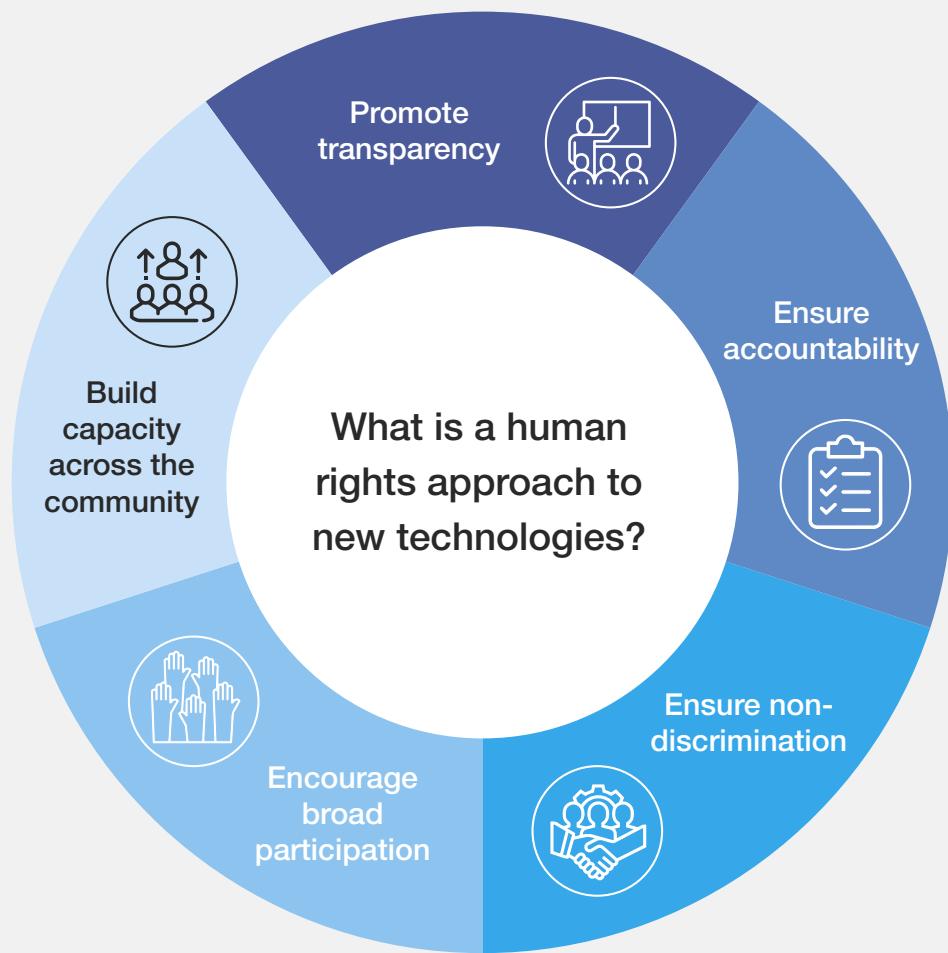
Trustworthy AI¹⁰⁰ and the OECD's Principles on AI.¹⁰¹ Ethical principles, laws and human rights are often inextricable, working in congruence across their spheres of application, and the way in which they apply to AI is no different. Applying this in the right context clarifies how technologies may give rise to different ethical concerns and considerations in the development and application of AI.



Fundamentally, it is the law – not an individual's ethical or moral compass – that governs how we develop and use AI. But a good company goes beyond the law's bare minimum. Ethical frameworks for AI can fill the law's interstices, thereby helping to uphold the spirit of the law. The best ethical frameworks provide practical guidance to those responsible for developing and deploying AI systems, assisting these people to make decisions that will protect fairness, equality and human rights.

Edward Santow, Human Rights Commissioner,
Australian Human Rights Commission

FIGURE 13 | Human rights approach to emerging technologies



Source: Adapted from Australian Human Rights Commission, “Human Rights and Technology”, Discussion paper, 2019, p. 26, https://tech.humanrights.gov.au/sites/default/files/2019-12/TechRights_2019_DiscussionPaper.pdf (accessed 11 November 2021)

Ethics and risk management: Managing and reducing risk has for years been a key component of business ethics management. It has demonstrated its efficiency and efficacy in addressing reputational, financial and regulatory risk, with clear operational and leadership support. AI ethics would benefit from taking a similar approach. As most new risks associated with AI have ethical implications,¹⁰² clear guidance on analysis and mitigation should be shared throughout the C-suite, AI practitioners and society, illustrating ways to deal with ethics in the context of AI and recommending actions and practices that would ultimately mitigate existing risks and potential new ones. As a result, these solutions can be considered ethically defensible and morally responsible but not morally accountable.

Ethics and the law and compliance: In the context of ethics, laws and regulation represent the public standards that are morally binding on all citizens of

a certain jurisdiction and enforced by institutional mechanisms, including punishments such as fines and imprisonment.¹⁰³ For the development and use of AI systems, laws and regulation are one of the layers of governance, alongside social morality and self-regulatory approaches (Figure 14) at the organizational level via codes of practice, compliance and risk management, industry standards and certifications.¹⁰⁴ While there is general consensus that current laws and regulations are not yet ripe to address the particular concerns associated with AI, regulators worldwide are making important progress to create robust and practical solutions to regulate AI. For example, the EU institutions are pushing for a more hands-on approach to regulating AI, with targeted changes to the European liability framework. A very risk-centric, AI-specific regulatory framework (e.g. the use of “high-risk” technologies and applications) is coming together.¹⁰⁵

FIGURE 14 | 3 layers of ethical governance – 3 interconnected levels



Laws and regulation - Public standards that purport to be morally binding on all citizens in virtue of their formal enactment and which are standardly backed up by institutional enforcement mechanisms including, at the limit, punishments such as fines and imprisonment



Social morality - Not all of the socially entrenched standards that properly govern our lives are, or should be, legal standards



Individuals and associations (e.g. businesses, universities, professional bodies, etc.) - Whatever social modes of regulation exist on these matters, individuals and associations will still need to exercise their own moral judgement

Source: Tasioulas, John, "First Steps Towards an Ethics of Robots and Artificial Intelligence", *Journal of Practical Ethics*, vol. 7, no. 1, 2019, <http://www.jpe.ox.ac.uk/papers/first-steps-towards-an-ethics-of-robots-and-artificial-intelligence> (accessed 4 November 2021)

The ethics of AI is not only about banning, restricting or constraining scientific development. In fact, the ethics of AI helps to develop a vision of a good humanity and, as computer science researcher Virginia Dignum says, humans should aim for the “development of intelligent systems along fundamental human principles and values, to ensure human flourishing and well-being in a sustainable world”.¹⁰⁶ “For that, all AI needs to

become responsible AI.”¹⁰⁷ The past successes and pitfalls of business ethics and bioethics demonstrate that “instead of trying to reinvent ethics, or adopt ethical guidelines in isolation, it is incumbent upon us to recognize the need for broadly ethical organizations. These will be the only entrants in a position to build truly ethical AI . . . Until this does occur, we can look forward to many future AI scandals and failures.”¹⁰⁸



4.3 Risks



The rapid advances in AI are creating enormous opportunities to improve our daily lives and create economic opportunities. However, the risk of abuse or misuse of AI requires a human-centred and principled approach to ensure that AI is used responsibly. Our employees know that building and using AI responsibly results in better products and outcomes for our customers. We ask hard questions to work through the new and complex ethical challenges AI brings and we are not afraid to draw red lines in premature or inappropriate use cases.

Brad Smith, President, Microsoft Corporation

4.3.1 What makes AI risky?

A number of AI system characteristics can lead to new types of risks.¹⁰⁹ Some of them include:

- **Continuous self-learning and change:** Unlike other technologies, AI systems can continuously evolve based on data provided to them, for example during usage. This is similar to how auto-correction evolves over time as one uses one's smartphone. Self-learning can lead to improved accuracy and quality of the decisions the AI system makes, but can also lead to behaviours that were not expected.¹¹⁰ In a sense, AI systems are like living organisms whose behaviour can change over time based on the experiences (the data, in the case of AI) they have.
- **Potential changes in the context:** Not only AI systems themselves may change over time, but so can the context within which they operate. When this happens, the system's performance (e.g. accuracy of predictions or quality of decisions) can deteriorate, possibly in unexpected ways.
- **Imperfect accuracy:** In almost all cases, there is no guarantee that the predictions, recommendations or decisions an AI system makes are accurate. It is almost certain mistakes will happen.
- **Deployed at scale:** AI systems making decisions can be deployed at scale, possibly affecting millions of stakeholders. Any inaccuracies can translate into a large number of mistakes.
- **Challenging transparency and accountability:** For many AI systems, so-called "black box" AI, it is not possible for humans to comprehend how they make predictions or decisions. Moreover, because these systems

AI system characteristics can manifest in various forms of risks when deployed at scale in different contexts.

can consist of multiple components possibly developed by various organizations and based on data across different sources, it can be impossible to identify responsibilities for errors. This can hinder accountability.

- **Challenging implementation of ethical behaviour:** AI systems make decisions often using complex mathematical equations. Despite impressive scientific progress, it is not possible to guarantee that the decisions these systems make are always ethical, especially when they operate on new data and possibly in changing environments once they are deployed. For example, there is no guarantee that an AI system will be fair across different groups of people over time, even if it has been tested and shown to be so using previous data.
- **Vulnerable to new types of attacks:** While significant progress has been made in developing AI systems that can solve complex problems, such as the recognition of objects in images or the diagnosis of diseases, work allowing the development of new types of attacks to these systems has also advanced. For example, adversarial learning methods can fool AI systems into making wrong decisions by manipulating – possibly in difficult to detect ways – the data that is used as inputs to these systems to make decisions.
- **Multiple and possibly conflicting objectives:** Trade-offs typically need to be made when developing AI systems. For example, increased accuracy of the predictions of a system may come at the cost of explainability and transparency; increased safety may come at the expense of potential loss of privacy; improving fairness may require loss in efficiency. Identifying and balancing the trade-offs is an important business decision.

4.3.2 What are the types of AI risks your business needs to consider as you deploy AI?

“

The C-suite needs to actively drive and engage in the end-to-end integration of a responsible and ethically led strategy for the development of AI in order to balance the economic potential gains with the once-in-a-generation transformation it can make on business and society. One without the other represents fundamental reputational, operational and financial risks.

Anand Rao, Global Leader, Artificial Intelligence, PwC

The risks from AI systems – especially the interplay between AI technology and how society uses the technology – are one of the key dimensions of categorization. The risk categories can be grouped into six areas (Figure 15), as outlined by PwC artificial intelligence expert Anand Rao:¹¹¹

- **Performance risks:** Risks associated with how the model performed is grouped under this category. Risks from model errors, bias in the data or the models using the data, lack of interpretability or explainability, potential brittleness or stability of model results are all examples of performance risks. The impact of these risks are physical, emotional, financial, medical harm to individuals or society and the resulting financial, reputational and operational harm to corporates that own these models.
- **Security risks:** Risks associated with the model security are grouped under this category. Risks from adversarial attacks, trojan poisoning, model inversion, deepfakes, etc., are all good examples of security risks. The security risks impact individuals, societies, corporates and nations.
- **Control risks:** Risks associated with the inability to control the AI when it malfunctions fall under this category. Examples around AI going rogue, model drift and lack of human agency in AI-driven processes are control risks that need to be monitored and intervened before they can cause harm to humans or make wrong decisions.
- **Economic risks:** Economic consequences of AI, including job losses, increased economic disparity, and winner-takes-most¹¹² dynamics are all economic risks of AI. These risks are broader than just individuals or companies and affect the entire nation or region.
- **Societal risks:** Risks to political stability, polarization of views, deepfakes, etc., that impact the reputation of individuals, corporates and society are grouped as societal risks.
- **Enterprise risks:** Risks associated with value alignment, goal alignment, and broader risks of artificial general intelligence and artificial super intelligence are grouped under this category.

FIGURE 15

Six categories of AI risks



Source: Adapted from Rao, Anand, “Five Views of AI Risk: Understanding the darker side of AI”, Towards Data Science, 29 November 2020, <https://towardsdatascience.com/five-views-of-ai-risk-eddb2fce43c2> (accessed 11 November 2021)

④ AI poses risks to individuals, groups, corporates and nations.

While AI presents significant opportunities,¹¹³ it can also cause harm. **Harm** is defined as an “adverse outcome or impact”. The “likelihood of harm” together with the “consequences of harm” are defined as the **risk**. In a formal mathematical sense, the likelihood is equated with probability and severity with the consequences. AI poses risks to individuals, groups, corporates and nations.

According to Rao:¹¹⁴

AI could pose a risk to an individual's **safety, security, reputation, liberty and equality**. Poor performance of AI could result in physical harm¹¹⁵ (e.g., accidents involving autonomous vehicles), emotional harm¹¹⁶ (e.g., mis-categorizing our emotions), financial harm¹¹⁷ (e.g., certain minority groups being flagged more for ‘suspicious’ financial activity), and medical harm¹¹⁸ (e.g., certain minority groups not receiving adequate treatment for cancer). Distinct from the impact on individuals, AI could also discriminate against specific groups of individuals, e.g., females or minority groups or age groups or socio-economic groups. Adversarial examples¹¹⁹ (e.g., strategically placed stickers on a Stop sign can fool a deep learning algorithm to treat the sign as a 30 mph sign), trojan poisoning¹²⁰ (e.g., hidden triggers embedded in neural networks to make them act erratically and maliciously), and model inversion¹²¹ (e.g., reverse engineering the ML algorithm to determine how it was trained) are some examples of security risks. The rise of deepfakes¹²² can destroy the reputation of individuals. Surveillance cameras coupled with AI¹²³ could rob individuals of their liberty and freedom. AI algorithms could institutionalize our historical biases¹²⁴ (e.g., black defendants pose a higher risk of recidivism than they actually do) and exasperate inequality.

AI can also increase the scale and speed of social inequity and amplify the winner-takes-all dynamics among companies and even countries. According to Rao, “Increased use

At the corporate level, AI could pose **financial, operational, reputational, and compliance** risks. A deepfake voice¹²⁵ of a CEO of a company was used to defraud a company for over \$240,000 recently, causing the company both financial and reputational harm. AI going rogue¹²⁶ or model drift¹²⁷ if not monitored appropriately could result in significant operational risks. Compliance risks and fines for corporates will likely increase in the future as more countries start enacting specific regulations for algorithmic accountability.¹²⁸

At the national level, AI could pose **national security threats, threaten political stability, increase economic disparity, and increase prospects of military conflict**. Automated decision making, intelligent malware, data diet vulnerability and a number of other factors associated with AI can pose national security threats. With hyper-personalization, AI can create ‘echo chambers’ resulting in increased polarization of views within a country. This coupled with automated dis-information could quickly threaten the political stability¹²⁹ of a country. AI induced automation can result in significant job losses for roles that are made up of predominantly repetitive manual and cognitive tasks. This could lead to greater unemployment and economic disparity.¹³⁰

of AI in military systems including autonomous weapons systems, robot soldiers,¹³¹ micro drones¹³² and other technologies create ethical, legal, operational and strategic threats.”¹³³

4.3.3

How can you assess what AI risks your business is exposed to?

One mechanism that has received significant focus in proposed regulation is the use of AI assessments. Many different assessments exist, the bulk of which are intended to be completed by those responsible for the first line of defence (see section 4.4.1 below).

Data Protection Impact Assessment

Some advocate for the expansion of Data Protection Impact Assessments (DPIAs), which are mandated for sensitive uses of personal data under GDPR. Some organizations have already been expanding DPIAs to include elements of ethics. One benefit of this approach is that the DPAI already includes mechanisms to weigh the risks and benefits of a specific application. However, these are limited to the use of personal data and would need to be expanded to consider the broader use of analytics and AI. DPIAs are designed to have

open ended, free text questions intended to initiate discussion as well as act as documentation. They are completed in parallel with development.

Risk assessment

Focusing on the risks primarily, a risk assessment is often a collection of a few questions intended to capture risks in the categories of materiality (e.g. financial risk), reputational risk, societal risk and scope (e.g. number of users). These assessments yield a risk tier that provides guidance on the downstream governance requirements for a given application. The proposed requirements may include the rigour of a review (e.g. a peer review vs an independent review) and are typically performed towards the beginning of application development, though the application may be reassessed as development progresses.

Ethics assessment

A review specifically to identify novel ethical risks, which would require adjudication by an Ethics Board or other independent committee, is another form of assessment. This assessment is typically performed early on in development to capture sensitive applications before they progress too far.

Bias assessment

Some proposed regulations, like a New York City Council proposed law on hiring algorithms,¹³⁴ would require a bias assessment, which is intended to be an independent review of the outcomes generated by a system for disparate impacts to specific populations. This assessment would be performed at the end of development, but prior to a system's deployment.

4.3.4 What steps can your organization take to manage AI risks and ensure the development and use of ethical AI?

Example steps to consider as a starting point include:

1. Align on AI principles relevant to the business
2. Confirm adequate top-down and end-to-end governance by instantiating stage gates throughout the AI application development and deployment process
3. Design for robustness and safety by incorporating assessment and risk-tiering processes
4. Exercise control and value alignment by enabling ongoing monitoring and adhering to the practices established by the three lines of defence structure
5. Respect privacy by considering what "should" vs "can" be done with data and elaborating on existing data governance practices

6. Be transparent by enabling the traceability, explainability and communication of information, decisions and actions of the AI system and the data that feeds the AI system, as well as the visibility into how (and which) broader systems leverage AI¹³⁵
7. Extend security practices to protect against AI-specific risks
8. Enable diversity, non-discrimination and fairness
9. Clarify and engender accountability by implementing the use of impact assessments and external auditing; establish specific requirements for the three lines of defence structure; identify and educate stakeholders in the three lines on the roles and responsibilities
10. Foster societal and environmental well-being by considering a broader scope of metrics, like ESG.

4.4 Governance

4.4.1 What guidance will help you think about your organization's AI governance?

Top-down governance: Three lines of defence

Governance and risk management are not new in the corporate world. Well accepted standards, guidelines and regulations help to ensure the smooth functioning of corporations. The Three Lines of Defence model was developed in 2008-2010 by the Federation of European Risk Management Associations (FERMA) and the European Confederation of Institutes of Internal Auditing (ECIIA) as a guidance on the 8th EU Company Law Directive article 41.¹³⁶ It was adopted by the Institute of Internal Auditors (IIA) in 2013 with their position paper.¹³⁷ Since then, the model has become the standard way of assessing and managing risks and performing governance.

In 2020, the IIA updated the guidance and released a paper on the Three Lines Model.¹³⁸ The document describes six key principles of governance, key roles in the Three Lines Model, the relationships between the roles and how to apply the model. It clearly articulates the responsibilities of the management, internal audit function and the governing body.

The Three Lines Model has been extensively used in many organizations and has been applied to a variety of risks, including technology risks and model risks, in financial services organizations. Credit risk, market risk and operational risk models in banks have been routinely governed

based on these three lines of defence. Rather than invent an entirely new governance structure, process, roles and responsibilities, this model has been adapted to apply to AI governance.

- **First line of defence – creators, executors and operations:** those specifying, designing, building, deploying and operating data, AI and ML models, automations and software. The first line also includes the operations team involved in operating and monitoring the data, software and models.
- **Second line of defence – managers, supervisors and quality assurance:** those assessing the risks of data, AI and ML models, automation and software, as well as those responsible for developing the strategy. The ongoing monitoring is also reviewed by the second line of defence. In addition,

the second line is responsible for checking that the first line has built their systems in alignment with expected practice.

- **Third line of defence – auditors and ethicists:** those overseeing the other two lines of defence to help ensure compliance with laws, policies and the organization's strategies, as well as the ethical and responsible use of technology.
- **Ethics Board** – a diverse and inclusive **group of executives and staff** within the organization. Some organizations may also choose to appoint external members to the board.

In addition to these roles within the organization, companies will also interact with external auditors, certifiers or other assurance providers as well as regulators. Figure 16 outlines the various roles and the key responsibilities of each.

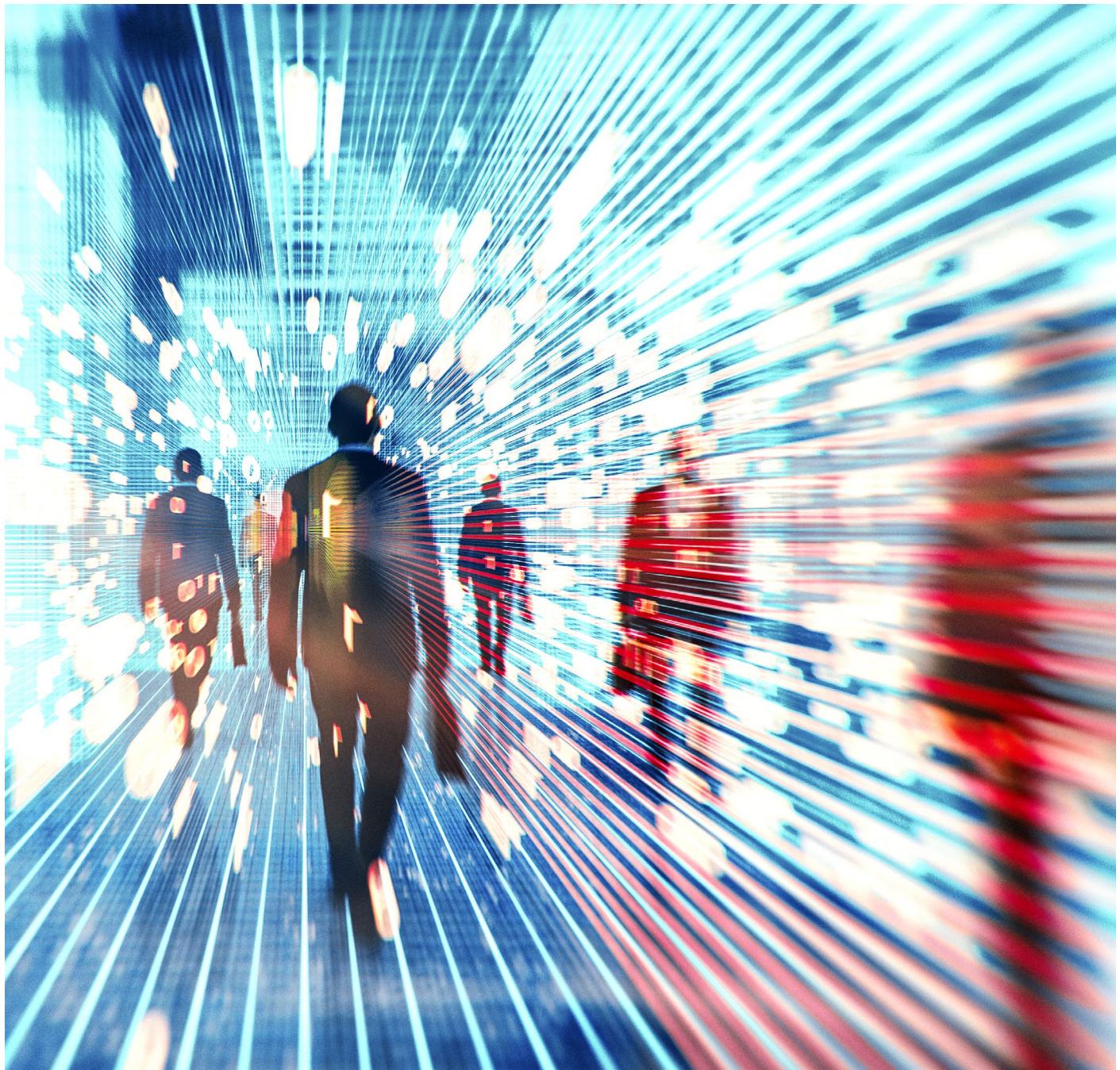


FIGURE 16 | Three Lines Model for AI governance



Source: Adapted from Rao, Anand, "Top-down and end-to-end governance for the responsible use of AI", Towards Data Science, 10 January 2021, <https://towardsdatascience.com/top-down-and-end-to-end-governance-for-the-responsible-use-of-ai-c67f360c64ba> (accessed 11 November 2021)

End-to-end governance

In addition to the top-down governance prescribed by the Three Lines Model, also needed is end-to-end governance from cradle-to-grave or from inception-to-retirement of the models and AI-embedded systems (Figure 17). The starting point for governance is not when beginning to build a model but much earlier in the life cycle.

It should really start from the **strategy** – the corporate or business strategy for the company overall and specifically for the group responsible for the data, automation, analytics and AI strategy.¹³⁹ Every company, especially those that consume or generate data and insights, should have a policy on ethics, and internal policies and procedures for adapting, adopting and practising ethical behaviour. An integral component of this stage is also understanding any regulations and best practices or guidelines from industry bodies or professional associations. For example, the Data Science Code of Professional Conduct¹⁴⁰ by the Data Science Association, and the Oxford-Munich Code of Conduct for Professional Data Scientists¹⁴¹ are good starting points for a code of conduct. Also worth tracking are the standards being developed

by the Institute of Electrical and Electronics Engineers (IEEE), especially the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems.¹⁴²

The next stage is **planning**. The quartet of data, automation, analytics and AI should be part of the planning phase. Software and AI models need to be treated very differently.¹⁴³ The experimental nature of AI models requires a portfolio approach. At any point in time, organizations that are mature in the adoption and roll-out of AI have a portfolio of models that are at different stages of evolution – conception, experimentation, deployment, production or retired. The ROI needs to be measured with respect to the overall portfolio and, depending on the overall strategy, adjusted for the right mix of business use cases, efficiency vs effectiveness initiatives, etc.

The next stage of end-to-end governance is the overall **ecosystem**, meaning the ecosystem in which the AI models will be embedded as well as the context in which they will be used by others within and outside the company. The broader social impact of AI being released by the company should be evaluated here. The IEEE's well-being metrics¹⁴⁴

are a strong contender here. The ecosystem must also cover the context in which the AI-embedded software will be used. For example, whether the system has automated, assisted, augmented or autonomous intelligence will determine the level of governance and escalation required. Change management for people who will be using the AI systems is a critical element for the successful adoption and continuous improvement of the combined human-AI ecosystem. Finally, given the availability of open source and vendor-based AI tools and techniques, having a good understanding of the minimum procurement standards for sourcing AI models is absolutely essential. Although targeted at public-sector organizations, the World Economic Forum's AI Procurement in a Box, co-developed with the UK's Office for AI is a good starting point.¹⁴⁵

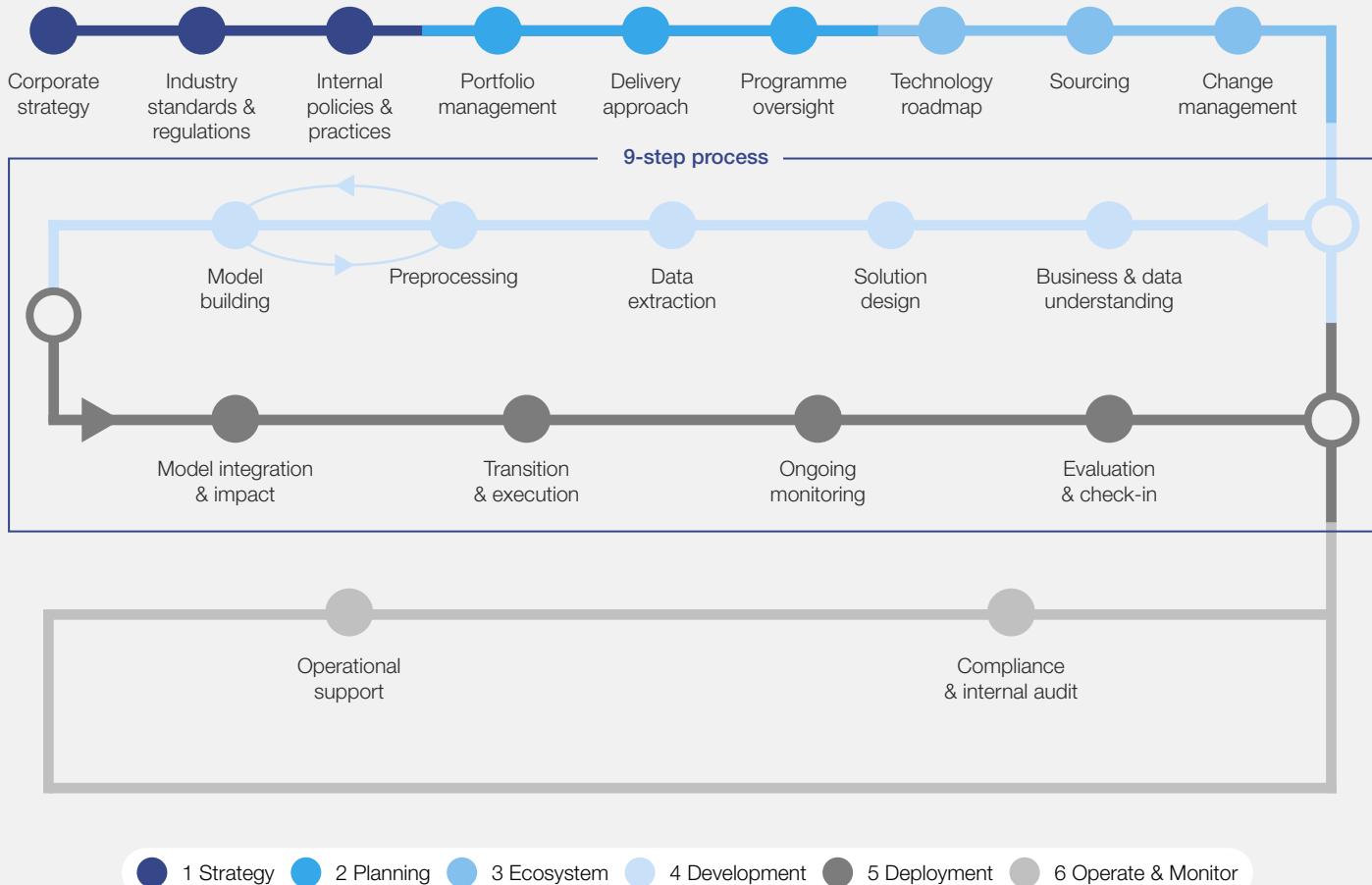
Next are the core model **development** and **deployment** stages. At this point, the governance focus shifts from the broader set of stakeholders to the development team. This team will be made up of not just the data scientists, data engineers, technologies, product managers, operations team – the technical teams – but also the business domain experts and ethicists. The governance oversight involves the phases of value scoping (includes business and data understanding, solution

design), value discovery (includes data extraction, preprocessing and model-building), value delivery (model deployment, transition and execution) and, finally, value stewardship (includes ongoing monitoring and evaluation, and check-in).¹⁴⁶ In these stages, AI applications should undergo standardized testing protocols, teams should complete standardized documentation, and maintenance plans for ongoing monitoring and upkeep should be developed.

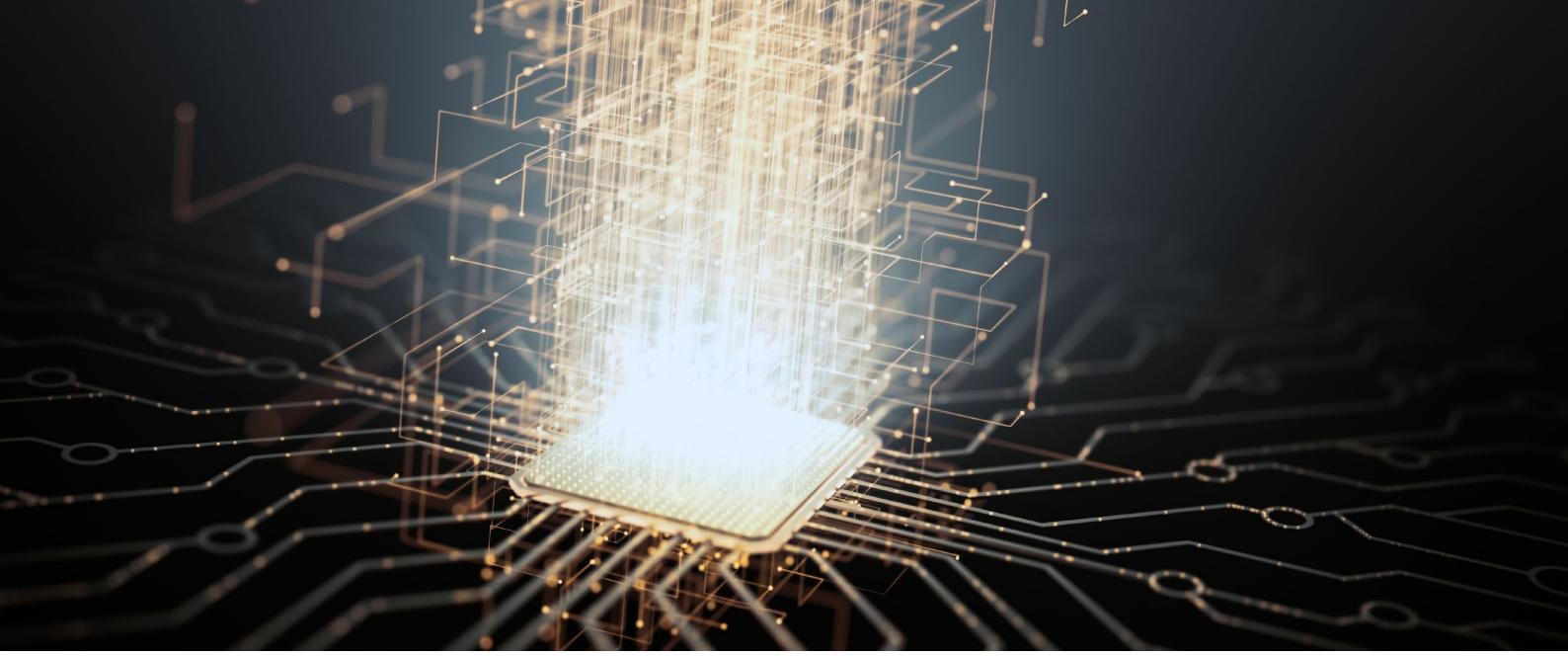
The final stage in the end-to-end governance process is the **operate and monitor** stage. The governance here is at two levels. One level is the governance of specific models and the extent to which the data, decisions, usage, algorithms and the context around which the AI models are being used are changing. The other is the portfolio level: monitoring the value delivered by the AI models or AI-embedded software and retiring models or initiating new ones going back to the value scoping stage of model development or even further back to the overall strategy of the company. In addition to this operational and strategic support, a compliance and internal audit review also needs to be conducted periodically. As discussed in the previous section, these fall under the second line and third line of defence.

FIGURE 17

End-to-end governance for responsible AI



Source: Adapted from Rao, Anand, "Top-down and end-to-end governance for the responsible use of AI", Towards Data Science, 10 January 2021, <https://towardsdatascience.com/top-down-and-end-to-end-governance-for-the-responsible-use-of-ai-c67f360c64ba> (accessed 11 November 2021)



4.4.2 How can AI governance help your organization manage risks proactively and prepare for future regulation?

Ethical failures will cost companies a great deal in government fines and penalties as well as in the erosion of share prices and reputation damage.¹⁴⁷ Increasing AI means increasing risks as well, be they in the legal or reputational domains. Costly and large compliance failures abound and are expected to multiply in the near future. Facebook, for example, was investigated extensively due to the privacy breach of 50 million users in the Cambridge Analytica case.

In the last few years, however, businesses have realized the need for ethical data governance and for the proper and agile governance of AI systems that leverage that data. As a result, numerous ethical principles have been released. PwC's ethical traceability matrix¹⁴⁸ has tracked over

100 organizations with defined sets of ethical AI principles, including corporations, public institutions and think tanks. These principles have much in common, although some include additional distinctive guidelines. The IEEE's Ethically Aligned Design¹⁴⁹ and the European Commission's Ethics Guidelines for Trustworthy AI¹⁵⁰ are two of the most commonly referenced sets of ethical AI principles. PwC has also published a set of nine core ethical AI tenets, grouped into epistemic and general principles (Figure 18).

According to an article authored by PwC¹⁵¹ for the IEEE International Conference on Big Data 2020 (IEEE BigData 2020), a leading forum on the latest research in big data:

To move from principles to practice, and to address the emerging risks that AI may bring, [organizations] are building capabilities to develop AI responsibility by addressing the strategic aspects of AI development, instituting controls for performance and technical characteristics of AI systems, as well as the broader governance and controls required for effective compliance and organizational oversight.

Organizations are increasingly considering the ethics of the applications and data they look to build and use. This extends past “*what do we have to do*” often dictated by compliance to regulation, to the “*what do we believe we should do*”. The corresponding change provides a more values-based and strategic assessment of the requirements for an organization, for an application, or for data. By implementing principles with concrete requirements, organizations can enable robust and consistent mitigation of ethical

risks and establish a mechanism to trade-off between competing priorities.

Governments globally have also been establishing national AI strategies to promote AI development within the country. These national AI [public policies] are further complemented by other soft-law, such as guidance released by regulators for specific issues, like the Consumer Financial Protection Bureau (CFPB) providing guidance on adverse action notices when using AI/ML models¹⁵² or the Information Commissioner's Office in the UK providing guidance on AI and data protection.

Most regulation to date is around data protection and privacy, though increasingly there are proposed regulations which would require assessments and governance for AI as well as limitations for its use. One active space is facial recognition, where several states and cities are banning its use by police and government¹⁵³ or limiting its use for commercial purposes.

Most importantly, regulators must be important actors in the AI governance ecosystem. Approximately **60% of organizations using AI have attracted legal scrutiny for their AI systems**,¹⁵⁴ this happened in a regulatory environment not unlike the Wild West when it comes to AI-specific regulation.

Detailed legislative proposals are emerging. One example is the first proposed AI regulatory framework by the European Parliament, which establishes a risk-based framework for AI as well as institutes' liability structures for AI-related damage. It covers AI-related intellectual property rights. This will likely have the triple effect of leading to further EU

regulation, determining access to the world's second largest market; putting the EU's weight behind and serving as an inspiration regarding its priorities, especially its strong ethical focus; and speeding up the adoption of similar AI regulatory frameworks in other democracies. In other geographies, such as in the United States, some government guidelines, especially on the procurement of ethical AI systems, are already enforced, suggesting that executives would do well to gain a competitive edge by proactively demonstrating their diligence with regard to the ethics superiority of their AI wares. Many enterprises have adopted similar AI procurement guidelines on their own.



Europe is a world pioneer in setting high-quality and ethical standards in order to protect its principles and fundamental human rights. By establishing a framework for ethics during the whole life cycle of any AI system, we can provide the safeguards to ensure that our societies harness the benefits of AI while mitigating its risks for a human centric, trustworthy, responsible and safe AI without compromising privacy and our digital rights. We need to set clear rules for solutions and decisions made by AI and Super Intelligence, to ensure human control, more options and rights for citizens, by law, by default and by design. Furthermore, Data & Metadata ethical Governance is needed for expanding the scope of GDPR and protect our Democracies from the misuse of tech and biometrics data for profiling, predictive policing, exclusion, surveillance and manipulation. Awareness is crucial to overcome the risks and achieve AI transformation in all sectors for Good for the People.

Eva Kalli, Member of the European Parliament

4.5 Operationalization of responsible AI

4.5.1 What is the business case for responsible AI?

A recent report indicates that the top three challenges and barriers all business organizations face in AI adoption are project management, AI risks and ethics management and regulatory constraints.¹⁵⁵ However, as organizations mature in their AI strategy and adoption, the importance of regulatory affairs, risks and ethics rises steeply. For leaders strategically planning for AI, it would be prudent to mitigate the forecasted risks arising from these domains well in advance to not stifle their own progress. The time has passed in the first few years when companies were exploring the imperative of doing AI ethically and responsibly. By 2020, 53% of organizations already had a leader responsible for the ethics of AI systems.¹⁵⁶ However, transparency is still lacking: only 46% of organizations have the ethical implications of current AI systems independently audited.¹⁵⁷

Practising responsible AI and revenue are interconnected.

The positive, multifactor correlation between ethics practised and revenue is increasingly common knowledge. As many as 94% of C-suite executives agree that responsible AI will produce greater ROI for shareholders, strengthening the direct correlation

between responsible AI and an investment world that has already developed a growing appetite for sustainability in the broader sense of the term.¹⁵⁸ Similarly, today's customer is likely to pay a premium for products that are responsible, secure, transparent and ethical, and this practice is even more valid in business-to-business transactions. Customers also like to spread the news about ethical AI practices: in retail, "companies that are seen as using AI and ML ethically have a 44-point Net Promoter Score (NPS) advantage over those seen as doing the opposite. For tech-savvy users, it is even higher, up to 84 points";¹⁵⁹ as well-informed digital citizens, this trait has a significant impact on revenue.

As they learn about AI, customers increasingly care, and act.

As many as "66% of customers expect AI models to be 'fair and free of prejudice and bias against them or any other person or group'; 67% expect organizations to take ownership of their AI algorithms when they go wrong."¹⁶⁰ A consumer push that the tech industry is increasingly heeding, dubbed "extended responsibility", concerning the unintended

• The increasing push from stakeholders for ethical AI, building trust in a brand to increase product quality and revenue are among the factors driving responsible AI.

consequences of AI in action, is for companies to be accountable for both AI's mistakes as well as its unintended use. If customer/user feedback and developers are brought together under the joint objective of responsible product design, especially when the process itself is enhanced with AI, the customer acquisition and retention benefits for the enterprise are unquestionable.

Responsible AI is essential to safeguard and increase brand reputation and trust.

In the current period when many companies are still early on the AI learning curve, 22% of organizations in the last 2-3 years have already faced a customer backlash due to decisions reached via their AI systems.¹⁶¹ These public opinion backlashes may get out of hand and, with high-risk AI domains and applications, can trigger a self-ban or self-restriction avalanche like the one involving many leading tech companies developing facial recognition technology in the summer of 2020.

Careful planning for responsible AI boosts product quality.

Infusing product design with ethics does not necessarily cause a slowdown. Streamlined ethical reviews can actually boost product innovation and enhance the quality continuously throughout its life cycle. A total of 97% of executives agree that this "ethics to innovation" link is essential, including the side effect of leading to increased product security and transparency.¹⁶² Practising responsible AI has demonstrated an almost doubled success rate (1.7 times) for companies in scaling up their AI solutions¹⁶³ and it also saves considerable cost instead of starting from scratch in case of an ethics backlash. As many as 90% of retail executives state they have witnessed the misuse of AI.¹⁶⁴ Product recalls and contract cancellations are the Sword of Damocles for executives not taking ethical and responsible AI at face value: even in this nascent age for ethical AI, close to a half of C-suite executives have already confessed

to cancelling AI projects at tremendous cost due to ethical concerns.¹⁶⁵

The best talent cares about responsible AI.

According to Deloitte insights, a total of 85% of enterprise leaders "believe that the future [overlapping with the present] of work raises ethical challenges – but only 27% have clear policies and leaders in place to manage them."¹⁶⁶ The readiness gap even increases when it comes to priorities: "75% of organizations say ethics related to the future of work are important or very important for their success over the next 12 to 18 months, but only 14% say they are very ready to address this trend."¹⁶⁷ The top drivers of ethical challenges are "legal and regulatory requirements, rapid adoption of AI in the workplace, changes in workforce composition, and pressure from external stakeholders."¹⁶⁸ Many of these factors come together under the umbrella of automation concerns. With the fear of being out-automated or otherwise made redundant a top-of-mind concern across the workforce in the Western world at least, the C-suite needs to speak clearly on the comprehensive talent strategy, talent management and upskilling solution they have for the AI Age.¹⁶⁹

Great talent is elusive: AI talent is even more elusive.

Ethics play an increasing role in the acquisition of talent and in retention and engagement, since employees increasingly scrutinize the ethics, beliefs and practices of their employers.¹⁷⁰ A rapidly rising ethical phenomenon is pressure coming from the workforce itself, and the more the enterprise depends on the loyalty of top talent, the higher the chances that employees will push back, or even revolt on ethical grounds. An apt example is the multiple waves of employee activism and walkouts that tech leaders have experienced regarding major defence and intelligence AI projects or sensitive AI solution deals with China, sometimes leading to major contract cancellations.



4.5.2 | What are some tools and practices your organization can use to manage AI risks and ensure the deployment of ethical AI?

Vision and values statement

Organizations must ensure that they remain grounded in socio-economic, moral and environmental priorities when addressing problems with AI and looking to empower the entire organization for ethical and responsible design and use. This requires the right vision, a vision built on strong ethical principles, adapted to the context of the organization and contextualized¹⁷¹ and agreed with all relevant stakeholders from customer to employees, leadership to suppliers.¹⁷² Having a strong ethical vision for AI, driven by C-suite executives, represents the foundation for a fair, transparent, beneficial, safe and robust outcome with AI.

Principles, codes of conduct and frameworks

The ethical principles defined by organizational values need to be translated in organizational policies, codes of conduct and frameworks to allow the operationalization of those principles.¹⁷³ A code of conduct is a policy that describes the ethical conduct expected from all stakeholders in the AI process. An organizational framework should

include clearly articulated principles tailored to the company's context as well as the perspectives of different stakeholders and guidance on how the principles should be translated in design and governance requirements.¹⁷⁴

In the last few years, hundreds of initiatives have attempted to define a set of ethical principles as the golden rules of morality at the organization level. These principles are meant to be followed when designing and developing AI solutions. To this end, it is necessary to take a systemic and pragmatic approach that, on the one hand, acknowledges the fact that moral values are the fruit of a societal, cultural and individual evolution in which their relevance is and should be continually negotiated within a community of stakeholders. On the other hand, it should be practical enough to be operationally embedded into the AI life cycle, beyond checklists and assessments, and into the permanent ethical, moral behaviour of people, including in their processes and technology itself, an "ethos of ethics".¹⁷⁵



AI is transforming every aspect of our lives, has played an essential role in the COVID-19 response and holds promise to do the same for other crises like climate change. But AI also brings real challenges for governments seeking to design policies that maximize the benefits of AI while minimizing the risks. The OECD has worked hard to help policy-makers begin tackling these challenges, adopting in 2019 the first intergovernmental standard for the responsible stewardship of trustworthy AI. But the OECD AI Principles are just the start; much work is ahead to turn the principles into practice so that AI systems are designed to be robust, safe, fair and trustworthy.

Ulrik Vestergaard Knudsen, Deputy Secretary-General,
Organisation for Economic Co-operation and Development (OECD)

FIGURE 18

The 9 ethical AI principles

The epistemic principles

Interpretability (Explainability, Transparency, Provability)



An AI system should be able to explain its model decision making overall and what drives an individual prediction to different stakeholders.

Reliability, Robustness, Security



AI systems should be developed such that they will operate reliably over long periods of time using the right models and datasets.

The general ethical AI principles

Accountability



All stakeholders of AI systems are responsible for the moral implications of their use and misuse. There must also be a clearly identifiable accountable party, be it an individual or an organizational entity.



Human Agency

The degree of human intervention required as part of AI solutions' decision-making or operations should be dictated by the level of perceived ethical risk severity.

Data Privacy



Individuals should have the right to manage their data that is used to train and run AI solutions.



Safety

Throughout their operational lifetime, AI systems should not compromise the physical safety or mental integrity of humans.

Lawfulness and Compliance



All the stakeholders in the design of an AI system must always act in accordance with the law and all relevant regulatory regimes.



Fairness

The development of AI should result in individuals within similar groups being treated in a fair manner, without favouritism or discrimination, and without causing or resulting in harm. It should also maintain respect for the individuals behind the data and refrain from using datasets that contain discriminatory biases.

Source: Adapted from PwC analysis, published in Golbin, Ilana, and Maria Luciana Axente, "9 ethical AI principles for organizations to follow", World Economic Forum Agenda, 23 June 2021

External and internal boards and committees

Ethics bodies have many names and shapes,¹⁷⁶ but they remain an integral component of the data and AI ethics ecosystem that is vital to operationalizing ethics governance within the organization. They can be internal and external, top-level or cross-organizational, and their mandates are an amalgam of:

- Representing a broad set of perspectives and expertise
- Identifying new opportunities for using data and AI to better support the organization's mission and values
- Developing tools, frameworks, practices, standards, programmes and training to help employees identify and report ethical issues
- Acting as an agile governance body that serves as a continuous repository for institutional knowledge.

Data and AI governance is very much an evolving field; few templates of an ethics body exist. Ethics bodies are often established reactively and in haste, frequently as a band-aid solution to a full crisis,

which can be counterproductive. A well-functioning and effective committee is adequately resourced, aptly positioned in the organization, sufficiently empowered and thoughtfully designed, and its responsibilities and purview are clearly defined.

Culture of ethics

AI is changing not only business models and processes but, even more fundamentally, every aspect of society. Cultures are at the heart of this change and are where ethical skills, knowledge and behaviour should be recognized, rewarded and appreciated. It will take time and patience for organizations to consider and appreciate AI ethics as a core competency across diverse roles in the AI life cycle, but ethically aligned upskilling coupled with formal ethics education will raise the awareness of these skills' importance and will create the need for industry to apply AI ethics when making decisions about AI products. The ultimate goal is to encourage a more participatory, collaborative and diverse mindset when building AI products, a fundamental shift from the "Can I build it?" to the "Should I build it?" mindset. This can be achieved by raising ethical awareness in the organization and providing AI ethics champions as examples, which will foster the right cultural context and behaviour more quickly.



In our dream world, success will lead to better-informed policy-makers and new corporate models of organization that build ethics into all stages of design and corporate leadership.¹⁷⁷

Jeffrey Behrends, Fellow-in-Residence,
Edmond J. Safra Center for Ethics, Harvard University

Ethics education and training

Because technologies like AI magnify the outcomes they deliver, whether beneficial or harmful, the process of value alignment and ethics adoption is the responsibility of all those involved in the field: the direct stakeholders (developers and users) as well as indirect stakeholders (citizens, owners, unions, regulators, governments). Only by educating the stakeholders on ethics for AI, including how to navigate the moral dilemmas and what is acceptable or permissible, will they be able to acquire a mindset and behaviour that allow responsible AI to be developed. Formal training programmes and curricula in ethics should be embraced, as well as other activities that teach ethical thinking, analysis and reasoning, such as communities of practice, responsible AI events, book clubs, team debates and hackathons. Ultimately, the emphasis should

be on embedding ethics into all technology (and science, engineering and mathematics) courses at all levels, following the example of Harvard University's Embedded EthisCS course,¹⁷⁸ and bringing that training to corporations for executives at all hierarchy levels.

Furthermore, current digital upskilling programmes should consider the impact that technologies have or can have, the limits of individual and collective responsibility involved in them and the levels of participation (whether co-designing them, challenging them or making decisions about them, etc.). Such upskilling should not only provide domain knowledge of specific technologies but should help to develop a broader mindset for working with technology, similar to developing adaptability or empathy, and recognizing its impact.



Ethical reasoning is part of what you do as a computer scientist.¹⁷⁹

Alison Simmons, Samuel H. Wolcott
Professor of Philosophy, Harvard University

The experts: ethicists

An important distinction can be made between the profession of AI ethicist and the field of AI ethics itself. The field involves a diverse set of roles: technologists, data scientists, software engineers, evangelists, policy-makers and researchers, with the mission to develop responsible AI for the benefit of humanity.¹⁸⁰ The AI ethicist's main role is to translate abstract ethical theories and principles into practical

moral consideration of the use and design of AI and to navigate moral dilemmas, especially in novel AI use cases and applications, like facial recognition, virtual companions, companion robots or predictive algorithms in the justice system. To do so, the AI ethicist needs to deploy multidisciplinary knowledge of AI technology, business, policy and law, and play a convener role in many other disciplines that contribute to the design, development and use of AI systems.



What does it mean to [be] an AI Ethicist in industry? It means to be an individual with a robust knowledge of ethics who possesses the ability to apply and communicate ethical principles in the context of artificial intelligence within the corporate structure, and to do it all with bravery.¹⁸¹

Olivia Gambelin, Founder and Chief Executive Officer,
Ethical Intelligence Associates

Reporting and advice channels

These are important for fostering and building a transparent, trustworthy and inclusive environment where gathering information about AI ethics is normal practice. Ensuring appropriate means and ways for employees to receive advice on ethical dilemmas or to report breaches pertaining to AI and data is a good measure that will help to identify potential ethical issues and solve them before they escalate. A common practice is for organizations to set up ethical hotlines¹⁸² or other channels for employees to report or alert management to the presence of ethical abuses. Hotlines, anonymous forms or reports to direct supervisors can help to identify ethical issues pertaining to the development and use of AI.

Stakeholder consultation and dialogue

Never before has the role of stakeholders in shaping the moral stance of an organization been more important, with AI touching more groups ever more profoundly and with lasting effect. From children who use virtual assistants not built for them,¹⁸³ to women more likely to lose their jobs to automation and AI,¹⁸⁴ to workers surveilled via digital collaboration tools,¹⁸⁵ more stakeholders are affected by AI initiatives, such that various means to engage with them are necessary. Surveys that gather views on specific issues to include them in design decisions, public consultations, citizen juries¹⁸⁶ and workshops are forms of stakeholder consultation and dialogue that promote greater accountability and transparency in the age of AI.



Since technology is so interwoven into social destiny, its potential harms often fall upon the already vulnerable. We have to broaden our focus. It's not just about business or even user needs any more: responsible AI can only come from understanding and mitigating the damage that might befall underrepresented groups and indirect stakeholders. This means embracing approaches such as co-design and radical product inclusion. Design with, not just for.

Cennydd Bowles, Managing Director, NowNext

Audit and reporting

Auditing as a “structured process whereby an entity’s present or past behaviour is assessed for consistency with relevant principles or norms”¹⁸⁷ is a good tool to ensure the ethical alignment of AI-based systems but should not replace ethical reasoning. In fact, by improving the traceability of AI systems via evaluation and documentation, auditing can support ethical discussions on transparency, fairness or accountability with a view to building stakeholder trust in AI products. Moreover, designing for auditability could help society mitigate the ethical challenges associated

with AI-based systems while reaping the full benefits of automation. However, the term “audit” may be interpreted in multiple ways, which causes confusion to corporations and consumers, as well as consumer advocates. In the assurance field, an audit is a formal evaluation by a third party to a defined set of standards. Without standards, which some organizations like the IEEE are defining, it is challenging if not impossible to perform an audit. ForHumanity has created a taxonomy to clarify these differences.¹⁸⁸ Others are also proposing taxonomies that consider the consumer of the audit, the nature of the audit and requirements for output.

4.5.3 | How can your organization adapt its product development practices to ensure they are ethically aligned?

The components of AI ethics management create the operational set-up for ethical analyses and reasoning to happen at the product development level. This is the level at which development teams come together to frame problems and find innovative solutions with AI to address ethical dilemmas and learn to mitigate harms and risks. It is therefore crucial for the product development process to be ethically aligned (Figure 19), with checks and balances embedded at every

step. And while embedding ethics in the software and data science life cycle is a new endeavour, available tools, frameworks and techniques can help educate the team and guide the process. Development teams should experiment using tools already created, from impact and risks assessment to diagnostics and consequences games (games that illicit the intended and unintended consequences of AI solutions), or create new resources.

FIGURE 19

New mindset for ethically aligned product development

Responsibility/Ethics *in* Design

- Consideration in development *processes* of the ethical and societal implications of AI as it integrates and replaces traditional systems and social structures

Responsibility/Ethics *by* Design

- Integration of ethical *reasoning* abilities as part of the behaviour of artificial autonomous systems

Responsibility/Ethics *for* Design(ers)

- Research integrity of *researchers* and manufacturers, and certification mechanisms

Source: Adapted from Dignum, Virginia, "Responsible Artificial Intelligence", UMEA University, https://ec.europa.eu/jrc/communities/sites/default/files/03_dignum_v.pdf (accessed 11 November 2021)



In order to develop a product that is ethically aligned, stakeholders' feedback is crucial to design a system that takes ethical and social issues into account. To ensure representation of stakeholders, organizations should enact a planned and controlled set of activities to account for the interests of the full range of stakeholders or practitioners who will be working alongside autonomous and intelligent systems and incorporating their insights to build upon, rather than circumvent or ignore, the social and practical wisdom of involved practitioners and other stakeholders.¹⁸⁹

IEEE, Ethically Aligned Design, First Edition

How can your organization implement responsible AI?

Over 200 organizations globally, including the OECD, have published principles for trustworthy, responsible and ethical uses of AI. Operationalizing these principles has become critical due to increases in disruptive technologies, regulation and consumer distrust of algorithms, prompting companies to reveal the practices they have adopted to ensure the responsible use of technology

Case Study 1

Microsoft

Microsoft's responsible AI programme

As part of its work to empower people to achieve more, Microsoft has focused on developing and deploying cutting-edge AI in a way that is responsible. Following the publication of the company's AI principles,¹⁹⁰ Microsoft has spent

the last few years expanding its responsible AI programme to put these principles into practice and ensure people can realize the potential of AI in a way that is responsible. Excerpts of the core building blocks of this programme, as outlined by Microsoft, are reproduced below.

The building blocks of Microsoft's responsible AI programme¹⁹¹

Governance as a foundation for compliance

Microsoft's responsible AI governance approach borrows the hub-and-spoke model that has worked successfully to integrate privacy, security and accessibility into our products and services.

Our "hub" includes: the Aether Committee, whose working groups leverage top scientific and engineering talent to provide subject-matter expertise on the state-of-the-art and emerging trends regarding the enactment of Microsoft's responsible AI principles; the Office of Responsible AI, which sets our policies and governance processes; and our Responsible AI Strategy in Engineering (RAISE) group, which enables our engineering groups to implement our responsible AI processes through systems and tools. The three groups work together to set a consistent bar for responsible AI across the company and they empower our "spokes" to drive initiatives and be accountable for them.

Developing rules to enact AI principles

In the fall of 2019, we published internally the first version of a Responsible AI Standard, a set of rules for how we enact our responsible AI principles underpinned by Microsoft's corporate policy. . . . Through a phased pilot across 10 engineering groups and two customer-facing teams, we learned what worked and what did not. . . . There was a thirst for more tools, templates, and systems, and for a closer integration with existing development practices.

Just over a year later, we're previewing version two of the Responsible AI Standard with our employees. . . . It will mandate that teams building AI systems meet requirements that accrue to principle-specific goals.

For each requirement in the Responsible AI Standard, we will build out a set of implementation

methods that team can draw upon, including tools, patterns and practices crowdsourced from within and outside the company and refined through a maturity process. We expect this to be a cross-company, multi-year effort and one of the most critical elements for operationalizing responsible AI across the company.

Drawing red lines and working through the grey areas

In the fast-moving and nuanced practice of responsible AI, it is impossible to reduce all the complex sociotechnical considerations into an exhaustive set of pre-defined rules. . . . Our sensitive uses process¹⁹² requires that use cases that meet review criteria are reported to our Office of Responsible AI for triage and review. . . . Since July 2019, we've processed over two hundred use case reviews, including an uptick in reviews since March 2020 as more Microsoft teams and customers sought to use AI technologies amid applications and opportunities with harnessing data and AI methods to mitigate challenges with COVID-19.

Outcomes of the process include our declining opportunities to build and deploy specific AI applications. . . . For example, Microsoft President Brad Smith spoke publicly about how, through our sensitive uses review process, we determined that a local California police department's real-time use of facial recognition on body-worn cameras and dash cams in patrol scenarios was premature, and he shared the fact that we turned down the deal.

We also came to appreciate the importance of three key lessons: 1) by digging into the details of use cases, we've been able to understand and articulate their different risk profiles . . . ; 2) we've learned the important roles that benchmarking and operational testing play . . . ; 3) we've learned how we need to communicate with our customers to empower them to deploy systems responsibly.

These learnings helped inform new practices at Microsoft. For example, we developed Transparency Notes to help teams communicate the purposes, capabilities and limitations of an AI system . . . Our Face API Transparency Note¹⁹³ was our first attempt at this new practice, and we now have a growing number of Transparency Notes being prepared across our platform offerings.

Evolving our mindset and asking hard questions

In 2020, our mandatory Introduction to Responsible AI training helped more than 145,000 employees learn the sensitive use process, the Responsible AI Standard and the foundations of our AI principles.

Additionally, we introduced Envision AI, an applied workshop and practice for completing impact assessments. Developed by our Project Tokyo team¹⁹⁴ and Office of Responsible AI, Envision AI takes participants through real scenarios that emerged while our Project Tokyo team was immersed in designing an approach to intelligent personal agent technology.

The kind of mindset shift we are guiding involves an ongoing process of dialogue, integration and reinforcement. At times, our teams at Microsoft have experienced galvanizing moments that accelerated progress [and realize] that a human-centred approach to AI results in not just a responsible product, but a better product overall.

Pioneering new engineering practices

Privacy, and the GDPR experience¹⁹⁵ in particular, taught us the importance of engineered systems and tools for enacting a new initiative at scale.

. . . As we have been rolling out our responsible AI program across the company, the existence of engineering systems and tools to help deliver on our responsible AI commitments has been a priority for our teams. . . . In recognition of this need, we are embarking on an initiative to build out the “paved road” for responsible AI at Microsoft – the set of tools, patterns and practices that help teams easily integrate responsible AI requirements . . . AzureML serves as the foundation for this paved road, leveraging the early integrations of our open-source tools, Fairlearn¹⁹⁶ and InterpretML,¹⁹⁷ so that our customers will also benefit for our development of engineering systems and tools.

Scaling efforts to develop AI responsibly

As we look ahead, we’ll focus on three things: 1) consistently and systematically enacting our principles through the continued rollout of our Responsible AI Standard; 2) advancing the state of the art of responsible AI through research-to-practice incubations and new engineering systems and tools; 3) continuing to build a culture of responsible AI across the company.



Deutsche Telekom (DT) is a leader in operationalizing AI ethics. Its approach is centred on building trust in the use of AI, with a commitment to transparency, explainability, data privacy and

security. The company's four-year journey to operationalize AI ethics and nine key lessons to ensure the ethical use of AI technology can serve as an example for organizations globally.

Lesson 1. Commit to digital ethics as a strategic advantage. Digital ethics can be used a competitive advantage. Creating a competitive advantage based on AI ethics requires building trust and AI acceptance. The result is a tangible financial benefit to the organization. In any given year, DT has over 600 IT projects that use AI throughout the company.¹⁹⁹ AI projects are used to detect cybercrime, to optimize network planning and, increasingly, to improve the customer experience with tools like customer chatbots and digital assistant products. In the context of such large-scale use of AI, DT treats AI ethics as a competitive advantage.

Lesson 2. Co-create AI Ethics Guidelines with technical and non-technical staff. DT started its journey by organizing six months of workshops with data scientists, engineers, the management team and compliance personnel. The group focused on creating a narrative on AI's innovation opportunities and challenges. Technologists broadly saw a new set of engineering tools to deliver more powerful applications. Others voiced concerns over job security, bias and company reputational risks. This led DT in 2018 to publish its "Guidelines for Artificial Intelligence" with nine principles designed for integration into the life cycle of AI product development, deployment and usage by data scientists, developers, engineers and project managers. The company's Management Committee approved the Guidelines and distributed them both internally and externally:

Deutsche Telekom Guidelines for Artificial Intelligence²⁰⁰

1. **Responsible** – provide a clear definition of who is responsible for which AI system or feature and its decisions and actions
2. **Careful** – use technology in accordance with company values and human-defined rules
3. **Supporting** – use AI to benefit customers who must come first
4. **Transparent** – ensure awareness of customers when they are using AI and over the use of their data
5. **Secure** – ensure customer data is protected against unwanted external access

6. **Reliable** – ensure AI algorithms are transparent, auditable, fair and fully documented
7. **Trustworthy** – ensure humans always remain in control of AI and have the ability to deactivate and stop the systems at any time
8. **Cooperative** – use the complementary strengths of AI and humans to cooperate for better decision-making and effectiveness
9. **Illustrative** – share and enlighten employees, companies and society to better deal with the challenges of the transformative power of AI.

These principles were based on five meta principles: 1) **responsibility**; 2) **security** and **defence**; 3) **explainability**, to ensure a transparent and reliable AI system by defining and explaining the processes; 4) **self-determination**, to protect against manipulation and influencing the behaviour of users; and 5) **bias avoidance**, to ensure that AI decisions do not discriminate against any groups unfairly.

Lesson 3. Obtain credible, visible leadership and support for the AI ethics programme.

The DT Management Board supported the AI ethics programme in various ways. The company appointed a Head of AI Risk and Compliance, who oversees a team in the legal department. This top-down mandate, supported by the Chief Technology Officer, provided legitimacy to the company-wide AI ethics focus. The Head of AI Risk and Compliance is a technologist by training and an experienced business leader, adding credibility to the effort by working together with the technology teams to balance innovation with the requirements of AI ethics by design and default.

Lesson 4. Build AI ethics controls into the technology life cycle.

DT implemented an ethics control framework that includes review and approval mechanisms throughout the AI technology development life cycle. These controls apply at all stages of the product:

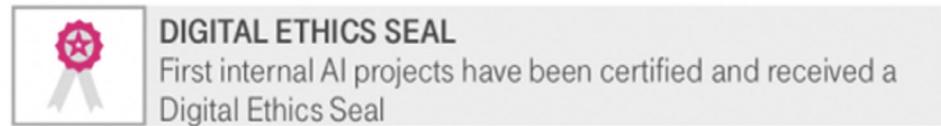
Central Innovations Projects must commit to ethical AI guidelines to receive funding.

1. At project inception, project funding is linked to stating an understanding of and agreeing to the Guidelines for Artificial Intelligence.

2. The AI Compliance team reviews a monthly report listing all AI projects and their self-reported compliance with the Guidelines for AI.
3. Enforcement is conducted by the AI Compliance team, which aims to audit 10-15% of the 50 new or updated monthly AI projects. The team checks such areas as bias, security and safety on an ongoing basis.
4. At the go-live phase, project teams must confirm that they have complied with the Guidelines for AI. The project can be certified and awarded a Digital Ethics Seal (Figure 20). While this is not mandatory, over time the absence of a seal will likely elicit surprise or disapproval.

FIGURE 20

Digital ethics seal



Source: Deutsche Telekom, "Digital Ethics: Shaping AI Responsibility", presentation at the British Embassy in Berlin, 19 December 2019, p. 9

Lesson 5. Provide broad education and engagement on AI ethics both internally and externally. DT Launched a Digital Ethics Centre in Berlin to make the risks of AI more tangible and less theoretical. The Centre is designed to educate internal staff as well as corporate customers, partners and other key stakeholders. It showcases AI risks, such as the use of AI in children's dolls. Because of the COVID-19 pandemic, the Centre is being transformed into a virtual exhibition that allows everyone to experience the exhibits online.

DT also delivers training programmes, including:

- A half-day face-to-face training with data science staff in groups of 20-30 employees
- A Digital Ethics eLearning programme for larger audiences (500 developers recently completed the course)
- Mandatory management compliance training, required for certain DT managers, with added AI and data ethics topics alongside more traditional compliance subjects, such as bribery and corruption.

The company is working to provide all new hires with training on AI data ethics risks. DT is committed to a wider dialogue on the impact and governance of AI.

Lesson 6. Implement explainability processes. DT describes its AI systems in terms of the processes used to build the system and how the AI algorithms make decisions. Explainability

is regulated under the GDPR, under which customers are entitled to receive common-sense explanations for why an algorithm has made a particular decision. DT has set out to provide AI Explainability Statements with a particular focus on bias avoidance.

Lesson 7. Ensure AI systems have a “kill switch”. All of the DT AI systems can be stopped at any time and can revert to functioning under legacy, rule-based operations. DT has taken a considered approach to how humans should interact with AI systems and the level of automation. The company minimizes full automation with “humans-out-of-the-loop” and focuses on using AI with “humans-in-the-loop” using the AI system for completing tasks or in a supervisory “human-over-the-loop” role.

Lesson 8. Extend the AI Guideline to suppliers. To help ensure that AI ethics are reflected across the increasingly complex supply chain, it is critical that vendors be certified as following responsible AI practices to ensure that no flow-through AI risks exist. DT has implemented a vendor approved status for those who meet their Guidelines for AI. Vendors are now increasingly leveraging their status as an “approved DT AI vendor” for their own market differentiation and advantage.

Lesson 9. Start the AI Ethics journey now. DT's AI ethics programme is in its third year. The journey has been an evolution, with continual learning and iteration. The company has issued its third iteration of the Guidelines for AI, increasingly tailored to roles for testers, coders, analysts and project managers, etc.

Implementation of AI

The practical implementation of AI sets it up for success.



5.1 Introduction

As organizations turn to AI to improve performance, innovate, reduce cost, and mitigate risks, it is important that they understand the foundational requirements for the successful implementation of AI throughout their business. AI is not a tool that works independently of other business functions or technologies; rather, it is dependent on factors such as the organization's strategy, people, data, technology, and ecosystem integration. To ensure the organization is effectively and responsibly leveraging this technology it must consider each of these components.

The AI life cycle is the iterative process AI solutions follow, from the initial ideation of the tool in the minds of organizational stakeholders, to the use and maintenance of the tool in an operational environment. In some cases, it can also include the decommissioning of the tool, if the enterprise determines it is no longer required. The AI life cycle is broken down into 4 stages:

Stage 1

Design

5.1.1 How can you set up an AI implementation for success?

While this module focuses on the practical implementation of AI, from model development to decommissioning, a model is only successful if it solves the right problems for the organization, in the right order of priority, with the right stakeholders, bringing together the right technical and soft skills, aligned under a leadership team that understands both the potential as well the limitations and risks of these technologies.

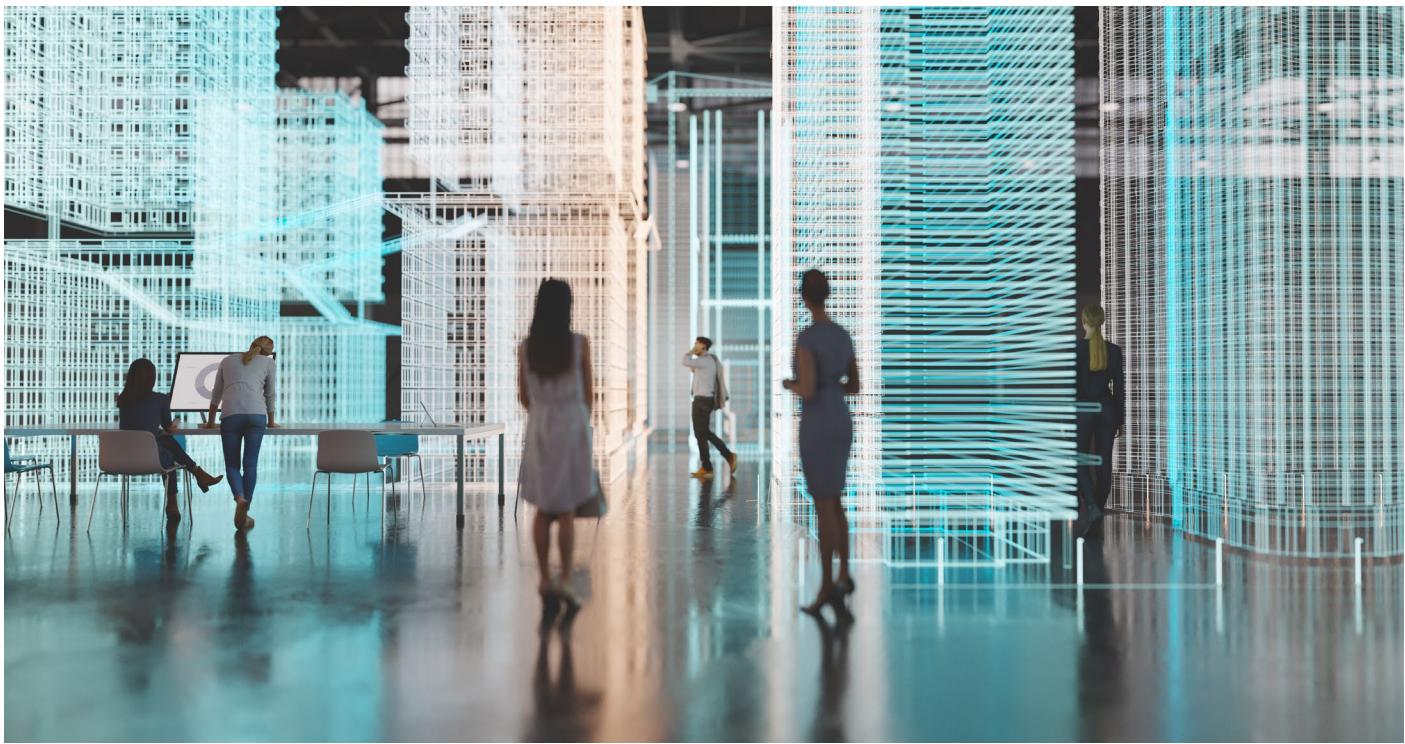
Before forging ahead on the implementation of an AI solution, it is important to check that you have:

- Clarified your business strategy and vision and used it to help define concrete problems you want AI to help you solve (see [section 2.3](#) "What is an AI strategy and why does it need to be aligned with your organization's larger corporate strategy?")
- Defined a prioritized roadmap of potential AI solutions, including an impact and cost analysis and established metrics for success (see [section 2.8](#) "What steps are involved in designing an AI roadmap?")
- Raised awareness of the business risks inherent in developing an AI solution (see [section 2.10](#) "What are the main reasons that AI initiatives fail to deliver business value and how can you mitigate these risks?")
- Hired the right talent to execute the use case (see [section 3.9](#) "What are the most important skills to hire for in an AI team?")

- Assembled the right project team (see [section 3.8](#) "Who are the main stakeholders and what are the roles in an AI project team?")
- Aligned with the right stakeholders (see [section 3.13](#) "How can the AI team align with business, on the one hand, and engineering, on the other?")
- Integrated ethical principles into your AI development process (see [module 4](#), "Responsible AI").

Taken together, these actions can be thought of as the **design stage** of an AI implementation. Once you are confident in your ability to complete the above items, it is time to dive into the **development stage**. Developing an AI model includes multiple steps, from collecting, storing, cleansing, analysing and synthesizing the available data and understanding its assumptions as well as how to govern that data appropriately, to selecting a model, training it on the data, preparing to move it to production when it can begin generating business insights, and choosing the right tech, data and AI partners to work with. Next comes the **deployment stage**, which includes pushing your model into production and scaling it and, finally, the **monitoring and feedback stage**, which can include continuous iterations on the original model, as well as measuring value and even retiring the model completely, when appropriate.

An AI model can be successful if it solves the right problems for the organization, in the right order of priority, with the right stakeholders.



Stage 2

Development

5.1.2

Where does the data to train AI come from and how can it be sourced?

Data is a fundamental requirement to train and use AI models, thus organizations must adopt robust practices that allow them to gather and activate high-quality and large quantities of data. Often, collecting, storing and cleansing the data necessary to build a model forms the bulk of the work needed to build a working model.

With regard to collecting data that can be used to develop new AI use cases, organizations need to ensure they have enough data to train robust AI models (accuracy) and the right quality of data for the AI models to serve their intended purpose (validity). To ensure proper data standards are enforced, organizations must implement data governance. This section outlines some of the most important elements data executives must consider.

Depending on the type of AI application being developed or implemented, the model will require specific types of data on the intended subject. Because of the nature of ML, the organization must have enough quantity and richness of data to both train and test the models. Not having enough data often leads to inaccurate results and limited reliability. Organizations can turn to internal and external resources to collect first-, second- and third-party data. In all cases, an understanding of the overall “data supply chain”²⁰¹ is necessary to ensure quality management and accountability.

First-party data refers to data directly generated by the organization itself (e.g. transactions, such as customer purchases). Organizations tend to rely the most on first-party data as it is directly sourced, making it trustworthy and free to use (after considering all the ethical concerns). When first-party data is insufficient, organizations turn to external secondary- and third-party data.

Second-party data is the practice of buying or sourcing external entities' first-party data. The external entity collects their own first-party data and provides it to others via legal agreements or data marketplaces. Third-party data is the practice of sourcing data from an aggregator of second-party data that does not have a direct relationship with the data source. Third-party data is frequently more robust and complete as it combines various sources, but it can be risky as it is often difficult to confirm the data collection processes.

In addition to data sourcing, the entire data supply chain must be considered. Sometimes referred to as “DataOps”, the data supply chain is a combination of the practices, policies and technologies required to handle, analyse and build on the organization's data. Optimizing this supply chain improves business agility, reduces data deficits and gives the business confidence in the data being used. While this section of the toolkit covers data, the next section focuses on technology considerations.

5.1.3 How do you know that you can trust the data or that it is of high quality?

Organizations can use a combination of techniques to source the right amount of data on the required subject. But in addition to having enough data, organizations must also ensure that it meets high quality standards. For data to meet these standards, organizations should strive to ensure data is being measured against the following six data quality dimensions:²⁰²

- Completeness: the ability to ensure the data is comprehensive with minimal missing values. This is often expressed as a percentage of fields that have a value (field with a value divided by total expected fields).
- Uniqueness: the measure of ensuring data is only recorded once with minimal duplicative values. This is represented as a percentage (expected unique entities divided by actual unique entities).
- Timeliness: the measure of delay between when the data is generated/accurate and ready to be used. Timeliness is an important dimension for data, which is time sensitive and quick changing. A large delay may lead to inaccurate data.

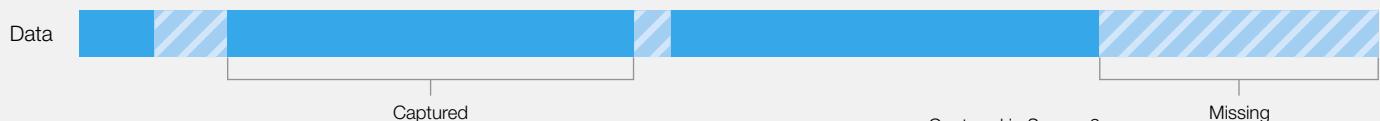
- Validity: the practice of ensuring the data is captured in the correct format or syntax. The "Date of Birth (DoB)" field, for example, should be captured in a date format.
- Accuracy: the ability to measure if the data represents the "real world subject" being studied. Although this may be hard to quantify, data gathering techniques and entry rules should be in place to reduce bias and risks of inaccurate collection.
- Consistency: the measure of ensuring alignment among data stored in multiple places about the same topic. This is often measured as a percentage (the number of matches/the total number of duplicate fields).

This list of dimensions is not exhaustive and not every data field will require that all dimensions be met, but these dimensions provide insight into the characteristics of "good data". They can be measured and scored in various data governance tools, such as a data quality scorecard or data quality dashboards.

FIGURE 21

Example of data quality

Completeness



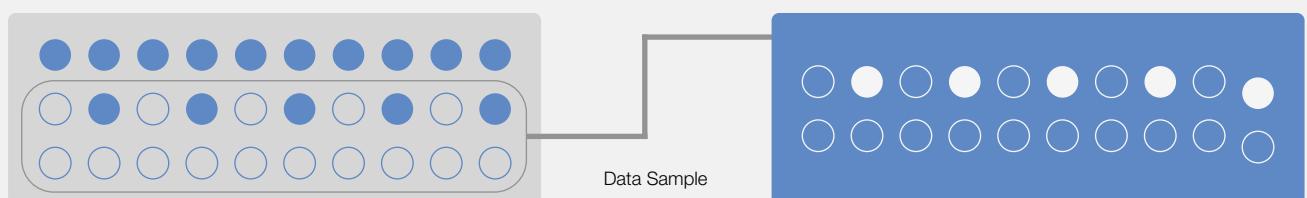
Uniqueness



Timeliness



Accuracy



Note: Sources 1 and 2 depict data from different sources. Source: World Economic Forum and Deloitte

In the example in Figure 21, the data can be considered inaccurate as it does not reflect the real world.

5.1.4 | What does good data governance consist of?

For an organization to truly become data driven and discover practical means to apply data, it must ensure that appropriate practices are in place for legal compliance, risk management, access and security, quality control and usability of data. Often this is done by implementing formal data governance practices, including roles and specialized tooling/systems. Data governance stakeholders are tasked with defining, managing and maintaining the organization's data policies, procedures and data governance tools.²⁰³ Although the roles, responsibilities and structure of a data governance team may differ, strong data governance or management at a minimum addresses the following aspects:²⁰⁴

- Data quality: the usability and level at which the data is fit for its intended purpose. Strong data quality is important to ensure the data is trustworthy, actionable and standardized. This requires standards, processes and people to make sure the data is accurate, timely, consistent and complete. Data quality is typically checked periodically and measured by adherence to business rules.
- Metadata management: the administration of data that provides important information about a company's data assets, such as who is

responsible for the data, where it comes from, how often it is maintained, what the intended purpose is, and on privacy, retention and security concerns, etc. Best practices include collecting, maintaining, using and refining metadata on a consistent basis across the enterprise in a manner that promotes optimal leverage of the company's data assets.

- Data storage and retention: the housing, archiving, retention and disposition of structured, semi-structured and unstructured data, related to the wider processes and tools used for Data Lifecycle Management and Records Management. A multilayered governance organization and process should be in place to verify that data storage is properly managed and meets the business needs and any regulatory requirements.
- Data access, privacy and security: the ability to provide efficient access to trusted data and safeguard important information from corruption, compromise or other loss. It is important for the organization to balance analysts' ability to access the data they require while maintaining data privacy and security. This includes access control, anonymizing personally identifiable information where applicable, and cybersecurity.

5.1.5

How can you enforce good data governance?

A strong data governance framework is a prerequisite to building trusted AI.

Data governance provides oversight over the entire data life cycle, from collection to use and disposal. Along with the policies and procedures established by the organization's data governance leaders, it is necessary to implement the right technical tools, such as data dictionaries, data catalogues and data quality scorecards. These tools help ensure the orderly maintenance and easy sourcing of data assets for data users.

The data governance team is tasked with developing and enforcing data standards and procedures throughout the organization. This can

be done using various operating models, ranging from a centralized data governance team, which is leveraged by all lines of business, or a decentralized team, which has data governance talent within each line of business operating independently. Regardless of the operating model used within the organization, a data governance team typically consists of multiple roles, including but not limited to data stewards, custodians, owners, analysts and scientists. The responsibilities within these roles range from maintaining the data to analysing and consuming the data. The correct operating model depends on each individual organization's needs and structure.

5.1.6

What technical questions do you need to answer when building AI?

Implementing AI that "works" may require that a certain level of technical requirements also be met. Depending on the particular use case or application being developed, an array of technical resources can support the development and implementation of AI as well as answer the team's questions. This section highlights some of the questions, but it is important to note that it is not a list of prerequisites but an introduction to basic technical questions that often arise when developing AI applications.

Although AI sounds complex and is at the cutting edge of current technology, the AI ecosystem has reached a point where accessing the tools, libraries and platforms required to develop AI models is relatively easy. Depending on the complexity of the model or application the organization is looking to build, developers and data scientists will need a combination of data, computing power, storage and development environments to put into production and host AI models.

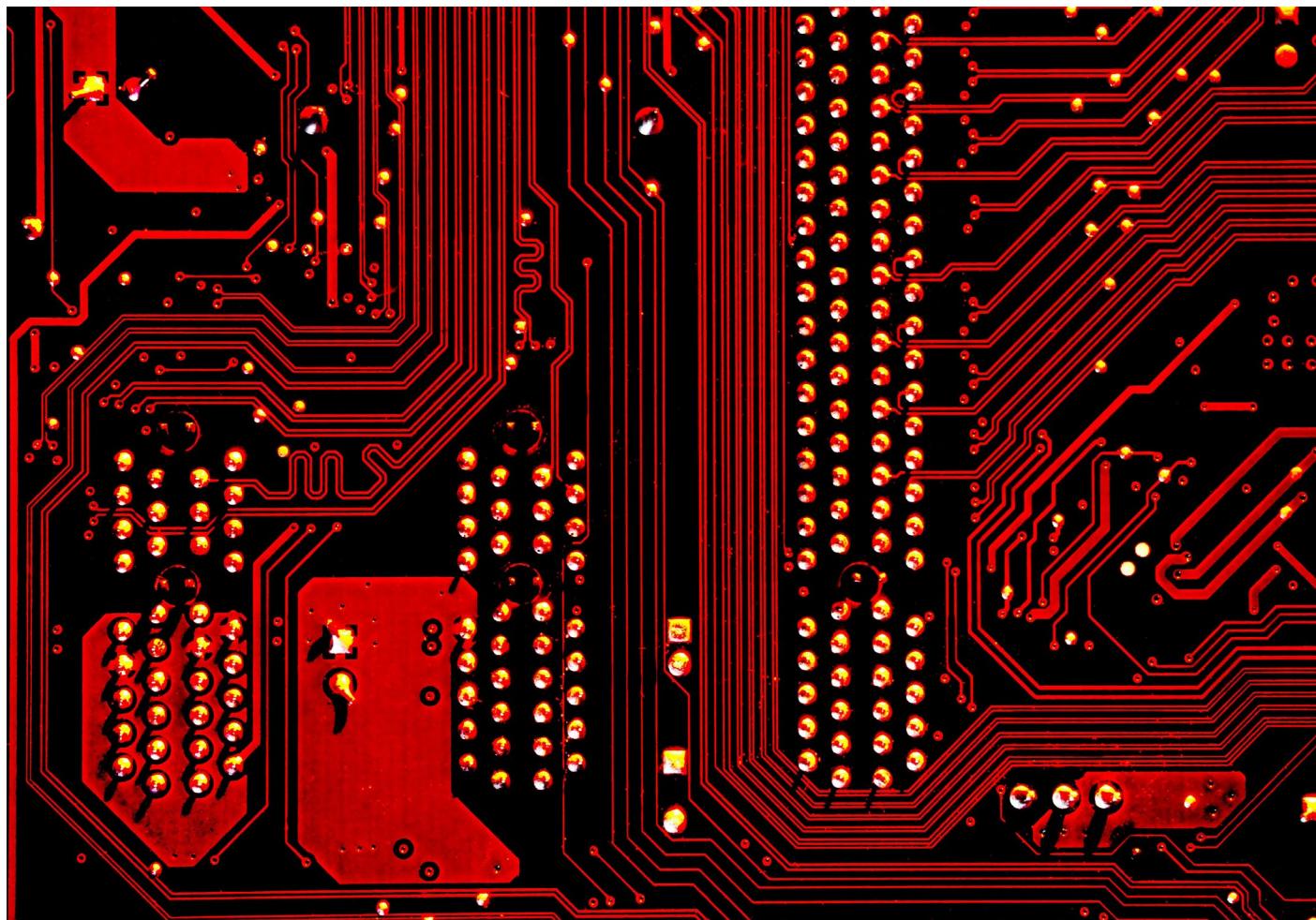
Where and how should you store the data to build AI models on?

Where and how data is stored also contributes to accessibility, security and deployment. Traditionally, organizations stored their own data on local servers, which required significant management of storage facilities, security and access control. With the emergence of cloud services, such as Amazon's Web Services and Microsoft's Azure, organizations have opted to outsource their data storage and computing needs. This has led to lower storage costs, improved security and increased flexibility. While cloud storage is not required for AI development, it improves the organization's ability for self-service. Before adopting cloud storage, however, it is important for the organization's IT and data teams to consider the type of data architecture and tools that best fit their use cases. Doing so will help to determine whether a cloud environment is required for the business and, if so, which provider has the best offering.

There are three types of cloud environments: public, private and hybrid. A public cloud, as described above, entails a third party (such as Amazon or Microsoft) providing storage and computing resources to external parties. All infrastructure on a public cloud is owned and managed by the provider, and users of the public cloud pay service fees that depend on usage. The second type, a private cloud, is when cloud computing resources

(e.g. hardware, infrastructure) are dedicated to a specific organization, within a private network. It can be maintained and serviced by a third party, but ownership of the resources belong to the organization. The third type, a hybrid cloud, combines the attributes of both a private and public cloud. This occurs when an organization outsources to a public cloud because its private resources do not suffice.²⁰⁵

In addition to offering data storage, cloud providers have realized the benefits of allowing access to computing power and development tools. While simple AI applications can be built on a local computing device, such as simple time series forecasting models and image recognition applications, often more complex models that will run on real-time data require additional computing power to meet user needs. Cloud providers have developed solutions to utilize advanced technology through their cloud service offerings. This includes infrastructure as a service (IaaS), which allows external parties to use the cloud providers' infrastructure for storage and computation; platforms as a service (PaaS), which are environments that allow external parties to develop their applications and access development tools (e.g. Azure Databricks); and software as a service (SaaS), which are software solutions that are not locally hosted but accessed through the cloud.



5.1.8 | What types of AI models can you choose from?

“A number of different types of ML algorithms exist, each suited to solving certain kinds of problems.

Statistical and mathematical modelling are designed to fit a variety of objectives, from descriptive to predictive to prescriptive analytics, each one increasingly complex:²⁰⁶

- Descriptive: describing what has already happened. In essence, this is what most organizations have thought of as “business intelligence” for the last few decades.
- Predictive: anticipating what might happen based on statistical probability. For example, some operators of heavy machinery, such as transit agencies, use predictive analytics to anticipate when an asset will require maintenance.
- Prescriptive: recommending a course of action to achieve certain goals. For instance, ride-hailing and delivery service providers use AI to estimate the shortest time between two points, optimizing delivery time.

The goal of developing models in ML is to derive actionable insights to help drive better business decisions. Throughout the model development process, it is important to embed the responsible AI principles discussed in module 4. Ensuring principles guide development from the ideation to the retirement stage of an AI implementation will help mitigate strategic, financial and reputational business risks.²⁰⁷

Selecting the appropriate type of algorithm involves the company’s data scientists and analysts determining how best to solve the business

problem identified, given what is already known and the data held. A number of different types of ML algorithms exist; each is suited to solving certain kinds of problems. A description of some of the most common high-level approaches follows:

- Supervised learning: an algorithm using labelled training data to learn the relationship between a given input and a given output, so that it can then predict outputs given new data. An example is training an image recognition algorithm by feeding it thousands of images of different objects, and telling the algorithm what object is in each image (thus labelling the image).
- Unsupervised learning: an algorithm using unlabelled training data to infer patterns and structure in that data. An example is developing an algorithm to segment customers by providing it with available customer data. It is considered unsupervised as it is not possible to label the data (given that it is not possible to know in advance what the segments will be).
- Reinforcement learning: an algorithm consisting of a simulated environment, an agent and a desired objective. The agent simulates various actions; it is rewarded for actions that accomplish the objectives and/or punished for actions that do not. The goal for the algorithm is to maximize reward, which teaches it to achieve the objective. Note that given its trial-and-error nature, reinforcement learning depends much less on data than supervised and unsupervised learning.

5.1.9

How do you train and test your AI model?

Training an AI model is the process by which an organization provides the algorithm with a subset of its data set (called training data) so it can learn the relationship between the inputs and the target variable (the feature of the data about which an organization wants to gain a deeper understanding). For supervised ML, the training data is labelled with known outcomes (unsupervised learning does not include a target variable).

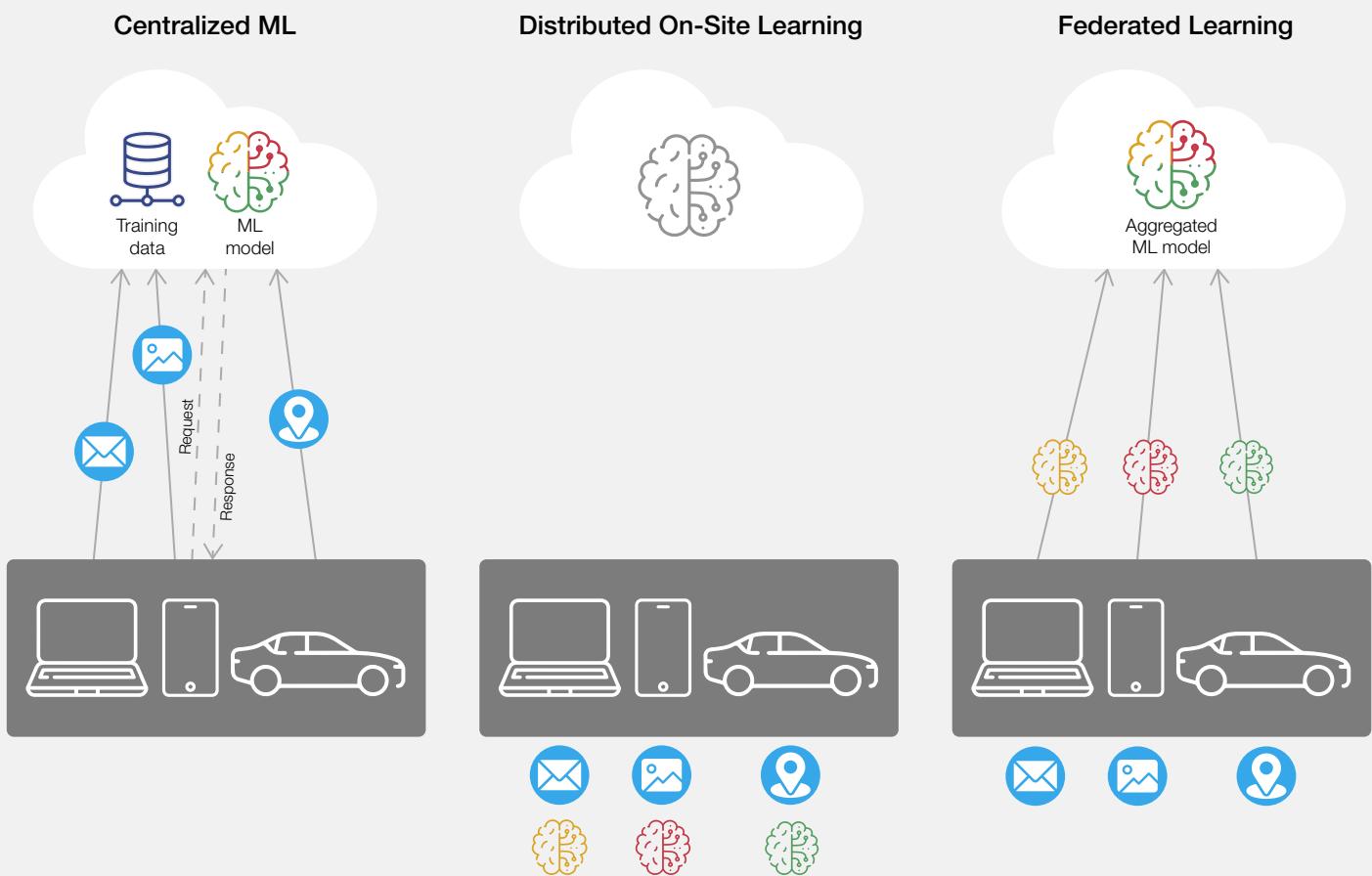
Model training and testing (sometimes called “model tuning”) is both art and science. Data scientists need to make many decisions, including

what model independent variables (also called “features”) to use. The tuning process can be iterative in nature, but tools do exist that help data scientists automate it to some extent.

Centralized vs diffused vs federated learning models

Once buy-in has been secured and the organization has adopted AI, questions related to learning models for AI, privacy and the storage of data on the system and within the system become imperative. The three main models are centralized, diffused or distributed, and federated (Figure 22).

FIGURE 22 | Centralized vs distributed vs federated learning models of AI



In centralized learning (left), data is sent to the cloud, where ML model is built. The model is used by a user through an API by sending a request to access one of the available services. For distributed on-site learning (middle), each device builds its own model using its local data set. After the first interaction with the cloud to distribute a model to the devices, no more communication with the cloud is needed. In federated learning (right), each device trains a model and sends its parameters to the server for aggregation. Data is kept on-devices and knowledge is shared through an aggregated model with peers.

Source: Adapted from Abdulrahman, Sawsan, et al., “A Survey on Federated Learning: The Journey From Centralized to Distributed On-Site Learning and Beyond”, *IEEE Internet of Things Journal*, vol. 8, no. 7, 2021

Each model has benefits, but with cybersecurity and privacy (including regulation such as GDPR) becoming a greater concern, the federated learning model has grown in popularity. Given it is a decentralized approach to training, the raw data is kept solely on the end users' devices. The “learned” data, which would later benefit the larger network, is then aggregated and can be shared with other clients.²⁰⁸

Even with a federated learning model, organizations will need to be clear on what data they are willing to share with the network, how they plan to anonymize it (if they do) and the duration of time it can be shared. Additionally, if a third-party platform is providing the cloud, clarity will be needed on what access the platform provider has to the data, and how this is managed.

5.1.10

Is your model interpretable and explainable?

Once your model starts producing insights, having confidence in its conclusions is paramount. Explainability and interpretability are the concepts used by data scientists and other AI professionals to describe understanding how a model arrives at conclusions and whether those conclusions are trustworthy:

- Explainability: the degree to which a human can understand how a model arrived at a specific decision

- Interpretability: “the extent to which the internal mechanics of an AI system can be explained in human terms”.²⁰⁹

In scenarios where it is important that the end user be able to understand the decision and expect an explanation (e.g. an applicant for government unemployment insurance), the expectation can be that the decision-making process is clear to stakeholders who have no background in AI.

Different stakeholder groups will have different motivations for ensuring explainability and interpretability. System developers will want to verify how an AI system is working and whether it is still producing insights “correctly” from a technical perspective, whereas business users will want assurance that the outputs of the system are trustworthy and that they meet the ethical (and, if applicable, regulatory) standards to which the business adheres.²¹⁰

Both explainability and interpretability are important to ensuring model insights are correct and usable. The more interpretable the model is, the more

confident data scientists and developers will be that it is functioning as intended. The more explainable the model is, the more business users will be able to understand the insights provided by the model and feel confident to act on them. Explainability and interpretability ensure that AI systems are trustworthy and meet ethical and regulatory standards, and that insights are readily communicable to non-technical resources and external stakeholders. However, one needs to also consider potential new risks that explainability may lead to, although it is still too early to clearly understand how explainability may actually affect the behaviour of people and the successful and safe adoption of AI.



5.1.11 How do you choose the right vendors and partners for implementing AI?

As in the implementation of other technologies, AI often requires working with third-party vendors or partners, including entering into cloud, academic, strategy or implementation partnerships. This section discusses vendor and sourcing considerations.

The most important component in finding the right vendor or partner is knowing what it is that your organization is trying to accomplish and what gaps must be addressed in that journey. A thorough understanding of this allows the organization to choose the right type of partner (e.g. does the organization need a cloud provider, academic institution or a consulting firm?) and the vendor that best fits its needs. Depending on this need, the selection criteria can be developed. For example, an organization just beginning their AI journey may wish to get support on building an AI and data strategy and thus seek the support of a consultancy. The assessment criteria for this would depend on the consulting firm’s track record with AI, use case development, experience with implementation, understanding of the organization’s industry and relationship with the organization. For an organization looking to switch from a private to a

public cloud, the criteria would entail understanding data security, accessibility, AI development tools and the possibility of providing training support.

In addition, partnerships may also differ from the traditional paid services model. For example, it is common for consultancies to create co-development engagements in which they help organizations develop the end product but secure the right to redistribute the finished product to others. Such services tend to cost less but do not guarantee ownership of the final product. It is also important to note that some collaborations may not be monetary. For example, if an organization is looking to hire strong AI talent, it may find it useful to build relationships with academic institutions to recruit graduates. Establishing relationships with such institutions enables the organization to source talent and develop internal capabilities.

The AI ecosystem has matured to a point where organizations can find the support they require. Although some vendors provide similar offerings, their differences will likely make one more suitable for your objective. It is important to align a partnership and ecosystem with the AI ambition.

5.1.12

What happens after you've selected and trained your model?

One of the trickiest but most important steps in the AI life cycle is deployment. Deployment is “the method by which [to] integrate a machine learning model into an existing production environment to make practical business decision based on data.”²¹¹ Integration usually “requires coordination between data scientists, IT teams, software developers and business teams”.²¹² Businesses can run into trouble if their IT systems, which often run on legacy software, are incompatible with traditional

model-building languages like Python. Re-coding the model to adjust it may extend timelines for deployment by weeks or months, “forcing data scientists and programmers to spend valuable time and brainpower rewriting them”.²¹³ Successful deployment is essential to drive business value from an AI model, but is also one of the most challenging aspects of an AI implementation. It is critical that integration and deployment be considered at the start of the AI life cycle rather than at deployment.

5.1.13

How do you scale your AI deployment?

Ensuring that the AI solution you are building can scale appropriately is critical to deriving business insights. Scaling AI depends on much more than having the right technical infrastructure in place; it requires a profound change management effort to ensure the organization is ready to leverage AI. Key aspects include:

- Talent and business process change: Ensuring that AI is useful to an organization necessitates that its adoption permeates the organization. Because AI is a relatively new aspiration for the majority of organizations, demand for talent that specializes in AI is very high. This talent will often be difficult to find and more expensive for an organization. In addition to technical or specialized talent, the organization's workforce more broadly may require upskilling to be able to effectively interact with AI or use its insights.
- Technical performance: Since AI models require intensive computer processing, ensuring that the computer storage and processing capabilities within the organization are adequate is critical. For many organizations, AI either acts as a use case for cloud or is fully enabled once they move to the cloud.

- Data volume and access: AI models often require hundreds of thousands, if not millions, of data points to perform effectively. Organizations may face complex technical challenges as they scale their data volume and access capabilities. Established technologies like relational databases, which have been used widely for decades, can be challenged by the sheer volume of data required for AI modelling. Upgrading to architecture that will enable the AI journey is important. Another aspect of data volume that may pose challenges is the effort required to make sure data is reliable. Cleaning and preparing raw data may necessitate new tools or services.
- Data privacy and security: Consolidating the substantial amounts of often sensitive data required for AI in one place can create security vulnerabilities that an organization must anticipate. Mitigating the risk of a security breach or data leak is also made more complex by the fact that many organizations will use vendor services like cloud storage, making them dependent on the security competencies of those vendors. Privacy considerations are also paramount as clients and end-users expect their information to remain confidential.

Ongoing monitoring and feedback

After an AI model is deployed, it must be monitored continuously and tweaked to ensure that the insights a business derives continue to be reliable and in line with the organization's values and ethical principles. The final stages of the life cycle ensure that the model is maintained through to its retirement.

5.1.14 Why do models need continuous monitoring, tuning and retraining?

The insights of many ML models are produced using a constant influx of new data. These models will need careful monitoring, tuning and retraining to ensure that they are still producing the correct insights (e.g. as customer demographics change with time, a customer segmentation model will need to be readjusted). The organization's monitoring team will check for ongoing statistical representativeness, possible biases, feedback loops

and other characteristics that are model dependent. Regular auditing includes the review of underlying assumptions made in early model selection and training phases in light of the current outcomes from the model. The input of new data is likely to require retraining the model over time as data shifts occur. Explainability, interpretability and privacy considerations must be addressed each time the model is retrained.

5.1.15 How do you measure the value of your AI model?

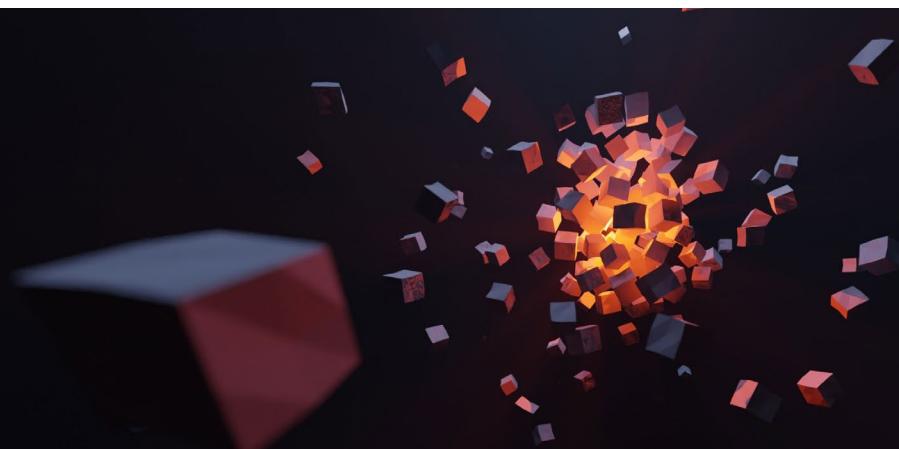
While your data scientists will have specific metrics to measure the technical performance of the model, business metrics should be used to understand the impact of AI on the organization. Often an organization does not need to create new metrics. The business objectives it looks to achieve using AI should correspond to existing business metrics. If the business objectives are brand new, the organization should consider creating new metrics to track these objectives along with the impact of AI.

In the spirit of data-driven decision-making, organizations can choose to leverage data to measure the impact of AI. For example, an AI model used to improve customer retention could be trialled with a subset of randomly selected customers, and results could be compared against a sample where insights from the model were not leveraged. This is similar to the web-design concept of A/B testing.

5.1.16 When should you retire a model?

While not typically included in the life cycle, a key question that business leaders should ask themselves is when to retire an algorithm/system. If the system breaches the organization's stated ethical principles or is no longer functioning properly, it may pose a significant business risk. Mitigating the risk of AI gone wrong could be both costly and damaging to an organization's public

relations. As part of its governance processes and ethical principles, an organization should have criteria by which it evaluates the continued functioning of the system, and a set of mitigation strategies should a system no longer function as intended. These can include a range of actions from taking the system offline to tune and retrain it to fully decommissioning it.



5.2 | Sample case studies of ethical AI

Businesses around the world are reimagining what they can do and achieving it with AI. The following are three real life examples of analytics solutions that have been built and deployed successfully with sound principles in mind.

Case Study 1

Xayn

How to personalize search results with privacy in mind

Start-up Xayn seeks to create a privacy-minded search engine that does not sacrifice the functionality or usefulness of the personalized search.²¹⁴ Launched in late 2020 as a free application, it is currently backed by funding, although this will change as it scales. Its future monetization model looks to combine “business-to-business and business-to-customer monetization similar to Zoom”.²¹⁵

In the current search marketplace, established search engines scrape user data like click history to personalize their search results. They also use that data combined with the search history to serve targeted advertising back to the user. The loss of privacy and commercialized aspect of these services have worried consumers, but few attractive alternatives exist.

Privacy-protecting search providers like DuckDuckGo do not store any data about the user but, as a result, cannot offer attractive personalization features that users may want.

How a privacy-centric search engine works

Xayn is built on a “custom open-source algorithm supported by four AI models focused on language processing, interest clusters, domain preferences, and personalized contextualization”.²¹⁶ The search algorithm in the platform’s feed learns by combining the encrypted AI models trained by each individual user and synchronizing them into a collective model that is then fed back to the user’s device.²¹⁷ No usage data or results are shared with the company itself, while users enjoy the benefit of smart and personalized search.

Case Study 2

A major bank

How to de-bias data to prevent poor AI decision-making

A major North American bank wanted to use AI to automate the granting of additional credit to qualified applicants. As a result of clear regulations in the financial sector regarding the characteristics that can be used to evaluate a customer for granting credit, the bank had to consider several ethical design principles during the AI life cycle, including fairness and user data rights.

The bank used a credit science algorithm to automatically determine the eligibility of each client, which is standard practice in the industry. However, training the algorithm on existing data could create a bias against populations that are under-represented in the data, potentially affecting the ability of minorities to receive automatic credit upgrades.

Two techniques were leveraged to remove data “blind spots”:

- Random testing to ensure all data had the equal probability of being selected for algorithm training
- Over-weighting methods, such as bootstrapping, to ensure the algorithm became familiar with under-represented groups and assessed applicants belonging to such a group fairly.

The bank ensured that its data engineers and scientists were aware of the structural data biases that could exist and knew how to identify and address them as well as how to extract actionable unbiased insights from the algorithm going forward.

Case Study 3

A large retail lending service provider

How an organization can ensure data security when using sensitive data in AI modelling

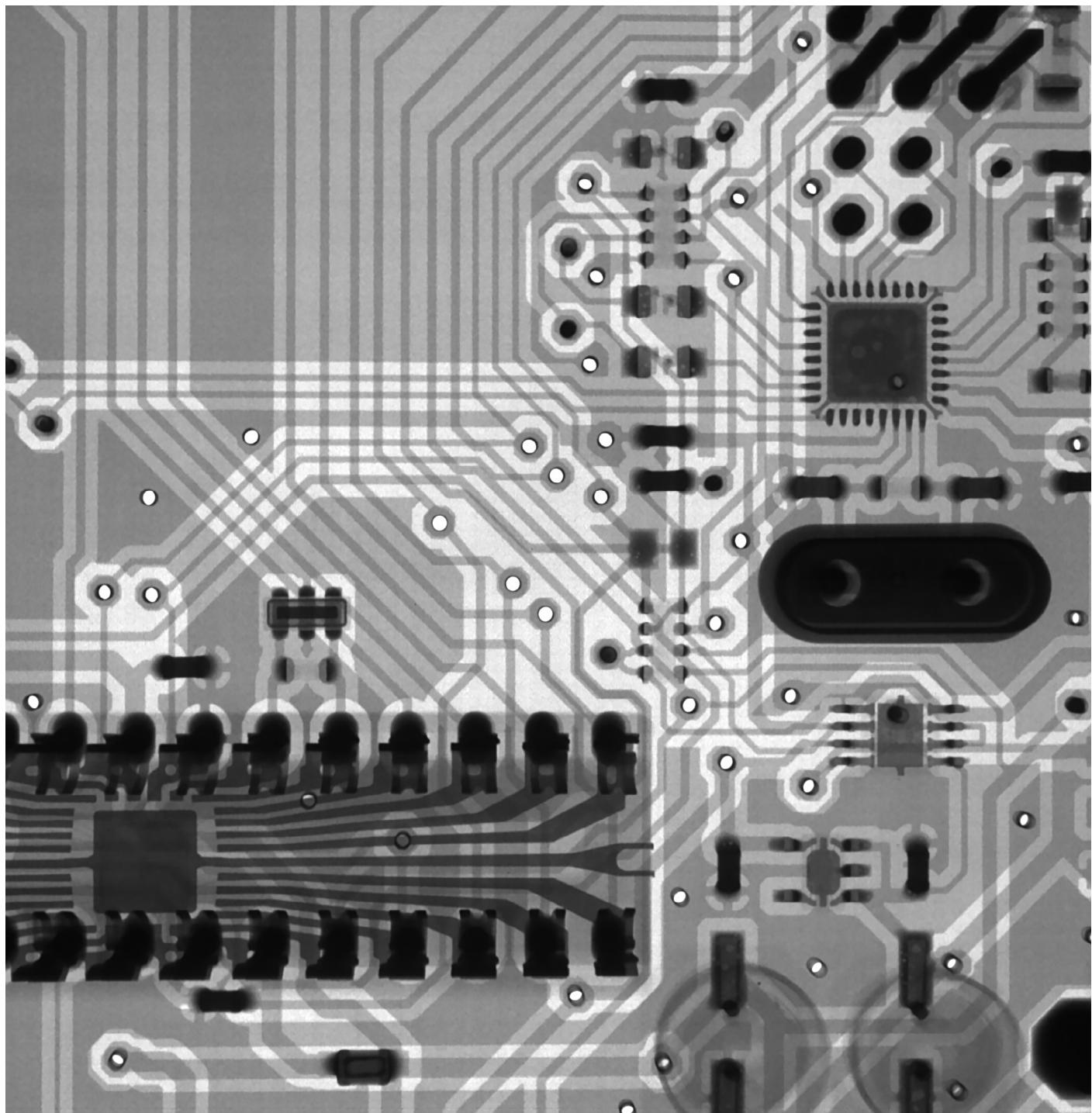
A large retail lending service provider in North America wanted to use AI to gain a deeper understanding of its customers by analysing their transactional data. Since transactional data provides very private details about an individual’s lifestyle and preferences, the provider had to carefully consider how to protect the security of this data when designing and training its algorithm. Studying granular transactional data could expose private customer details to those who were creating the AI model, breaching their privacy.

To ensure that the customer data was protected, even from the team building the AI model, anonymized data was used in development. This ensured individuals’ privacy and kept their lifestyles and preferences private.

Another privacy-minded technique the service provider employed was using aggregate-level data when training the model. This prevented discrimination against borrowers whose transactions could reveal personal distress (e.g. ambulance charges) or addictions (e.g. frequent lottery, casino or liquor store purchases).

AI in sustainable development and industrial AI

Companies can leverage AI to optimize their use of resources and minimize their environmental impact.



6.1 AI, sustainability and the C-suite

6.1.1 Why and how should you capitalize on the opportunities of sustainable development?

 Risks and rewards are motivating companies to transition to more climate-centric practices.

As the anticipated consequences of climate change become more tangible, activism from the public sphere will continue to influence calls for a clearer picture of how the economy is affecting the sustainability of our planet.²¹⁸ Companies will likely feel increasing pressure from both the individuals that support them and the bodies that regulate them:

Bottom-up: Many consumers have demonstrated an important trend to direct their income towards more sustainable products and services. Research conducted at the New York University Stern Center for Sustainable Business found that sales of “sustainability-marketed products grew over seven times faster than conventional products, demonstrating consumers’ strong preference for these products and their willingness to pay higher prices.”²¹⁹ In addition to this consumer influence, companies unable to meet employee expectations risk potential talent shortages and reduced employee engagement. Internal activism is prompting employers to support the sustainability transformation, materially shaping an organization’s strategy and reputation as staff hold their employers accountable to a degree never witnessed before.²²⁰ One example is the Amazon “Employees for Climate Justice” movement.

Top-down: Non-governmental organizations have independently produced a number of frameworks and standards for corporations to seamlessly adopt ESG reporting. Nevertheless, at the International Business Council 2020 Summer Meeting, the five leading parties for sustainability accountability issued a statement of intent to co-deliver “a

single, coherent, global ESG reporting system” to achieve sustainable value creation.²²¹ Their intent will be catalysed by governments, regulators and accounting bodies in favour of streamlining the transparency of this non-financial disclosure. Once implemented, this novel reporting system will help federal regulatory bodies to standardize their approach to shaping a more climate-friendly business ecosystem through regulation.

Impact: The shift in consumer preferences and regulatory burdens has the potential to impose substantial risks on businesses. Increasing costs, suppressing existing products or services, and/or materially affecting asset values are possible consequences of these newfound risks. The French Government used its regulatory authority to push companies to adopt sustainable practices in 2017, enacting “new regulations requiring that French multinational companies identify and prevent adverse sustainability impacts resulting from their own activities and the activities of their subcontractors or suppliers.”²²² Another illustration of this movement comes from supranational organizations such as the World Bank Group that have encouraged governments to distribute COVID-19 stimulus packages to foster sustainable growth.²²³ These are just two examples to illustrate the risk and rewards motivating companies to transition to more climate-centric practices. The businesses that are able to address changes in regulation proactively will create the best opportunity to maintain or improve their competitive positioning in their domestic market while opening a path to a strategic advantage on a global scale.

6.1.2 What shift in paradigm is required to maximize benefits and mitigate risks?

As investment strategies and the guiding hand of governmental bodies steer companies in the direction of more sustainable practices, new challenges and opportunities will be uncovered, revealing novel avenues for growth and competitive advantages for those willing to advance in this area. The companies that will benefit the

most from the transition to the new business reality will be those able to proactively shift their thinking on typical business orthodoxies. Three paradigm shifts can help companies take advantage of the benefits derived from traditional and AI solutions related to environmental sustainability, outlined by Deloitte as follows:²²⁴



- 1. Company-scale action: Change how you capture value by rethinking relevant priorities and time horizons.** The shift to a more climate-friendly world requires organizations to consider their initiatives' cost-benefit to the triple bottom line – the value delivered to stakeholders, the planet and profits. Although this form of action is important for all industries, new research suggests that organizations producing greater proportions of carbon emissions may be more sensitive to future carbon costs. On the other hand, carbon efficiency can be synonymous with resource efficiency. This is something for all organizations to consider as an opportunity to boost financial performance and reduce risk.²²⁵ Companies seeking short-term benefits from initiatives with detrimental effects on the planet will be left behind when these same initiatives preclude them from reaping the rewards of longer-term, sustainable strategies.
- 2. Industry-scale action: Change how you create value by changing relevant constraints.** As more companies make changes to their business to accommodate new regulations and external pressures individually, opportunities for industry-level advancement will increase through associations to address sustainability challenges on a wider scale. This requires companies to evaluate the full environmental impact of their industry and hold each player accountable to achieve greater results than any one company can

achieve on its own. The US farming sector serves as an example. A growing movement called regenerative farming has the ability to transition the agriculture industry from being responsible for approximately 10% of annual US CO₂ emissions to becoming a positive force by sequestering emissions through innovative practices. Although this requires businesses to sustain a reduced profit ceiling in the short term, once these regenerative farms have matured (generally a two-year process), farmers have observed substantial reductions in operating costs that override their losses in the long term and increased profitability. They have established the Ecosystem Services Market Consortium, a national market to sell credits for greenhouse gas (GHG) reduction, which are redeemed through an IoT and computer vision system in a digital marketplace where buyers can purchase carbon credits from the farmers directly.²²⁶ When multiple players of an industry align to prioritize changes that can make a positive impact on climate change, the payoff for the planet is exponential. The advertising industry has shown that staying ahead of regulatory bodies to make an impact on climate change has been a fruitful strategy, reinforcing the bond between the advertisers, their workforce and their consumer base.²²⁷ All players in any given ecosystem should feel empowered to discuss and negotiate specific parameters that could create a fair result for the industry while promoting a positive impact on climate change.

3. Ecosystem-scale action: Change how you define “value” by changing your objective.

The concept of sustainable value creation is increasing in prominence, implying the importance for a company to consider its ability to create value over the long term. The emergence of this concept serves as a check and balance to ensure an organization's existence is not jeopardized by sustainable business practices. Sustainable value creation requires an entirely different approach to decision-making since the objective shifts from typical metrics, such as immediate profit. “While there are many instances where climate-friendly practices are also good business, there are also a set of actions that may be costly – at least by conventional financial measures . . . The costs of inaction – and the loss of a societal license to operate – are likely to be orders of magnitude beyond any near-term outlays”²²⁸ and are the risks an enterprise may face when considering their positioning in the context of sustainability

if they are not able to adapt in line with the pressures inherent to climate change.

Once these shifts in paradigms are achieved, businesses open themselves to a new foreground of opportunities and challenges. Business executives are challenged to become fully aware of the environmental, social and regulatory factors that may influence their organizations. As corporate sustainability strategies and initiatives are imagined, planned and deployed to tackle these concerns, the teams in charge must have a clear picture of all the tools available to them to best address their specific issues and opportunities. That's where AI comes into focus. The following section reviews the latest AI trends and highlights the opportunities and areas of oversight where AI can play a role in promoting sustainable development. For the purpose of this topic, AI should be regarded broadly as cognitive technology applications and advanced analytics systems.

6.1.3

What are the latest AI trends to support sustainable development?

*This section is based on the World Economic Forum's [Empowering AI Leadership Toolkit for Boards of Directors](#) with updated content to benefit the C-suite.

AI can not only help reduce a firm's impact on climate change and optimize the use of natural resources but can also reduce AI's own carbon footprint.

AI is a transformative technology that helps advance diverse operations in numerous sectors through cognitive applications, advanced analytics and intelligent reporting systems. Companies have used AI to re-engineer processes in many business and government functions. For example, AI has been leveraged to augment workforces with cobots (a collaborative robot), to hyper-personalize customer engagement and ultimately to create better products. With such prevalence and widespread effect, the question to answer becomes how AI affects the environment. Can AI be the catalyst that will break the link that was previously required for all past industrial revolutions to succeed: the link between economic growth and environmental degradation? AI can deliver advances in water and energy conservation among other use cases, but it is crucial to recall that AI comes with its own carbon footprint.²²⁹

Companies are using AI to:

- **Reduce their environmental impact.** A 2019 study by Microsoft and PwC estimated that the “responsible use of AI can lead to a 4% (2.4 giga tonnes) drop in worldwide GHG emissions by 2030”.²³⁰ To reduce a firm's impact on climate change, AI has proven valuable in optimizing carrier routes, managing autonomous transportation, and predicting and forecasting supply/demand and preventative maintenance schedules. Although the intentions of these actions are typically related to reducing operating costs through newfound efficiencies, these optimizations can lead to positive environmental impacts in the elimination of waste and reduction of emissions.

- Optimize the use of natural resources.

A recent World Economic Forum article reports that AI is helping companies forecast “the output of energy generated by such green sources as solar, wind and hydro-based energy, thus ensuring minimal waste of these natural resources. AI helps conserve water usage in residential, manufacturing and agricultural areas. Predictive AI algorithms have developed new agricultural processes such as precision farming, ensuring that the exact amount of water required is used and only ripe crops are picked. Algorithms also assist in farmland planning, monitoring the health of crops and livestock”.²³¹ AI also plays a vital role in power management and distribution, enabling smart grids to advance their efficiency, transparency and promote the use of renewable energy systems.²³²

- Optimize the usage of AI to reduce AI's own carbon footprint.

A 2019 study produced by the University of Massachusetts Amherst suggests that training a common large AI model would emit the carbon dioxide equivalent of the lifetime impact of five American cars.²³³ Companies are beginning to counteract this impact by using green renewable energy to power their AI models and are starting to factor the environmental effects into their cost/benefit analyses for deploying AI selectively and responsibly.²³⁴ As an example, tech firms (Microsoft, Google, Amazon, Apple and Amazon, among others) have committed to plans for carbon neutrality or even sub-neutrality to align with public regulations and social expectations related to the environment.²³⁵ To support data

teams in analysing and optimizing the CO₂ footprint of their computing, a partnership of top AI experts (Mila, BCG GAMMA, Haverford College and Comet.ml) launched an open

source software package called CodeCarbon. Leveraging tools like this is the first step to understanding where improvements can be made, and how to best address them.²³⁶

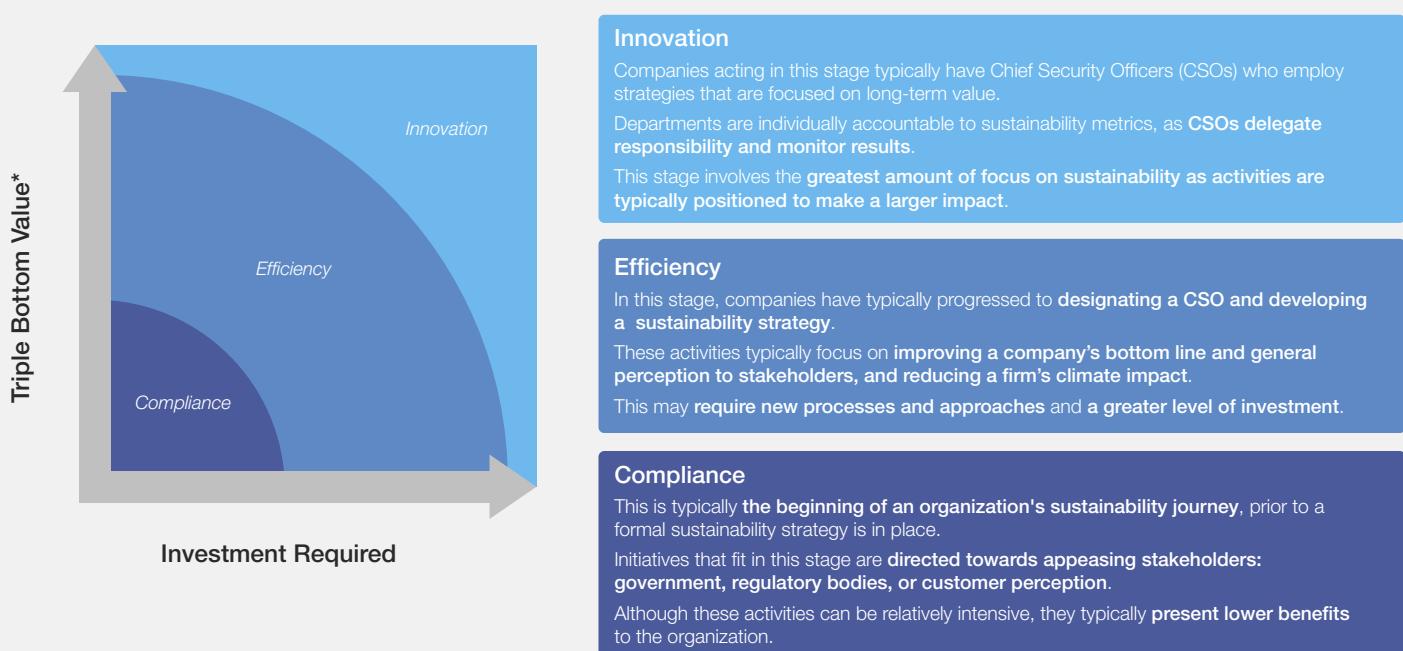
6.2 Three lenses for AI in sustainable development

6.2.1 How invested should your organization be in sustainable development with AI?

Concerning the specific sustainability strategies and initiatives that can be deployed, three stages of sustainability in which an organization can engage are noteworthy, as described by Dina Gerdeman in an article from HBS Working Knowledge.²³⁷ Depending on a firm's level of commitment and priorities, its reason for acting on sustainability is likely to fall somewhere on the spectrum of compliance, efficiency and innovation.

Activities that relate to each of these stages have their own justifications and potential benefits. However, it is important to point out that an organization may engage in activities that relate to different stages at the same time. Figure 23 describes the stages and the relationship between the effort required and the value gained for companies engaging in sustainability activities.

FIGURE 23 Three stages of sustainability for companies



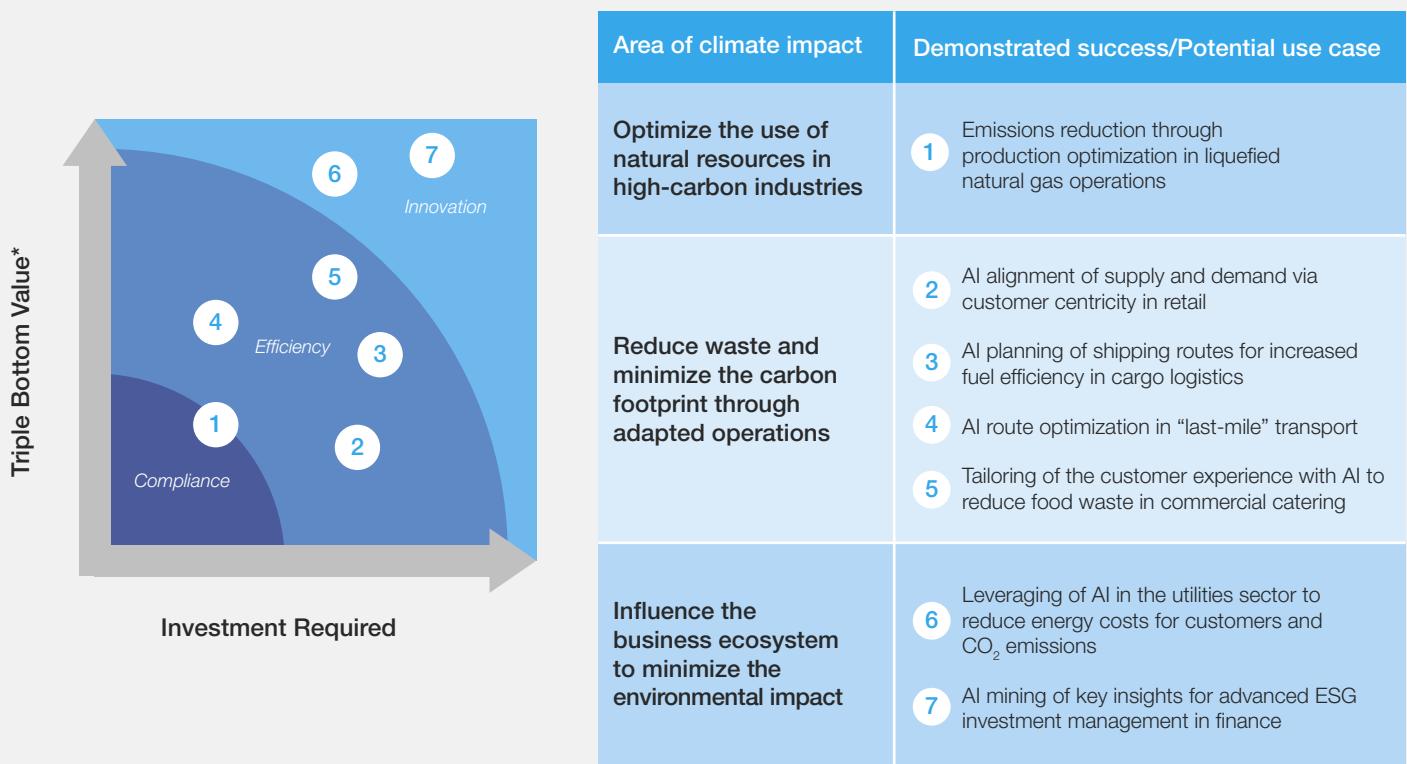
Note: *Total value increase for profitability, people and climate action

Source: Deloitte, based on content in Gerdeman, Dina, "Who Is the Chief Sustainability Officer?", Harvard Business School, Working Knowledge, 8 October 2014

Figure 24 depicts different types of specific AI use cases in relation to the investment required and the value derived.

FIGURE 24

AI use cases in the context of the three stages of sustainability action



6.2.2

Certain successful AI use cases depicting the three stages of sustainability provide good examples of the opportunities available within an organization and across the value chain of businesses.

What are some examples of organizations practising sustainable development with AI?

Compliance

- Emissions reduction through production optimization in liquefied natural gas operations
- As a larger contributor to climate change, oil and gas is an industry where emissions are directly linked to operations. However, the magnitude of impact has been found to be variable through innovative practices. One enterprise has demonstrated company-scale action by reducing their operations' CO₂ emissions through an AI-powered tool that delivers increased efficiencies through combining engineering expertise with critical insights from data to optimize the state of the refrigeration units.²³⁸ With economic, operational and engineering constraints considered, this tool equips asset engineers with a digital platform to troubleshoot and make operational decisions. The results of this tool have led to a boost in production capacity while empowering management to operate within an environment of constricted climate regulations.

Efficiency

- AI alignment of supply and demand via customer centricity in retail
- A \$2.5 trillion sector of retail classified as fast fashion has been criticized for its intensive environmental harm, as it is responsible for about 10% of the world's CO₂ emissions and results in massive amounts of ocean pollution.²³⁹ However, companies are beginning to initiate tactics to become more sustainable and have elected AI as an important tool to aid them on this journey. One of the most impactful applications of AI in this sector is to align their production to precise demand forecasts for particular products, as this can reduce waste across their entire value chain from production to distribution. This has been noted as a potential win-win opportunity to create more relevant offerings for their customers, while using less energy and producing less waste across the industry.²⁴⁰

3. AI planning of shipping routes for increased fuel efficiency in cargo logistics

Container shipping comprises approximately 3% of global CO₂ emissions.²⁴¹ Among other tactical solutions like “slow steaming”, some shipping companies have begun investing in technology to support them in achieving carbon neutrality. One example of the solutions deployed is leveraging AI to support captains in selecting the most fuel-efficient paths for transporting their cargo. The suggested routes are informed by dynamic factors such as wind, subsea currents, the squat effect (when fuel is consumed less efficiently because of shallow waters) and other environmental factors. When these factors are considered in sum, the optimal route is often much different than a straight line from point A to B and therefore produces substantial reductions in emissions.

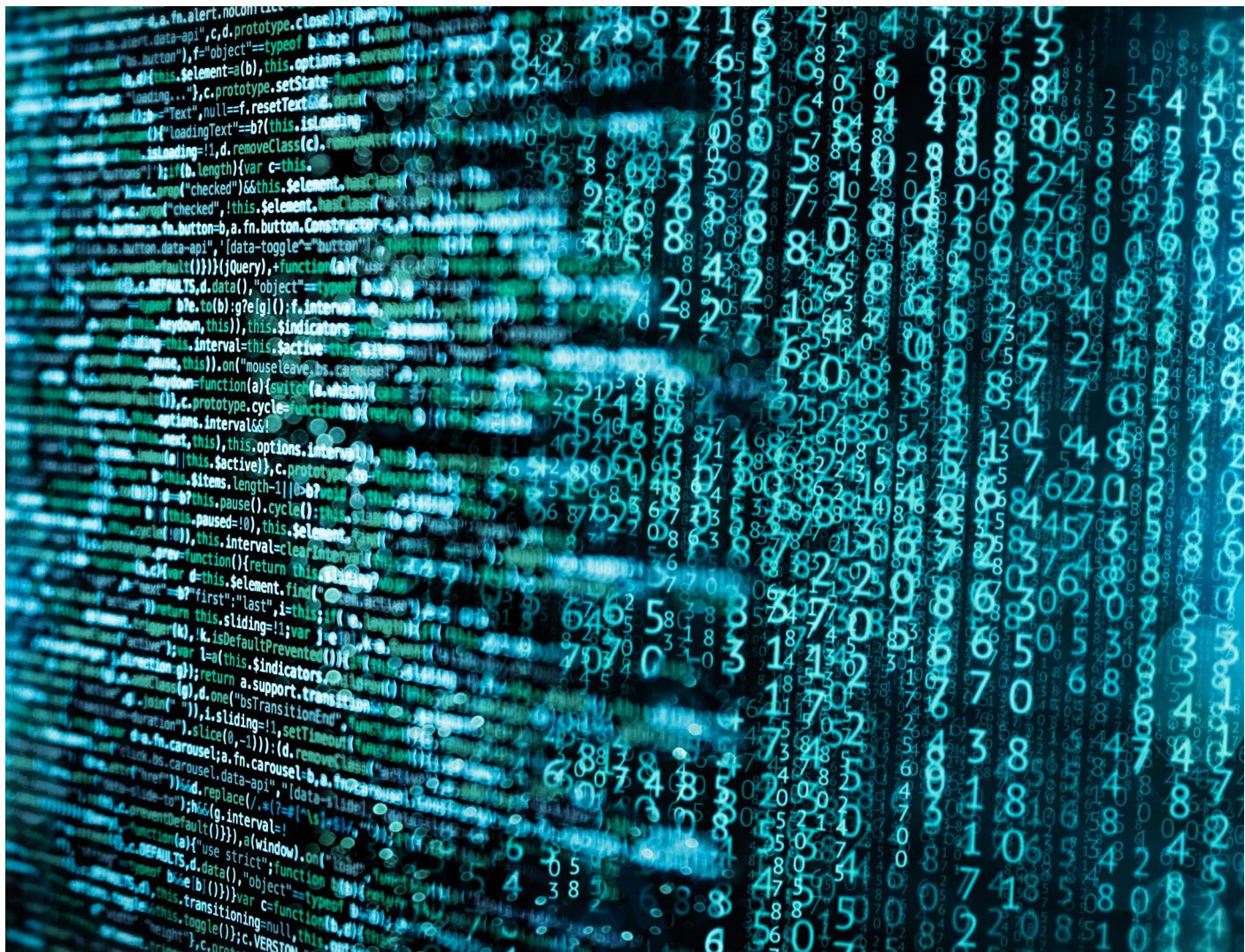
4. AI route optimization in “last-mile” transport

Last-mile delivery is an industry that is growing in size and carbon impact as e-commerce transactions become increasingly prominent, resulting in deliveries from distribution sites to the end consumers. A growing trend in reducing operational costs and combating the environmental impact of these deliveries has been to deploy digital platforms to increase the efficiency of their routes.²⁴² Drivers can be provided with dynamic

route selection support covering the entire cities or regions that they are responsible for. For drivers making over 100 deliveries a day, their greater efficiency on each trip compounds to yield significant benefits for both their productivity and fuel efficiency. One company that has implemented this type of technology has saved approximately 100 million miles and 10 million gallons of fuel per year since deployment.²⁴³

5. Tailoring of the customer experience with AI to reduce food waste in commercial catering

After finding that food waste accounts for about 1.6 billion metric tons of waste globally each year,²⁴⁴ one company set out to tackle this issue by surfacing actionable sensory data on their customers’ eating preferences. Once a customer’s preferences are understood, meals can subsequently be tailored to their preferences and habits to reduce the percentage of waste for each order. The company intends to use new technology so the right amount of food is produced at the right time. If executed successfully, it estimates that it could reduce food waste by at least 10%, which is equivalent to saving over \$200,000 every year. As projects like this are adopted at scale across industries, leading firms combat global sustainability challenges while positioning themselves to increase their customer base and develop new revenue streams.²⁴⁵



Smart metering
in the utilities sector
could reduce CO₂
emissions by 77% by
2035 in America.

Innovation

6. Leveraging of AI in the utilities sector to reduce energy costs for customers and CO₂ emissions

The utilities sector is an area that lags behind in innovation due to barriers to change, even when the leading solutions' benefits have been proven. Smart metering in this sector has been proposed as an important solution to help meet climate goals globally, as this technology could reduce CO₂ emissions by 77% by 2035 across America.²⁴⁶ Pioneering companies have demonstrated the value of smart energy grids through the deployment of connected meters, automation devices and sensors across their distribution grid. These enhancements provide a foundation for an innovative distribution management system and situational awareness tool used to monitor usage and reduce waste. They yield significant benefits for the company and customers. One implementation example resulted in savings of \$30 million in operational costs, and further savings generated from 170,000 fewer field visits needed to attend to the grid.²⁴⁷ These savings translated directly to reduced bills and better support for their customers. This innovation segment will continue to progress as governments are challenged by more frequent extreme weather events.

Another innovative company addressed its customers' substantial energy losses from radiators. By implementing sensors and advanced algorithms, it was able to help its customers identify leaks in the system. This allowed them to pinpoint overconsumption in individual radiators that return water with an excessive temperature, indicating a leak or malfunction. Fixing these types of leaks can help lower the required temperature in the system and recover large energy losses. According to the

company, "A reduction of just 1 degree Celsius corresponds to the annual consumption of 10,000 homes or 9,500 tonnes of saved CO₂".²⁴⁸

7. AI mining of key insights for advanced ESG investment management in finance

The finance sector has been influential in encouraging firms to adopt more climate-conscious practices as they direct greater proportions of their capital to sustainable businesses. Proactive investment management firms are beginning to deploy AI to better apprise how to distribute their funds. With the goal of mitigating risks associated with climate change, such as direct costs arising from increased insurance costs, the vulnerability of assets to extreme weather events, potential litigation from environmental damage as well as indirect costs, consumer preferences for business-to-consumer and business-to-business firms are shifting to a more climate-centric stance.

ML has been leveraged to assess the impact of pollution using satellite imagery for their prospective investments. In doing so, companies are able to attribute the ESG impacts to the granularity of cars in an individual parking lot. Natural language processing (NLP) has also been deployed to extract informative data on selected topics from public news and published company reports. Once this information is scanned with an NLP program, findings to support their funding decisions are automatically brought to surface. Business executives in the finance sector have spoken to the overabundance of ESG data available and the importance of deriving insights from these applications to support firms in selecting where to allocate their funds or, more broadly, where to focus their attention.²⁴⁹

6.3 Taking action

*This section is based on the World Economic Forum's [Empowering AI Leadership Toolkit for Boards of Directors](#) with updated content to benefit the C-suite.

6.3.1 What actions can your organization take to implement sustainable development with AI?

This section outlines two tools for executives planning to leverage AI responsibly for sustainable development.

Appendix 1: Preparedness Assessment Tool

Objective: Supporting C-suite executives' assessment of their firm's competency to deliver on sustainable AI initiatives while providing an opportunity to take stock of their organization's approach to potential impact areas

The Preparedness Assessment Tool helps executives assess whether they possess, or have access to, the knowledge required to independently judge decisions on using AI to sustain the environment. The tool prompts executives to locate where internal responsibility is held, and provides an opportunity to take stock of their current approach to addressing the potential implications of AI in the sustainability topics their organization may face. This exercise is key as boards begin holding their executives accountable for their ability to be proactive in considering these sustainability topics. The Preparedness Assessment Tool is available in Appendix 1.

Appendix 2: Guidance Tool

Objective: Supporting C-suite executives in their efforts to address potential sustainable development competency gaps revealed in the Preparedness Assessment Tool, through an “if, then” approach

After completing the Preparedness Assessment Tool, the same individuals are recommended to review Appendix 2, the Guidance Tool. This tool offers suggestions for further action in an “if, then” format and can be used to brainstorm to fill any gaps that may have surfaced during the initial exercise. The Guidance Tool is available in Appendix 2.

6.3.2 How do you set the agenda on sustainability through AI?

The following suggestions can help individuals who prepare executive-level discussions and set the agenda on sustainability through AI. Before leading the first meeting:

- **Prepare yourself:** Become familiar with AI, what it can do today to support sustainability, and what it will be able to do in the future as the field advances. Separate the hype from reality by looking at the research and the sources behind the claims, and the issues that complicate the implementation of the technology. Speak to senior IT, security and public affairs executives about any ethics issues on their minds.
- **Gauge executive interest in AI and sustainability:** Speak to other C-suite members. Learn what importance they place on AI and the concerns they have about planned AI investments and partnerships. Identify the board members who are most interested in moving forward with new AI investments, and those who have concerns or lack interest.
- **Set goals:** Think ahead about the desired outcomes from the board discussion. Set the initial agenda. Create a strategy for promoting the use of AI to improve sustainability efforts.
- **Design the presentation:** Arrange for a briefing on how AI is being used to improve the organization’s sustainability efforts. The presentation can include examples from competitors and potential use cases uncovered by researchers. It should include cost savings and other quantified benefits when possible. The presentation should also introduce major risks and responsibilities that the company will have to manage, and the requirements that must be met to run AI, such as the data for training AI systems.
- **Prepare the discussion:** Identify and prioritize relevant sustainability areas for pilots, based on high potential reductions in the carbon footprint; high potential optimization in natural resource usage; the availability of data; and the ability to implement and scale up if successful.
- **Delegate:** Decide which members of the executive team will be responsible for selecting and running the pilots as well as deciding what support is needed (e.g. technology, development platforms, innovation sandboxes etc.).
- **Engage:** Decide how the board will stay abreast of developments in sustainability innovation. Set follow-up agenda items. These can include:
 - **Sustainability awareness and culture:** Discuss how the company is developing a culture that supports a sustainable business model. This conversation can include current and future key performance indicators, sustainable practices assessments and rewards for pioneering sustainability initiatives.
 - **Renewable energy:** Discuss how the company is prioritizing the utilization of clean energy sources and how it is investing in renewable energy R&D.
 - **GHG standards:** Discuss which GHG protocol (corporate, value chain or product) standard can best support the company’s missions and goals for measuring and reporting emissions. Map and monitor environmental impacts across the value chain to understand which processes are (or can be) supported by AI investments.

6.4 Industrial AI

While the questions in the earlier modules of this toolkit are applicable across any mainstream AI and industrial AI implementation, this section covers further key considerations specific to industrial AI.

6.4.1 Introduction

Increases in computing power and storage along with new algorithmic advances have brought AI to the shop floor. A report by McKinsey with responses from 2,395 participants from a variety of industries shows that 20% or more of organizations' earnings before interest and taxes (EBIT) are invested in AI.²⁵⁰ AI offers the industry great potential, potential that should be strategized and handled with delicacy. Industrial AI is defined as "a systematic discipline which focuses on developing, validating and deploying various machine learning algorithms for industrial applications with sustainable performance".²⁵¹ This definition distinguishes itself from general AI on matters of infrastructure, data and skill sets.

The infrastructure of industrial AI lies in the domain of operational technology (OT). The information technology (IT) and OT domains are commonly confused. OT deals with data flow in the field and controlling and monitoring physical processes, whereas IT deals with data flow at the business level. Mainstream AI is concerned with building computerized systems that mimic human intelligence. In contrast, industrial AI uses technology to bring relief to the choke points of the industry, which could range from customer value creation to predictive analysis.

Industrial AI usually seeks solutions tailored to specific issues and, as a result, deals with reducible feature-rich data. The nature of the industry demands fast results and high accuracy. When combined with specific and associated data, this condition shrinks the possibilities of applicable algorithms. Furthermore, governing data is the major concern of implementation. Data lakes are repositories where data is collected, organized and analysed to generate higher value data that is later infused into an application area to achieve corporate objectives.

According to a recent *IEEE Access* article, "With the recent developments in information and communications technologies, particularly regarding IoT, big data and cyber-physical production systems, it is now feasible to implement the necessary flexibility, responsiveness and intelligence to face [the challenges of the industry]. . . . In the context of the Industry 4.0 paradigm, AI is being regarded as one of the key technologies to achieve these capabilities",²⁵² such as self-awareness, self-optimization, self-prediction, reliability, resilience and collaboration. Although the expectations for industrial AI applications are huge, enterprises must be able to manage the challenges to achieve successful and sustainable deployments.

What are the challenges to achieving the successful and sustainable deployment of industrial AI implementation?

Four dimensions of challenges can be discussed: data availability, data quality, governance, and security and privacy.

Data availability is the primary challenge; going beyond the pilot stage is only possible if sufficient data is available to train and validate the models. Data gatherers and providers are responsible for this stage.

The second challenge is data quality. To deliver useful results, industrial AI and ML models rely significantly on reliable, clean as well as frequently and correctly labelled or contextualized training data, making data quality a crucial component for the industrial success of these technologies.²⁵³ Data assessors are responsible for data quality in terms of intrinsic quality, contextual relevance, representational success and accessibility.

The third challenge is the governance of data among complex flow scenarios and multiple sources. According to data analysers and output generators,

the development environment should be smooth and accessible in an iteration cycle with continuous integration and deployment. They are in constant communication with data providers and data assessors during prototype and validation processes.

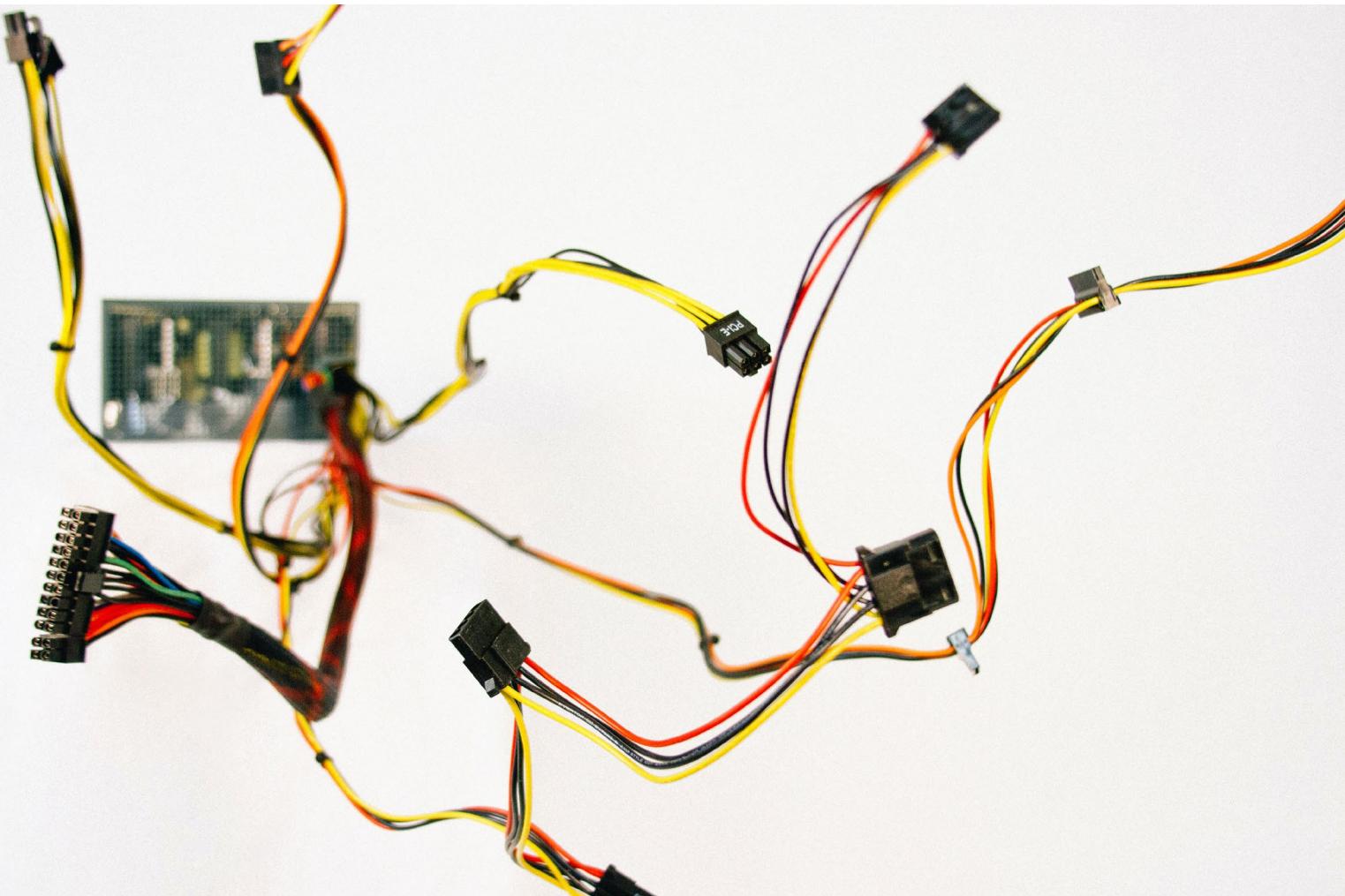
The fourth challenge is security and privacy. Focusing on them through penetration tests for volatile scenarios and defining the current vulnerability of systems to ease integration will greatly benefit enhancing data privacy and security. IT and cybersecurity experts are key people during the processes, assisting legal consultants to create regulations and data privacy protection for data flows.

A multistakeholder approach is at the heart of an implementation that aims to reap the full benefits of industrial AI. The project stakeholders include the data gatherers (operators), data providers (digital delivery managers), data assessors (domain leaders and data scientists), data analysers

● **Industrial AI uses technology to bring relief to the choke points of the industry, which can range from customer value creation to predictive analysis.**

(statisticians and data scientists), output generators (AI specialists), integration of workplace safety experts, legal consultants for regulation assistance and C-suite executives. These actors are critical to the development of an AI-enabled mindset in the implementation phase. Hence, creating an effective

organizational design and enabling communication between the stakeholders are key for business management. It is critical for people with intervention power to be involved in the effort. Increasing the sense of ownership greatly facilitates managing stakeholders from data sourcing to AI modelling.



What are the greenfield and brownfield environments for industrial AI?

Greenfield refers to software that is created from scratch in a completely new environment, whereas brownfield refers to any type of software that is built from old systems or coexists with current applications.

Brownfield development forces companies to live and work in a confined environment due to inherited hardware, embedded software and design constraints. It is expensive to dispose of all the legacy materials. Manufacturers are not ready to repeat that cycle for the sake of connectivity because some of it has decades of history, testing and implementation behind it. Manufacturers can rework their legacy systems while considering the serviceability of the original equipment, the cost of replacement and expected downtime losses. Retrofitting is a viable option when the current set-up can be maintained for the planned horizon with minimal effort.

Greenfield development has no limitations imposed by old systems or obligations to integrate with other legacy systems. Since manufacturers are entering an unexplored area with greenfield developments, the risks are considerably high even though the process is straightforward. As there is no obligation for integration with legacy systems, modularity and flexibility are critical and are positive features of the new systems. The extent of possibilities in terms of integration cannot be foreseen beforehand and the new system can quickly become obsolete just like the legacy system that was replaced by the greenfield approach. Taking into account that the legacy system being replaced was one of the best possible solutions at the time, the greenfield solution will then become the next legacy system. The lifespan of the new system will be determined solely by the flexibility and modularity that it provides.

6.4.2 Strategy

What is your action plan?

High-quality, on-demand and faster output generation industrial systems can be realized through industrial AI. Although reaping the full benefits could be achieved through correct business management and a clear roadmap, industries fall short on integration.²⁵⁴ The first step to overcome this roadblock is a gap analysis, a decision-making aid that helps measure the gap within the target and the current level of the project. The comparison of the present state and the ideal state highlights the areas of development and the opportunities for growth in the future.

To help create a framework for AI, the availability of data, skill sets and equipment should be taken into consideration when determining the status quo and the project's objective. These resources are needed to produce correct input for the AI algorithm.

To ensure that the investments will benefit the organization, examining the objective itself is essential. As with any investment decision, the

project should generate value, particularly from the customer experience perspective. If the objective is optimizing the usage of the sources, a brand-new customer-experience approach should be adopted to enhance sustainability. Industrial AI usually seeks solutions tailored to specific machines, therefore deals with a small fragment of featured and reducible data.

Through this analysis, an exploration of whether the product has a disruptive or incremental potential can be carried out. An example of incremental potential would be predictive or adaptive scenarios for boosting shop floor optimization.

Investigating the market for the presence of highly responsive buyers, market positioning and the sustainability impact of the product would also help to reach an estimated value. Strengthening the value proposition through these lenses would help build stakeholders' belief in the project.

How do you generate value in an industrial AI implementation?

It is essential to build the skills to address the gap between OT and IT and create and support the development of soft skills. Competence centres aim to close such skills gaps.

Value is usually generated through a new application or a cost-effective solution. ROI calculations could also generate a numeric value. When calculated transparently and broadly, the gap analysis will help minimize unnecessary costs arising from the project's implementation. The value proposition can also be strengthened by assessing the system's repeatability and the previously prevented issues, in addition to building ground for further process optimization.

The value of an industrial AI implementation does not lie solely in the optimization of the shop floor. When searching for optimization bottlenecks, the project can shed light on work order establishment or facility metrics, thereby providing visibility on enterprise and facility. Predictive maintenance, demand forecasting for dynamic planning and

adaptive heating, ventilation and air conditioning control for sensitive textile manufacturing applications are very difficult to apply without AI technology. Industrial AI is an essential technology. The industrial internet of things (IIoT) enables information about what and why some processes occur but lacks competency to answer questions about what will happen or how the system can react when that process does happen. Combining IIoT with industrial AI can lead to an outstanding level of control over processes and is necessary for a factory with industry 4.0 aspirations. By combining and contextualizing the real-time data acquisition to a specialized framework, a model can be created and transformed into a continuous flow to automate, predict and adapt the process in an optimized sequence.

How can you upskill the workforce for industrial AI?

Industrial AI implementation is no different from general AI implementation with regard to defining a roadmap and having a clear mission and objective while understanding the level of corporate readiness for implementation and deployment. However, the workforce should be upskilled to implement industrial AI. Human capital is a critical factor to sustain success within an organization.²⁵⁵ This fact is hardly surprising since industrial AI-based upskilling can reduce rework rates and increase productivity, and can build efficient teams while also enhancing workforce morale. An example of upskilling would be training the workforce about AI or ML algorithms that could be deployed to the shop floor to build the ground for AI literacy and consequently bring further development to the organization. Another example would be to have the required skills to address the gap between OT and IT and create and support the development of soft

skills for the communication of these departments. Competence centres, through their role as orchestrators between academia and industry, keep the pulse on future technologies in different industries. MEXT in Turkey, the technology centre established by the Turkish Employers Association of Metal Industries (MESS), with \$60 billion in revenue and more than 250 members, is as a good example of these centres.²⁵⁶ Among its wide range of consultancy services related to Fourth Industrial Revolution technologies, such as maturity assessments and road-mapping, it also aims to close the skills gap in the industry with its capability building centre. With its objective to train 250,000 people from different levels of the organization (operators, engineers, mid-level managers and C-level executives) in five years, MEXT has a crucial role to play in the upskilling of the industrial workforce for the digitalization of Turkey.

How do you assess infrastructure readiness for industrial AI?

After the gap analysis, the infrastructure and requirements available to enable industrial AI implementation should be clear. Answering the following questions will help to determine readiness:

- What is the readiness of the current data, storage, compute and network infrastructure?
- What resources are available internally to use (or are using) the cloud to develop and host AI software applications?
- How is the company positioned in the data value chain, and how should that influence short-term vs long-term infrastructure?
- How well is the company prepared to capitalize on forthcoming advances in AI hardware and data storage innovations?

- How will the company use new AI capabilities without causing IT and data inefficiencies?

The computing power required to implement industrial AI, whether it be edge intelligence, cloud computing or a hybrid model, should be decided on based on the resources and computing power currently available and what is desired as determined by the gap analysis. Edge, cloud computing and hybrid systems each have their advantages. An edge computing system could be beneficial in remote locations, when security measures are not outsourced and when low latency is concerned. Cloud computing could be preferred when a priority is to save on equipment and accessibility, and security features are outsourced. Hybrid solutions can be switched between on-premise and third-party service providers, providing control and flexibility. Other decisive factors would be the scalability of the project, the required latency, bandwidth and cost.

What is industrial AI's role in safety?

Industrial AI covers both the physical and the virtual world. Providing a safe and secure environment for each is key for a resilient system.

Industrial AI can also be used to help increase workplace safety measures; an applied use case would be to monitor workplace safety with computer vision. It is possible to measure/predict safety compliance to reduce the risk of job-site accidents, receive real-time unsafe act notifications and operationalize compliance. Industrial AI-supported safety risk mitigating systems can

significantly decrease the risk of accidents on the field and provide new insights to create innovative workplace safety solutions.

In the domain of cybersecurity, as witnessed by the Stuxnet case,²⁵⁷ cyberattacks are not limited to the digital world or online systems but can cause physical disruption to their environment. Security should exceed the boundaries of the organization and cover equipment providers, databases and software. This can serve as a critical decision factor in the choice of equipment and computing power.

How do you govern data for a successful industrial AI implementation?

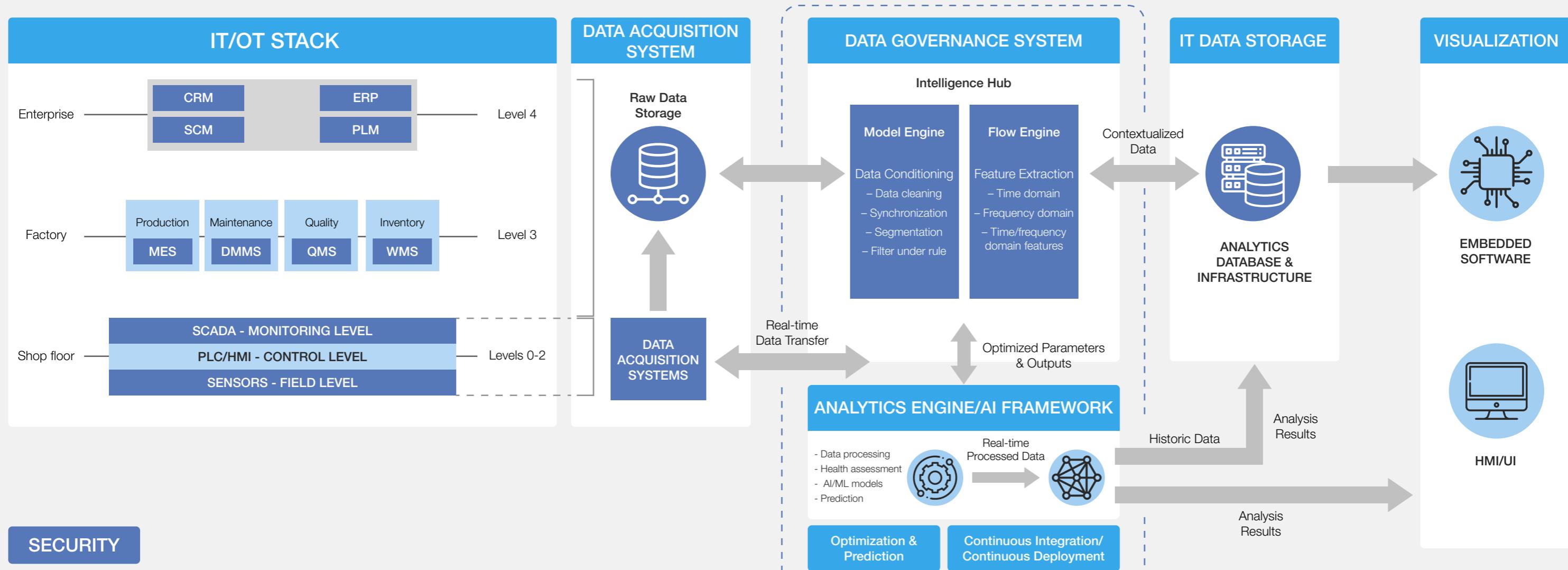
 **Providing a safe and secure environment for both the physical and the virtual worlds is key for a resilient system.**

Without data, AI cannot exist. As the main driver for implementation, data governance is of the utmost importance. In the case of industrial AI, data is derived from the OT layer almost continuously and delivered to the IT layer. Consequently, its path along these layers should be attentively organized. IT and OT convergence connect IT systems to OT systems, which allows these systems to transmit data to each other. This connectivity can then be used to enhance value and is necessary for industrial AI. Some benefits from this connectivity include gaining insights to foster innovation and new services while helping to manage various physical operations, such as real-time remote asset tracking and predictive and preventive maintenance. A lack of communication or understanding of the processes could produce unnecessary or misinterpreted data that will be delivered to the IT layer, creating false information. The first enquiry should be whether the data is frequent and

lightweight, followed by the latency. These insights on data would provide a correct analysis, which would permit transparency through the data path. The quality, reliability and flexibility of the data is further discussed in this toolkit. A similar approach should be taken in the case of industrial AI.

As mentioned, a few of the challenges of AI when governing the data flow include data quality, multisource complexity between domains and machine-to-machine variation. From operation to information, all levels of perfectly connected systems generate very fast and big data. To effectively manage the AI algorithms, the governance of this data plays a critical role. For instance, as an innovation accelerator, an AI integrated digital twin represents a synchronized real-time copy of physical systems to enhance productivity and foster innovation at reduced costs (see the case study).

CASE STUDY | AI integrated digital twin platform framework



To enhance warehouse operations from incoming logistics to the production cycle, a manufacturing company decides to remove the incoming quality control (IQC) process based on supplier outgoing quality performance. Reports requested from suppliers on quality standards and the verification of documents in quality management systems (QMS) show the existence of common prerequisites to complete both at the enterprise and factory levels. As a result of data sharing, companies can plan production efficiently, including rush orders or design revisions.

Based on the agreement with suppliers, the company starts to receive products with fine-grained quality reports defined in measurement protocols, and secures the quality standard in production. After the quality team removes the IQC process, a static sampling rule is required for the suppliers' shipments. However, costs from poor quality, originating from the quality of the material, prompts the team to reconsider whether to increase the sampling frequency or not. The team cannot determine whether the cost of increasing the frequency of the quality control process or the cost of the poor quality will be higher. Had the team been able to predict the batch quality and justify the economic feasibility of the quality control process, it could have increased the overall equipment efficiency (OEE) quality score. This requires Level 3 industrial AI implementation,²⁵⁸ delimited autonomy with the AI system warning of problems or deviations, and a human validating the solution recommended by the system.

To achieve this objective, the data scientist and the team need data from several systems, such as supplier outgoing quality data of the batch from quality management systems (QMS), machine data from manufacturing execution systems (MES) to derive finished goods' quality prediction, batch information from warehouse management systems (WMS) and supply chain management systems (SCM), product data from product life cycle management (PLM) and financials from enterprise resource planning (ERP), in order to understand the break-even point for the cost of poor-quality vs the cost of increasing sampling frequency for the supplier. Contextualizing all of these via a data governance system reduces the workload for the data scientist, since cleaning or inspecting the data is handled by the intelligence hub.

The model engine located in the intelligence hub covers the contextual dimension of the data quality, capturing relevant data to create meaning in the operations. The information is then combined and sent to the flow engine. The flow engine aims to attain the intrinsic value of the data with the feature extraction process. The useful information is gathered and is linked to frequency and time domains.²⁵⁹ In this way, the timeliness, accuracy and completeness of the data are assured, and the resulting data is delivered to the data scientist. To deliver functional results, industrial AI and ML models rely largely on precise, clean and properly labelled (using supervised techniques) training data, making data quality a vital aspect for these solutions' industrial success.

After the data analysis, the information about the batch is delivered to the ML expert for a prediction about the remaining batch. With this system's perpetual utilization, it is possible to make better predictions with less information.

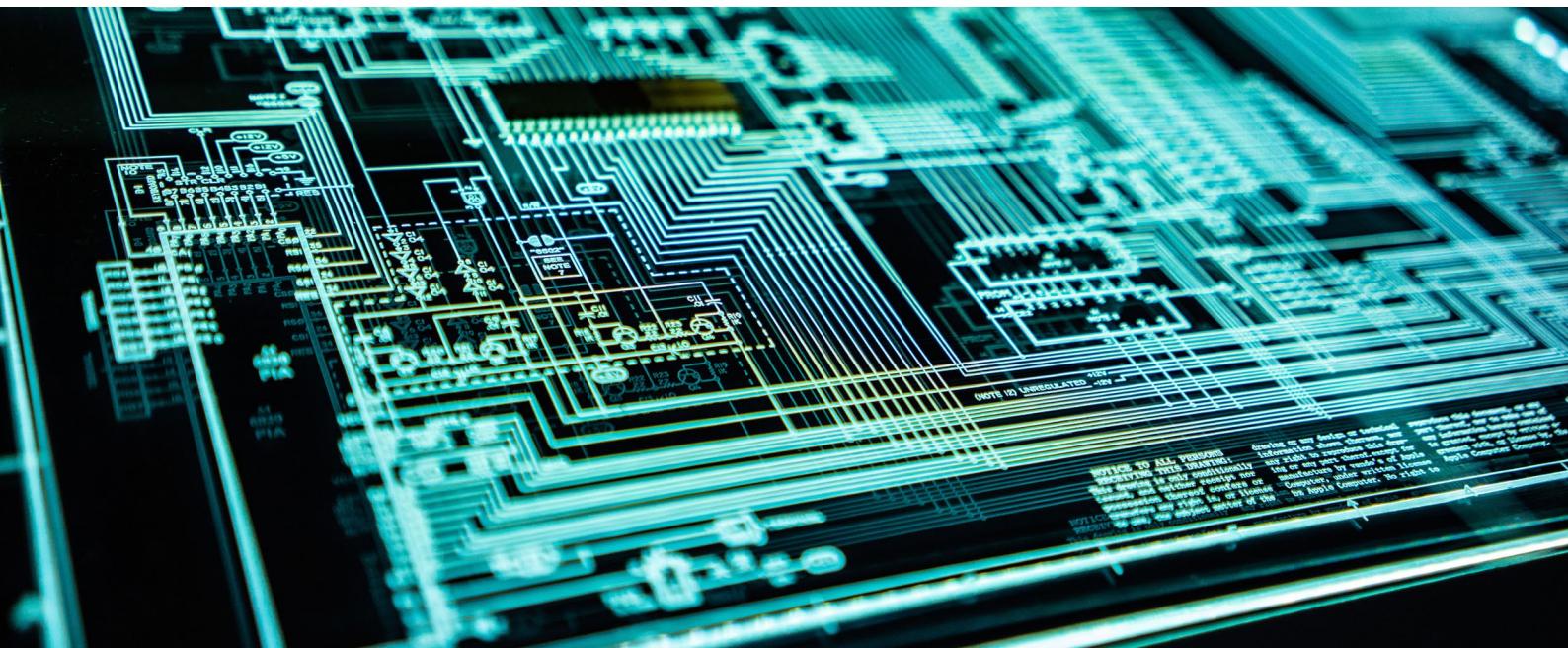
End-to-end data traceability provides “a new level of organization and control over the entire value chain of the life cycle of products. . . . The availability of all relevant information in real time by connecting all instances involved in the value chain [enables] deriving the optimal value-added flow at any time from the data . . . The connection of people, things and systems creates dynamic, self-organizing, real-time optimized value-added connections within and across companies”,²⁶⁰ like providing instantaneous feedback to suppliers about the batch quality prediction, enabling companies to have a new level of organization and control for further optimization opportunities through industrial AI solutions.

What is the role of traceability?

The traceability of the data has huge importance not only during the failure modes but also for the overall safety of the model. Even if the data comes from the desired source and is within the desired range and frequency, data may not be in the correct structure. This situation

The approach of Turkish digital consultancy company, Digitheta Software and Consultancy, which uses the three steps of evaluation, architecture and delivery, is a very good example for this use case. Contextualizing the collected data via labelling at the source, modelling the required data as a standardized payload, and transferring it to related systems based on determined conditions completely eliminate data cleaning and the diagnosis of undesired data for data scientists. Thetasphere Intelligence Hub – Maestro provides a highly flexible and scalable real-time data governance system that shortens the model development process and accelerates the learning curve of the AI models via contextualized and clean data.

can be overcome with traceability and fail-proof properties. As a result, when data fails to have the structural traits, it reports the change. Diagnostics within the AI implementation phase also help to correct and implement the solution, quickly enabling a resilient system.



6.4.3 | Organizational risks

How can you avoid a conservative approach and/or resistance to the project?

Resistance to change is an inevitable natural human instinct, but it can disrupt any project. Thus, potential resistance to change should be effectively mitigated. Being conservative is critical for correct business management. Effectively engaging all participants, including blue- and white-collar workers along with C-level executives to receive and respond to feedback, having a clear business vision from the beginning, and discussing and implementing a roadmap will also help build trust in the project.²⁶¹ As with any change, industrial AI implementation has its own risks. Risk can arise because of miscommunication between OT specialists, data

engineers and domain leaders. This results in false information, which can lead to problematic actions on the part of management. As with other stakeholders, a conservative OT specialist's lack of agreement or miscommunication with the data engineer can disrupt the project. But these risks can be mitigated through OT data change letters or release processes. These processes can help acknowledge and document the change of data across its path and promote transparency. A sensor failing to provide frequent data, for example, would be reported to the data scientist and domain leader by the operator to prevent the production of false information.

6.4.4 | Implementation

Why should you establish an industrial AI implementation?

When implementing industrial AI and determining a research question for an industrial AI project (Figure 25), the focus area should be where the organization can benefit from better decisions and where relevant data is available. The research question should be devised with the following queries in mind to determine the implementation properties.

1. What key bottlenecks is the organization facing?
2. Which areas of the organization will benefit from better solutions?
3. Is substantial and applicable data available in those areas?
4. What level of automation (Figure 26) does the organization need in decision-making for industrial processes?

FIGURE 25

Implementing industrial AI

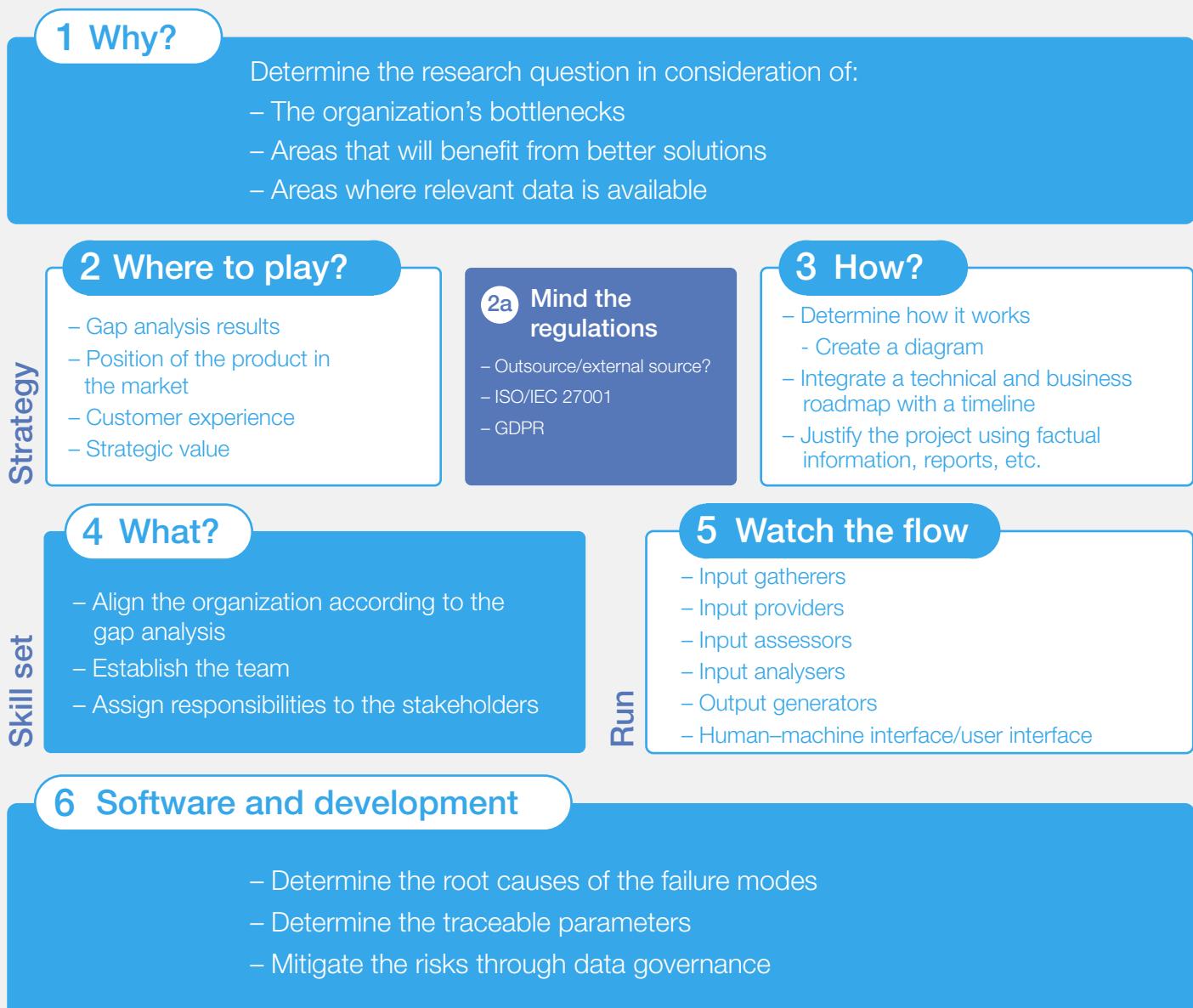
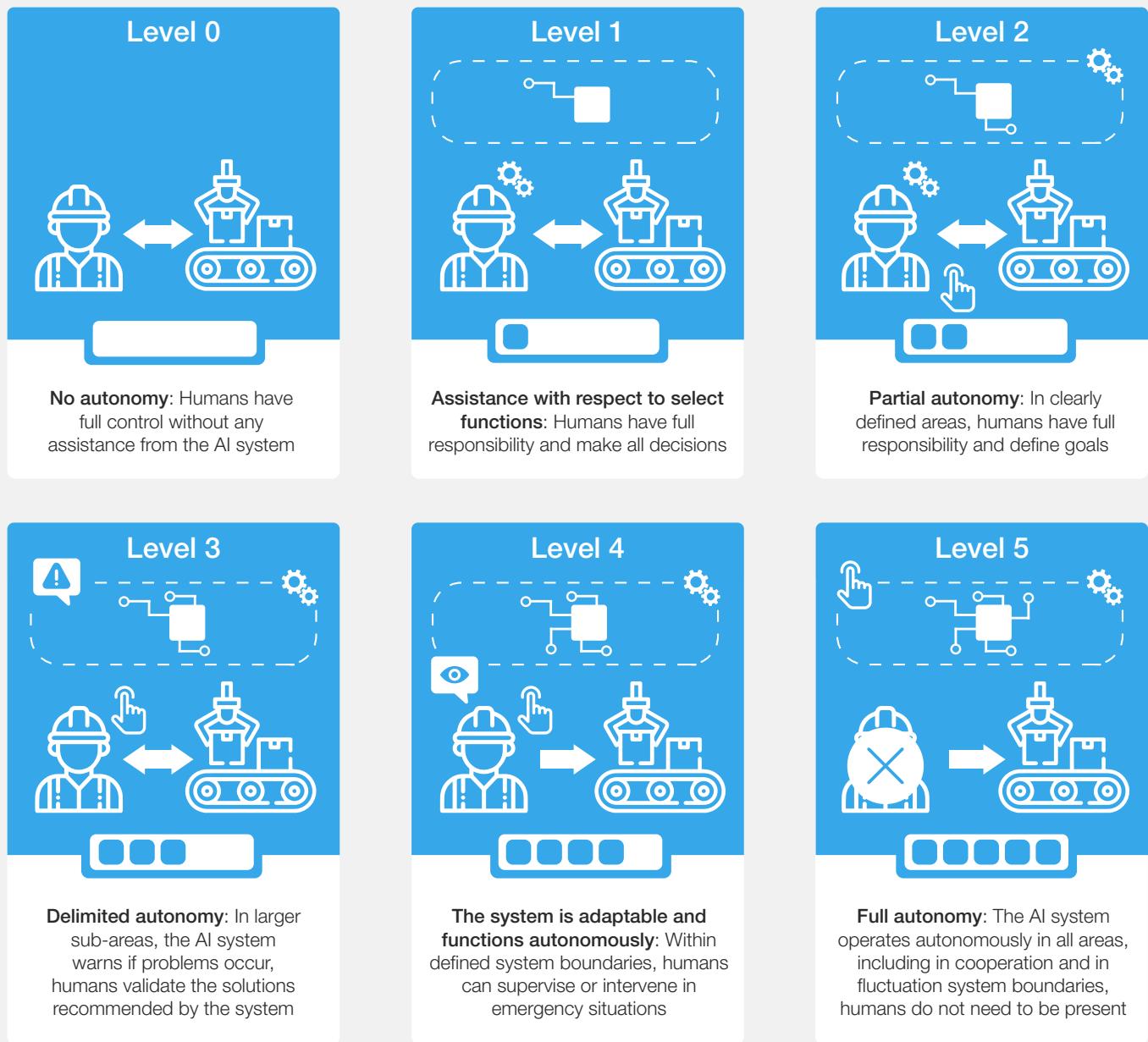


FIGURE 26 | Levels of automation



Source: MEXT analysis adapted from Peres, Ricardo Silva, et al., "Industrial Artificial Intelligence in Industry 4.0 - Systematic Review, Challenges and Outlook", *IEEE Access*, vol. 8, 2020, https://research.unl.pt/ws/portalfiles/portal/29075937/Industrial_Artificial_Intelligence_in_Industry_4.0_Systematic_Review_Challenges_and_Outlook.pdf (accessed 21 November 2021)

Regulations: What are the implications of ISO/IEC 27001, the GDPR and intellectual property rights?

International standard ISO/IEC 27001 ensures that organizations systematically examine the information for security risks, and design a comprehensive control mechanism and management process. If the project involves the collection of personal data, the GDPR could have further implications. Intellectual

property rights should also be considered in any AI implementation. More regulations will be imposed if the implementation is outsourced, since it involves data collection and data transfer. These regulations also pertain to backup data storage or analysis in an external source.

How will the project be conducted?

Justify the project using factual information, reports from well-known institutions, and numerical or survey-related materials. It should be explained in a diagram that indicates how the project will be conducted without disruptions. A technical and business roadmap with a timeline would help manage the business aspects and at the same time would motivate the project's stakeholders.

Following the gap analysis, required skill sets and resources should be obtained before establishing the team. The team should comprise a data scientist, domain expert, domain leader and

analyst, among others. Its coordination and communication, as well as its members' technical expertise on machinery and data, should be secured to ensure a good feedback loop from the sensor to the AI algorithm.

Holding the project's stakeholders responsible for the key data sets is recommended. Assigning those who will be responsible for the data sets is essential to ensure a smooth implementation. The principal actions to assign include gathering the input, providing it to the IT layer, assessing and analysing it, and gathering the output.

How is the implementation process going?

The implementation process starts from raw data or sensor triggered data collection. Input is then gathered in the desired amounts and range while checking the frequency of the data, and provided to the IT layer. Input assessors along with the domain leader assess and evaluate the data to identify the

unnecessary data and to coordinate the data with the project objective. The data is analysed by both data scientists and statisticians, after which the AI experts will turn it into output. The end output enters the human-machine interface or general interface where the experts can use it.

What do you need to consider during the pilot stage?

After the initial system architecture has been established, the failure must be identified before the system is commissioned. The root causes should be established and resolved. This can be

achieved through traceable data. As mentioned in module 2. AI strategy, determining and tracing the parameters are crucial to maintain the system's overall health.

Conclusion

In sum, industrial AI is built on the pillars of infrastructure, data governance and skill sets. Acquiring a clear strategy that is backed up by perceptive research questions is key for starting the project. A robust gap analysis, both from the infrastructure and skill set point of view, will help lay the foundation for the implementation. Data is the project's fuel, so acquiring clean, sufficient and the right type of fuel will help drive industrial AI implementation in an organization.

Emerging topic

Cognitive AI in manufacturing

Industrial AI deals with smaller data sets than general AI. A solution to this is to use cognitive AI, with human-like reasoning to solve problems, combining encoded human knowledge with ML techniques. This new AI technique achieves operational insights with uncertain and smaller data sets.

Appendix 1:

Preparedness Assessment Tool

This tool can be used by individual executives or as a cooperative C-suite exercise to assess the internal understanding of sustainability's potential impact on their organization as well as their organization's ability to address any inherent implications.

When reviewing the relevant sustainability-oriented innovation (SOI) topics below, complete columns a) through d) to assess your organization's preparedness for sustainable development:

- a. To what degree might this topic impact the organization in the future?
- b. Does the organization possess or have access to the knowledge needed to judge AI and sustainability issues?
- c. Who owns responsibility for matters related to this impact category?
- d. How does the organization's business model relate to the impact category and what is the organization's current approach to address potential areas of impact?

Impact category	Sustainability-oriented innovation (SOI) topics	a) Potential degree of impact (high, med, low)	b) Possession of this knowledge (abundant - A, sufficient - S, insufficient - I)	c) Where internal responsibility is held (e.g. VP/Director/Manager XYZ)	d) The organization's current approach to address potential impact (open answer to discuss the organization's approach to potential impact management)
-----------------	---	---	---	--	---

How is AI changing sustainability in our industry and markets?

01	Climate change	Clean power <ul style="list-style-type: none"> - Optimized energy system forecasting - Smart grids for electrical use - Optimized decentralized and peer-to-peer renewable energy systems Smart cities and homes <ul style="list-style-type: none"> - Smart traffic lights and parking systems for urban mobility management - Optimized sustainable building design - Energy-efficient building management systems - Analytics and automation for smart urban planning Sustainable land use <ul style="list-style-type: none"> - Early crop yield prediction - Precision agriculture and nutrition - Monitoring health and well-being in livestock farming Sustainable production and consumption <ul style="list-style-type: none"> - Supply chain monitoring and transparency - Active optimization of industrial machinery and manufacturing - Digital twins for lifespan performance optimization - Smart recycling programmes - Integrated municipal and industrial waste management Smart transport systems <ul style="list-style-type: none"> - On-demand shared transport mobility - AI-enabled electric cars - Autonomous vehicles for efficient transport - Optimized traffic flows 			
02	Biodiversity and conservation	Habitat protection and restoration <ul style="list-style-type: none"> - Precision monitoring of ecosystems - Bird habitat and migration pattern prediction - Simulation of animal and habitat interaction - Habitat loss detection and monitoring - Micro drones for pollination - Optimized breeding of plants 			

02 | *(continued)*

- | | | | |
|---|--|--|--|
| <p>Natural capital realization</p> <ul style="list-style-type: none"> - Registration and trading of biological and biomimetic assets <p>Pollution control</p> <ul style="list-style-type: none"> - Pollutant dispersal prediction and tracking - Analysis of urban run-off quality issues <p>Sustainable trade</p> <ul style="list-style-type: none"> - Food value chain optimization - Supply-chain monitoring and origin tracking - Detection of unauthorized animal capture and trade - Poacher route prediction and high-risk animal tracking - Plant species identification - Machine-automated land-use detection linked to ecosystem payments <p>Invasive species and disease control</p> <ul style="list-style-type: none"> - Machine-automated biodiversity analysis - Smart mosquito traps - Plant disease identification and detection | | | |
|---|--|--|--|

- Sustainable fishing**
- Overfishing prevention and control
 - Insights for fishermen
 - Aquaculture monitoring
 - Monitoring and detection of illegal fishing activities
 - Optimized patrol schedules
- Impact from climate change (including acidification)**
- Real-time monitoring of ocean temperature and pH
 - Phytoplankton distribution detection and prediction
 - Monitoring of ocean currents
 - Monitoring of coral reef ecosystems
- Species protection**
- Monitoring of location and quantities of ocean species
 - Prediction of the spread of invasive species
 - Monitoring and prevention of illegal trafficking of marine wildlife
 - Drones and AI to analyse whale health
- Habitat protection**
- Monitoring of marine habitats for change (e.g. marine dead zones)
 - Habitat conservation assessments
 - Coral reef mapping
 - Autonomous vehicle deep sea assessments
- Pollution prevention**
- Marine litter prediction
 - Robotic fish to fight pollution
 - Real-time monitoring of pollution levels

03 | Healthy oceans

04 | Water security

- | | | | |
|---|--|--|--|
| <p>Water supply</p> <ul style="list-style-type: none"> - Water supply monitoring and management - Water quality simulation and data alerts <p>Drought planning</p> <ul style="list-style-type: none"> - Drought prediction - Simulations of drought planning - Drought impact assessments <p>Adequate sanitation</p> <ul style="list-style-type: none"> - Drones and AI for real-time monitoring of river quality - Monitoring of adequate sanitation of water reserves - Real-time monitoring and management of household water supply <p>Water efficiency</p> <ul style="list-style-type: none"> - Residential water use monitoring and management - Optimization of industrial water use - Predictive maintenance of water plants | | | |
|---|--|--|--|

04	(continued)	<ul style="list-style-type: none"> - Early warning system for water infrastructure - Smart meters in homes <p>Catchment control</p> <ul style="list-style-type: none"> - Harmful algal bloom detection and monitoring - Streamflow forecasting - Automated flood-centred infrastructure 			
05	Clean air	<p>Monitoring and prevention</p> <ul style="list-style-type: none"> - Real-time air pollution monitoring and simulations <p>Filtering and capture</p> <ul style="list-style-type: none"> - Optimized sensor-based air purifying systems - Carbon capture, sequestration and use <p>Clean fuels</p> <ul style="list-style-type: none"> - Advanced battery and fuel-cell design - Advanced battery components - Pollution forecasting for transport management <p>Early warning</p> <ul style="list-style-type: none"> - Air quality alerts - 2-10 day pollution level forecasting 			
06	Weather and disaster resilience	<p>Prediction and forecasting</p> <ul style="list-style-type: none"> - Extreme weather event modelling and prediction - Weather-forecast-informed flight paths - Climate informatics for enhanced climate modelling <p>Resilience planning</p> <ul style="list-style-type: none"> - Impact and risk mitigation analytics - Emergency risk communication - Real-time disaster risk mapping - Real-time disaster response coordination <p>Financial instruments</p> <ul style="list-style-type: none"> - Rapid, multisource risk analysis - Analytics for financial parametric risk instruments - Analytics for claims analysis <p>Resilient infrastructure</p> <ul style="list-style-type: none"> - Automated mitigation of flood risk - Building-specific earthquake damage prediction - Disaster-ready urban infrastructure and buildings <p>Early warning systems</p> <ul style="list-style-type: none"> - Natural catastrophe early warning - Real-time enabled communication of natural disasters - Social media enabled disaster response 			

How are governmental and social pressures influencing sustainable development in our industry and markets?

07	Global commitments to the 17 Sustainable Development Goals (SDGs)	<p>Opportunities for AI to accelerate the achievement of the SDGs</p> <ul style="list-style-type: none"> - Value chain productivity evaluation and enhancement - Research and development - Solving of previously unsolvable business problems - Creation of new business capabilities 			
08	Reporting impact and performance + metrics	<p>Sustainability and impact reporting developments</p> <ul style="list-style-type: none"> - ESG reporting standards and frameworks - Global Reporting Initiative (GRI) - Carbon Disclosure Project (CDP) - IRIS+ impact accounting system - Impact Management Project (IMP) - SDG Compass 			

09	Ethical, legal and other AI responsibilities <ul style="list-style-type: none"> The challenge of AI expansion outpacing the development and deployment of legal and regulatory frameworks and the mechanisms designed to govern it <ul style="list-style-type: none"> - How management judges initiatives' value, risks, compliance with core mission and values, and the responsibilities and legal requirements to be met - How new sustainability-orientated processes might create conflicts with core values and principles - The risks and responsibilities of committing to sustainability-orientated business model strategies 			
----	---	--	--	--

What level of coordination of AI-enabled sustainability initiatives is necessary in our industry and markets?

10	Identification of new transformation and improvement opportunities <p>Sustainability-orientated innovation (SOI): <i>Deliberate changes to products, processes, services, organizations or wider systems to deliver environmental and social as well as economic value</i></p> <ul style="list-style-type: none"> - How management can target AI to support sustainability-driven business models for a better world - How to develop a culture that simultaneously supports SOI and the responsible use of AI and data - What management can learn from other companies' sustainable development strategies and whether AI has a role - Whether management is focused on growth or just cost reduction - How management can use AI to find new opportunities and balances the trade-offs regarding sustainability - How management judges initiatives' value, risks, compliance with core mission and values, and the responsibilities and legal requirements to be met - The prize of aligning business strategies with AI and sustainable development vs the cost of doing nothing - Whether establishing a task force or committee will help the board review the company's AI activities - Benefits of providing an educational programme on AI to board members (either internally or externally) - How to best prioritize the most important SOIs, and consider the impact on the triple bottom line (people, planet and profitability) 			
11	Investment in AI <ul style="list-style-type: none"> - Expenditures in AI to support the company's strategy and sustainability goals - AI investments being made to pursue sustainability-orientated development opportunities - How to strike a balance between short- and long-term investment horizons with the rate of AI's advancement - How to safeguard AI investments against obsolescence due to the rapid rate of tech advancement 			
12	Identification of implementation requirements <ul style="list-style-type: none"> - Whether sustainability-orientated process improvement plans are consistent with core values, and whether these initiatives are well aligned with strategic priorities and business needs - Whether AI-enabled processes for sustainability are in place and how readily the company can meet these challenges - Identification of the implementation and ethics requirements for successfully pursuing these opportunities (e.g. cultural changes; support and incentives for sustainability-orientated innovation; talent needs; cross-functional cooperation; requirements to obtain, manage and protect data) - How to determine when the technology to create these new AI-enabled sustainability-orientated innovations will be mature and scalable enough to use 			

Appendix 2: Guidance Tool

This tool is useful after the Preparedness Assessment Tool. The suggestions below serve as starting points to brainstorm responses to fill any potential gaps that may have surfaced from the initial exercise.

If the C-suite...

Then consider...

01	Needs more external information about AI and environmental sustainability-orientated innovation	<ul style="list-style-type: none">- Reading and subscribing to news sources and reports on AI trends and sustainable development- Attending professional events on AI opportunities, strategies and emerging risks- Meeting with AI researchers, regulators, trusted advisers and sustainability experts- Establishing an emerging technology committee to advise the board or assign a technology advisory role to an existing committee- Reading use cases on websites of AI vendors, consultancies and start-ups- Speaking with peers and executives on the boards of other companies in the same ecosystem- Commissioning (or discovering via conversations with start-ups) third-party perspectives
02	Needs more internal information about AI and environmental sustainability-orientated innovation	<ul style="list-style-type: none">- Requesting reports and updates from the executive team on AI pilots and implementations and on programmes to create a culture for AI success- Meeting with the executives (chief HR officer, chief information officer, chief risk officer, etc.) engaged in developing a mindset for AI success in innovation, risk management and use- Meeting with mid-level human resource managers and teams developing, implementing and using AI
03	Wants to better understand the potential ethical and legal risks of future AI applications and technologies for sustainability-oriented initiatives	<ul style="list-style-type: none">- Setting up an independent ethics board- Hiring AI ethicists as advisers- Asking legal counsel to report on AI's regulatory requirements and legal risks- Engaging with organizations devoted to promoting responsible AI and AI ethics- Attending or setting up a workshop exploring AI ethics issues
04	Needs to focus on strategy and business-model innovation for sustainability-oriented initiatives	<ul style="list-style-type: none">- Setting up design thinking and ideation sessions- Studying AI strategy innovations by competitors and in industry- Focusing on developing a culture and organization that supports innovation and experimentation- Reviewing executive compensation
05	Needs to increase benefits from strategic AI initiatives for sustainability	<ul style="list-style-type: none">- Rapidly improving new AI-enabled systems and processes- Investigating new business models- Examining new joint ventures and partnerships

06	Needs to better understand and follow through on the risk, compliance and responsibilities of AI for sustainability-oriented initiatives	<ul style="list-style-type: none"> - Working with trusted advisers and legal experts to identify risks, relevant legal issues and ethical concerns - Developing an AI governance framework - Establishing an ethics board - Requesting regular reports on risk and responsibility assurance (see the toolkit's sections on ethics, AI governance and risk)
07	Questions the value, terms and risks of a major AI contract for sustainability-oriented initiatives	<ul style="list-style-type: none"> - Bringing in third parties, including attorneys and consultants, to review the contract - Asking the chief financial officer to review and explain questionable parts of the contract - Seeking opinions from executives who were not consulted on the deal - Asking the chief information officer to report on the implications of the deal for the company's IT infrastructure and technology organization - Asking executives who oversee data, legal, ethics and public affairs to analyse the implications of the deal
08	Needs to understand how competitors are leveraging AI technologies for sustainability-oriented initiatives	<ul style="list-style-type: none"> - Hiring benchmarking services - Encouraging management to focus on competitor analyses
09	Needs to improve the stewardship of data for sustainability-oriented initiatives	<ul style="list-style-type: none"> - Appointing a chief data officer or other executive to take responsibility for data management
10	Needs more external information about AI and environmental sustainability-oriented innovation	<ul style="list-style-type: none"> - Establishing an emerging technology committee to advise the board, or assign a technology advisory role to an existing committee - Reading and subscribing to news sources and reports on AI trends and providers - Reading use cases on websites of AI vendors, consultancies and start-ups - Scanning articles in trade press and journals for professionals in target processes - Establishing an educational programme on AI for board members - Attending events on AI and on target processes - Meeting with venture capitalists, AI technology experts and AI researchers at major universities - Speaking with peers and executives on the boards of other companies in the same ecosystem - Commissioning (or discovering via conversations with start-ups) third-party perspectives
11	Wants to better align AI sustainability-oriented innovation activities with priorities	<ul style="list-style-type: none"> - Reviewing and reconsidering KPIs and metrics, including creating new metrics - Developing an operating model that integrates AI process changes into a new way of doing business - Setting up strategy reviews

Contributors

World Economic Forum

Mansour AlAnsari

Lead, Enterprise Analytics, Saudi Aramco;
Artificial Intelligence and Machine Learning Fellow

Theodoros Evgeniou

Professor of Decision Sciences and Technology
Management, INSEAD; Artificial Intelligence and
Machine Learning Academic Partner

Kay Firth Butterfield

Head of Artificial Intelligence and Machine Learning,
World Economic Forum LLC

Arunima Sarkar

Project Lead, Artificial Intelligence and
Machine Learning, Centre for the Fourth
Industrial Revolution India

Centres for the Fourth Industrial Revolution

Tunç Acarkan

Director, Technology Management, Centre for
the Fourth Industrial Revolution Turkey; Turkish
Employers Association of Metal Industries (MESS)

Mariam Al Muhairi

Head, Centre for the Fourth Industrial Revolution
United Arab Emirates

Arwa Al Qassim

AI Lead, Centre for the Fourth Industrial Revolution
United Arab Emirates

Maria Luciana Axente

Lead, Responsible AI and AI for Good, PwC

Titus Capilnean

Director, Corporate Marketing, Appen

Heba Chehade

Foresight Lead, Dubai Future Foundation

Lucas Cooke-Hromek

Senior Consultant, Deloitte Canada

Thomas Davenport

Visiting Professor, Saïd Business School,
University of Oxford

Sukhman Dulay

Consultant, Deloitte Canada

Ilana Alexandra Golbin

Director and Global Responsible AI Leader, PwC

Tim Gordon

Partner, Best Practice AI

Simon Greenman

Partner, Best Practice AI; Member,
World Economic Forum Global AI Council

Aruba Khalid

Associate Project Manager,
Dubai Future Foundation

Rashed Al Mazroui

Intern, Dubai Future Foundation

Manuel Mikoleit

Head, Competence Center Digitization &
Transformation (Compliance), Deutsche Telekom

Emily Ominsky

Responsible AI Asset Curator, Accenture

Michael Pascu

Senior Consultant, Deloitte Canada

Anand Rao

Global Leader, Artificial Intelligence, PwC, USA

Indre Raviv

Senior Vice-President, Marketing, CUJO AI

Murat Solmaz

Senior Partner, Digitheta Software and Consultancy

George Tilesch

President, PHI Institute for Augmented Intelligence

Monika Viktorova

Senior Consultant, Deloitte Canada

Caroline Zimmerman

Research Associate, Hoffmann Global Institute
for Business and Society, INSEAD

Acknowledgements

Cana Alpaslan

Business Analyst, Digitheta Software and Consultancy Inc.

Ece Akın Armutak

Manager, Technology Governance, Turkish Employers Association of Metal Industries (MESS)

Nihar Dalmia

Partner, Government and Public Services National Lead, Deloitte Canada

Elif Gedik

Learning and Development Specialist, Turkish Employers Association of Metal Industries (MESS)

Yasemin Oral

Head, Corporate Communications, Turkish Employers Association of Metal Industries (MESS)

Julián Torres Santeli

Senior Manager, Artificial Intelligence, Government and Public Services Deloitte Canada

Barış Velioğlu

Head, Software Development, Digitheta Software and Consultancy Inc.

Endnotes

1. Rao, Anand, and Gerard Verweij, "Sizing the prize: What's the real value of AI for your business and how can you capitalise?", PwC's Global Artificial Intelligence Study: Exploiting the AI Revolution, 2017, <https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html> (accessed 10 November 2021).
2. Board Agenda and Mazars in association with INSEAD Corporate Governance Centre, "Leadership in AI 2021: Boards, barriers and new beginnings", 2021, <https://www.insead.edu/sites/default/files/assets/dept/centres/icgc/docs/leadership-in-ai-2021.pdf> (accessed 17 November 2021).
3. World Economic Forum, "Empowering AI Leadership: An Oversight Toolkit for Boards of Directors", January 2020, <https://spark.adobe.com/page/RsXNkZANwMLEf> (accessed 1 November 2021).
4. World Economic Forum, "Empowering AI Leadership: An Oversight Toolkit for Boards of Directors", op. cit.
5. Organisation for Economic Co-operation and Development (OECD), "Recommendation of the Council on Artificial Intelligence", OECD Legal Instruments, OECD/LEGAL/0449, 22 May 2019, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449> (accessed 1 November 2021).
6. Candelier, François, Rodolphe Charme di Carlo, Midas De Bondt and Theodoros Evgeniou, "AI Regulation Is Coming: How to prepare for the inevitable", *Harvard Business Review*, September-October 2021, <https://hbr.org/2021/09/ai-regulation-is-coming> (accessed 11 November 2021).
7. Anyoha, Rockwell, "The History of Artificial Intelligence", Harvard University, 28 August 2017; Strickland, Eliza, "The turbulent past and uncertain future of artificial intelligence", Institute of Electrical and Electronics Engineers, *IEEE Spectrum*, 30 September 2021; Dilmegani, Cem, "Measuring AI advancement through games (2021 update)", AIMultiple, 1 January 2021 update; Silver, David, and Demis Hassabis, "AlphaGo: Mastering the ancient game of Go with Machine Learning", Google AI blog, 27 January 2016; Gershgorn, Dave, "The data that transformed AI research—and possibly the world", Quartz, 26 July 2017; The Medical Futurist, "FDA-approved A.I.-based algorithms", 2021; BBC, "AI: 15 key moments in the story of artificial intelligence", 2021; Thomson Reuters, "Our AI timeline", 2021; Gartner, "Hype Cycle for Artificial Intelligence, 2019", 2019.
8. Simonite, Tom, "AI Ruined Chess. Now, It's Making the Game Beautiful Again", Wired, 9 September 2020, <https://www.wired.com/story/ai-ruined-chess-now-making-game-beautiful> (accessed 11 November 2021).
9. IBM Research, "Watson and the Jeopardy! Challenge", YouTube, 6 November 2013, <https://www.youtube.com/watch?app=desktop&v=P18EdAKuC1U> (accessed 11 November 2021).
10. DeepMind, "AlphaGo" case study, 2021, <https://deepmind.com/research/case-studies/alphago-the-story-so-far> (accessed 11 November 2021).
11. OpenAI Five "OpenAI Five, 2016-2019", <https://openai.com/five> (accessed 11 November 2021).
12. Heaven, Will Douglas, "DeepMind's AI can now play all 57 Atari games—but it's still not versatile enough", *MIT Technology Review*, 1 April 2020, <https://www.technologyreview.com/2020/04/01/974997/deepminds-ai-57-atari-games-but-its-still-not-versatile-enough> (accessed 11 November 2021).
13. Cargill, "Cargill brings facial recognition capability to farmers through strategic equity investment in Cainthus", 31 January 2018, <https://www.cargill.com/2018/cargill-brings-facial-recognition-capability-to-farmers> (accessed 11 November 2021).
14. Blue River Technology [website], <https://www.bluerivertechnology.com> (accessed 11 November 2021).
15. Biobank [website], <https://www.ukbiobank.ac.uk> (accessed 11 November 2021).
16. Alsever, Jennifer, "Medicine by machine: Is A.I. the cure for the world's ailing drug industry?", *Fortune*, 20 January 2020, <https://fortune.com/longform/ai-artificial-intelligence-medicine-healthcare-pharmaceutical-industry> (accessed 11 November 2021).
17. Armstrong, Maggie Mae, "Cheat sheet: What is Digital Twin?", IBM, 4 December 2020, <https://www.ibm.com/blogs/internet-of-things/iot-cheat-sheet-digital-twin> (accessed 11 November 2021).
18. Sam, Saarinen, Evan Cater and Michael Littman, "Personalized Education at Scale", 24 September 2018, <https://arxiv.org/abs/1809.10025> (accessed 11 November 2021).
19. Massive Open Online Courses (MOOCs) [website], <https://www.mooc.org> (accessed 11 November 2021).
20. O'Neil, Cathy, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown Publishers, 2016.
21. Gender Shades, "How well do IBM, Microsoft, and Face++ AI services guess the gender of a face?", MIT Media Lab, 9 February 2018, <http://gendershades.org> (accessed 11 November 2021).
22. Boneh, Dan, Andrew Grotto, Patrick McDaniel and Nicolas Papernot, "Preparing for the Age of Deepfakes and Disinformation", Stanford University, Policy Briefs, November 2021, <https://fsi.stanford.edu/publication/preparing-age-deepfakes-and-disinformation> (accessed 11 November 2021).
23. Metz, Cade, "Meet GPT-3. It Has Learned to Code (and Blog and Argue)", *The New York Times*, 24 November 2020, <https://www.nytimes.com/2020/11/24/science/artificial-intelligence-ai-gpt3.html> (accessed 11 November 2021).

24. GPT-3, "A robot wrote this entire article. Are you scared yet, human?", *The Guardian*, 8 September 2020, <https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3> (accessed 11 November 2021).
25. Yeo, Catherine, "How Biased is GPT-3?", *Medium*, 3 June 2020, <https://medium.com/fair-bytes/how-biased-is-gpt-3-5b2b91f1177> (accessed 11 November 2021).
26. Business Wire, "IDC's Global DataSphere Forecast Shows Continued Steady Growth in the Creation and Consumption of Data", 8 May 2020, <https://www.businesswire.com/news/home/20200508005025/en/IDCs-Global-DataSphere-Forecast-Shows-Continued-Steady-Growth-in-the-Creation-and-Consumption-of-Data> (accessed 5 January 2022).
27. Immersive IT, "Prevent Data Silos for Business Success", Immersive IT blog, 2019, <https://immersive-it.co.uk/blog/prevent-data-silos-for-business-success> (accessed 11 November 2021).
28. Business Wire, "IDC's Global DataSphere Forecast Shows Continued Steady Growth in the Creation and Consumption of Data", op. cit.
29. Simpson Rochwerger, Alyssa and Wilson Pang, "From Google Translate to Amazon Recruiting: When Good AI Goes Bad", Enterprise AI, 4 June 2021, <https://www.enterpriseai.news/2021/06/04/from-google-translate-to-amazon-recruiting-when-good-ai-goes-bad/> (accessed 1 November 2021).
30. Manyika, James, and Jacques Bughin, "The promise and challenge of the age of artificial intelligence", McKinsey Global Institute, 15 October 2018, <https://www.mckinsey.com/featured-insights/artificial-intelligence/the-promise-and-challenge-of-the-age-of-artificial-intelligence> (accessed 1 November 2021).
31. Business Wire, "NewVantage Partners Releases 2021 Big Data and AI Executive Survey", 4 January 2021, <https://www.businesswire.com/news/home/20210104005022/en/NewVantage-Partners-Releases-2021-Big-Data-and-AI-Executive-Survey> (accessed 11 November 2021).
32. Evans, Richard, and Jim Gao, "DeepMind AI Reduces Google Data Centre Cooling Bill by 40%", DeepMind, 20 July 2016, <https://deepmind.com/blog/article/deepmind-ai-reduces-google-data-centre-cooling-bill-40> (accessed 11 November 2021).
33. Davenport, Tom, "On AI and Jobs, We Are All Augmentarians Now", *Forbes*, 8 June 2018, <https://www.forbes.com/sites/tomdavenport/2018/06/08/on-ai-and-jobs-we-are-all-augmentarians-now/> (accessed 1 November 2021).
34. Babic, Boris, Sara Gerke, Theodoros Evgeniou and I. Glenn Cohen, "Beware explanations from AI in health care", *Science*, vol. 373, no. 6552, 16 July 2021, pp. 284-286, <https://www.science.org/doi/abs/10.1126/science.abg1834> (accessed 11 November 2021).
35. IBM, "Everyday Ethics for Artificial Intelligence", 2019, <https://www.ibm.com/watson/assets/duo/pdf/everydayethics.pdf> (accessed 1 November 2021).
36. Ibid., p. 15.
37. Ibid., p. 21.
38. Ibid., p. 27.
39. Ibid., p. 32.
40. Ibid., p. 40.
41. White, Andrew, "Our Top Data and Analytics Predicts for 2019", Gartner, 3 January 2019, https://blogs.gartner.com/andrew_white/2019/01/03/our-top-data-and-analytics-predicts-for-2019 (accessed 1 November 2021).
42. Fountaine, Tim, Brian McCarthy and Tamim Saleh, "Building the AI-Powered Organization: Technology isn't the biggest challenge. Culture is", *Harvard Business Review*, July-August 2019, <https://hbr.org/2019/07/building-the-ai-powered-organization> (accessed 12 November 2021).
43. FutureLearn, "How to Foster a Data-driven Culture", Data Analysis and Fundamental Statistics, <https://www.futurelearn.com/info/courses/data-analytics-for-business-basic-analysis-and-statistics/0/steps/177286> (accessed 12 November 2021).
44. Dubai Future Foundation, *New Data Economy Report*, 2021, <https://www.dubaifuture.ae/insights/new-data-economy-report> (accessed 12 November 2021).
45. Thomke, Stefan, "Building a Culture of Experimentation", *Harvard Business Review*, March-April 2020, <https://hbr.org/2020/03/building-a-culture-of-experimentation> (accessed 1 November 2021).
46. FutureLearn, "How to Foster a Data-driven Culture", op. cit.
47. Thomke, "Building a Culture of Experimentation", op. cit.
48. Ibid.
49. Ibid.
50. Kohavi, Ron, and Stefan Thomke, "The Surprising Power of Online Experiments: Getting the Most Out of A/B and Other Controlled Tests", Harvard Business School, September-October 2017, <https://www.hbs.edu/faculty/Pages/item.aspx?num=53201> (accessed 12 November 2021).
51. Thomke, "Building a Culture of Experimentation", op. cit.
52. United Nations Educational, Scientific and Cultural Organization (UNESCO), "UNESCO launches worldwide online public consultation on the ethics of artificial intelligence", 15 July 2020, <https://en.unesco.org/news/unesco-launches-worldwide-online-public-consultation-ethics-artificial-intelligence> (accessed 1 November 2021).

53. Ibid.
54. United Nations Educational, Scientific and Cultural Organization (UNESCO), “AI Ethics: Another step closer to the adoption of UNESCO’s Recommendation”, 2 July 2021, <https://en.unesco.org/news/ai-ethics-another-step-closer-adoption-unescos-recommendation-0> (accessed 1 November 2021).
55. World Economic Forum, *The Future of Jobs Report 2020*, October 2020, https://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf (accessed 1 November 2021).
56. Lindzon, Jared, “How AI Is Changing The Way Companies Are Organized”, *Fast Company*, 28 February 2017, <https://www.fastcompany.com/3068492/how-ai-is-changing-the-way-companies-are-organized> (accessed 1 November 2021).
57. Ibid.
58. Kambil, Ajit, “4 types of executive sponsorship to catalyze change”, Deloitte, 11 October 2017, <https://www2.deloitte.com/us/en/insights/focus/executive-transitions/four-types-of-executive-sponsorship-to-catalyze-change.html> (accessed 19 November 2021).
59. Bughin, Jacques, Michael Chui and Brian McCarthy, “How to make AI work for your business”, *Harvard Business Review*, 15 August 2017, <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/how-to-make-ai-work-for-your-business> (accessed 19 November 2021).
60. Loucks, Jeff, Tom Davenport and David Schatsky, “State of AI in the Enterprise, 2nd Edition”, Deloitte, 2018, https://www2.deloitte.com/content/dam/insights/us/articles/4780_State-of-AI-in-the-enterprise/DI_State-of-AI-in-the-enterprise-2nd-ed.pdf (accessed 19 November 2021).
61. Zwetsloot, Remco, “Strengthening the U.S. AI Workforce”, Center for Security and Emerging Technology (CSET), September 2019, https://cset.georgetown.edu/wp-content/uploads/CSET_US_AI_Workforce.pdf (accessed 1 November 2021).
62. World Economic Forum, *The Future of Jobs Report 2018*, Centre for the New Economy and Society, 2018, https://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf (accessed 1 November 2021).
63. LinkedIn, “APEC – Closing the Digital Skills Gap Forum”, 2019, https://economicgraph.linkedin.com/content/dam/me/economicgraph/en-us/download/apec_ai_talent_insights_version.pdf (accessed 19 November 2021).
64. VTRAC Consulting Corporation, “How to Attract (and Retain) AI Talent”, 2019, <https://www.vtrac.com/2019/02/20/how-to-attract-and-retain-ai-talent>, (accessed 1 November 2021).
65. Ibid.
66. Ibid.
67. Ibid.
68. Ibid.
69. Bush, Matt, “Why Is Diversity & Inclusion in the Workplace Important?”, Great Place to Work, 13 April 2021, <https://www.greatplacetowork.com/resources/blog/why-is-diversity-inclusion-in-the-workplace-important> (accessed 15 November 2021).
70. Ibid.
71. Ibid.
72. Nonnecke, Brandie, “Artificial intelligence can make our societies more equal. Here’s how”, World Economic Forum, Agenda, 21 September 2017, <https://www.weforum.org/agenda/2017/09/applying-ai-to-enable-an-equitable-digital-economy-and-society> (accessed 1 November 2021).
73. ESI ThoughtLab, “Driving ROI through AI”, September 2020, https://econsultsolutions.com/wp-content/uploads/2020/09/ESITL_Driving-ROI-through-AI_FINAL_September-2020.pdf (accessed 15 November 2021).
74. Sirkin, Harold L., Perry Keenan and Alan Jackson, “The Hard Side of Change Management”, *Harvard Business Review*, October 2005, <https://hbr.org/2005/10/the-hard-side-of-change-management> (accessed 15 November 2021).
75. Johnk, Jan, Malte Weissert and Katrin Wyrtki, “Ready or Not, AI Comes – An Interview Study of Organizational AI Readiness Factors”, *Business & Information Systems Engineering*, vol. 63, no. 1, 2021, pp. 5-20.
76. Ibid.
77. World Economic Forum, *The Future of Jobs Report 2020*, October 2020, https://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf (accessed 15 November 2021).
78. Romsey, Joseph, “5 Tips to Help Workers Upskill and Adapt to Artificial Intelligence”, Society for Human Resource Management (SHRM), 30 November 2020, <https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/how-hr-can-help-workers-upskill-and-adapt-to-artificial-intelligence-5-tips.aspx> (accessed 2 November 2021).
79. Ibid.
80. Ibid.
81. This section is based on research conducted by Joseph Romsey. See “5 Tips to Help Workers Upskill and Adapt to Artificial Intelligence”, op. cit.
82. Boudens, Jeff, Rodgers Palmer and Brooke Weddle, “Mobilize your organization with a powerful change story”, McKinsey & Company, 28 October 2019, <https://www.mckinsey.com/business-functions/people-and-organizational-performance/our-insights/the-organization-blog/mobilize-your-organization-with-a-powerful-change-story> (accessed 15 November 2021).
83. Johnk, Weissert and Wyrtki, “Ready or Not, AI Comes – An Interview Study of Organizational AI Readiness Factors”, op. cit.

84. Organisation for Economic Co-operation and Development (OECD), “E-commerce in the times of COVID-19”, 7 October 2020, https://read.oecd-ilibrary.org/view/?ref=137_137212-t0fjgnerdb&title=E-commerce-in-the-time-of-COVID-19&ga=2.133446798.1677695305.1628073411-1364741124.1628073411 (accessed 15 November 2021).
85. Newman, Daniel, “5 Ways AI Is Transforming the Customer Experience”, *Forbes*, 16 April 2019, <https://www.forbes.com/sites/danielnewman/2019/04/16/5-ways-ai-is-transforming-the-customerexperience/?sh=269f394a465a> (accessed 15 November 2021).
86. Stover, Steve, “AI Success Relies on Strong Organizational Change Management”, *CMSWire*, 11 April 2019, <https://www.cmswire.com/digital-workplace/ai-success-relies-on-strong-organizational-change-management> (accessed 3 November 2021).
87. Sustainability Management School (SUMAS) Editorial Team, “Millennials and their Impact on Sustainability”, 22 January 2019, <https://sumas.ch/millennials-and-sustainability> (accessed 15 November 2021).
88. Authority of the House of Lords, Select Committee on Artificial Intelligence, *AI in the UK: ready, willing and able? Report of Session 2017-19*, 16 April 2018, <https://publications.parliament.uk/pa/ld201719/ldselect/l dai/100/100.pdf> (accessed 15 November 2021).
89. Floridi, Luciano, et al., “AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations”, *Minds & Machines*, vol. 28, 2018, pp. 689-707, <https://doi.org/10.1007/s11023-018-9482-5> (accessed 15 November 2021).
90. PwC, “PWC’s Responsible AI”, <https://www.pwc.com/gx/en/issues/data-and-analytics/artificial-intelligence/what-is-responsible-ai.html> (accessed 15 November 2021).
91. Bossmann, Julia, “Top 9 ethical issues in artificial intelligence”, World Economic Forum Agenda, 21 October 2016, <https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence> (accessed 15 November 2021).
92. European Commission, High-Level Expert Group on Artificial Intelligence (AI HLEG), “Ethics Guidelines for Trustworthy AI”, 8 April 2019, <https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html> (accessed 15 November 2021).
93. GitHub, “Awful AI”, <https://github.com/daviddao/awful-ai> (accessed 16 November 2021).
94. Future of Privacy Forum, “Unfairness by algorithm: Distilling the harms of automated decision-making”, 11 December 2017, <https://fpf.org/2017/12/11/unfairness-by-algorithm-distilling-the-harms-of-automated-decision-making> (accessed 15 November 2021).
95. Bonime-Blanc, Andrea, “It’s time we added a letter to ESG. Here’s why”, World Economic Forum Agenda, 14 October 2020, <https://www.weforum.org/agenda/2020/10/its-time-we-added-a-letter-to-esg-heres-why> (accessed 16 November 2021).
96. Scher, Stephen, and Kasia Kozlowska, “The Rise of Bioethics: A Historical Overview”, *Rethinking Health Care Ethics*, 2018, pp. 31-44, https://link.springer.com/chapter/10.1007/978-981-13-0830-7_3 (accessed 16 November 2021).
97. Ibid.
98. Future of Privacy Forum, “Unfairness by algorithm: Distilling the harms of automated decision making”, op. cit.
99. Scientific Foresight (STOA), “Is artificial intelligence a human rights issue?”, 19 March 2019, <https://epthinktank.eu/2019/03/19/is-artificial-intelligence-a-human-rights-issue> (accessed 16 November 2021).
100. European Commission, “Ethics Guidelines for Trustworthy AI”, op. cit.
101. Organisation for Economic Co-operation and Development (OECD), “OECD Principles on AI”, <https://www.oecd.org/going-digital/ai/principles/> (accessed 16 November 2021).
102. PwC, “PWC’s Responsible AI”, 2019, <https://www.pwc.com/gx/en/issues/artificial-intelligence/responsible-ai-placemat.pdf> (accessed 16 November 2021).
103. Tasioulas, John, “First Steps Towards an Ethics of Robots and Artificial Intelligence”, *Journal of Practical Ethics*, vol. 7, no. 1, 2019, <http://www.jpe.ox.ac.uk/papers/first-steps-towards-an-ethics-of-robots-and-artificial-intelligence> (accessed 4 November 2021).
104. Bonduelle, Yann, et al., “Comprehensive AI governance needed”, PwC and Eurasia Group, 2020, https://www.eurasia-group.net/files/upload/Comprehensive_AI_Governance_Needed_Now_Eurasia_Group_PWC_Japan.pdf (accessed 16 November 2021).
105. European Parliament, “Parliament leads the way on first set of EU rules for Artificial Intelligence”, Press Release, 20 October 2020, <https://www.europarl.europa.eu/news/en/press-room/20201016IPR89544/parliament-leads-the-way-on-first-set-of-eu-rules-for-artificial-intelligence> (accessed 16 November 2021).
106. Virginia, Dignum, *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*, Springer, 2019, <https://www.springer.com/gp/book/9783030303709> (accessed 4 November 2021).
107. Virginia Dignum in public presentations.
108. Lauer, Dave, “You cannot have AI ethics without ethics”, *AI and Ethics*, vol. 1, 2021, pp. 21-25, <https://link.springer.com/article/10.1007/s43681-020-00013-4> (accessed 4 November 2021).
109. Babic, Boris, I. Glenn Cohen, Theodoros Evgeniou and Sara Gerke, “When Machine Learning Goes Off the Rails: A guide to managing the risks”, *Harvard Business Review*, January-February 2021, <https://hbr.org/2021/01/when-machine-learning-goes-off-the-rails> (accessed 4 November 2021).

110. Babic, Boris, Sara Gerke, Theodoros Evgeniou and I. Glenn Cohen, "Algorithms on regulatory lockdown in medicine" *Science*, vol. 366, no. 6470, 2019, pp. 1202-1204, <https://www.science.org/doi/abs/10.1126/science.aay9547> (accessed 4 November 2021).
111. Rao, Anand, "Five Views of AI Risk: Understanding the darker side of AI", *Towards Data Science*, 29 November 2020, <https://towardsdatascience.com/five-views-of-ai-risk-eddb2fce3c2> (accessed 4 November 2021).
112. Cusumano, Michael A., Annabelle Gawer and David B. Yoffie, interviewed by Allyson MacDonald, "How Digital Platforms Have Become Double-Edged Swords", *MIT Sloan Management Review*, 8 May 2019, <https://sloanreview.mit.edu/article/how-digital-platforms-have-become-double-edged-swords> (accessed 4 November 2021).
113. Rao, Anand, and Gerard Verweij, "Sizing the prize: What's the real value of AI for your business and how can you capitalise?", PwC, <https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf> (accessed 16 November 2021).
114. Rao, "Five Views of AI Risk: Understanding the darker side of AI", op. cit.
115. Knight, Will, and Karen Hao, "Never mind killer robots – here are six real AI dangers to watch out for in 2019", 7 January 2019, <https://www.technologyreview.com/2019/01/07/137929/never-mind-killer-robotshere-are-six-real-ai-dangers-to-watch-out-for-in-2019/> (accessed 16 November 2021).
116. McEvoy, Fiona J., "Is Emotion AI a Dangerous Deceit?", 8 March 2019, <https://youthedata.com/2019/03/08/is-emotion-ai-a-dangerous-deceit> (accessed 17 November 2021).
117. Jimenez, Alison, "BankThink AI use carries bias risk for financial regulators", *American Banker*, 16 August 2019, <https://www.americanbanker.com/opinion/ai-use-carries-bias-risk-for-financial-regulators> (accessed 17 November 2021).
118. Montague, Enid, "Dangers of artificial intelligence in medicine", *The Hill*, 16 January 2020, <https://thehill.com/opinion/healthcare/478651-dangers-of-artificial-intelligence-in-medicine> (accessed 17 November 2021).
119. Hao, Karen, "A new way to train AI systems could keep them safer from hackers", *MIT Technology Review*, 10 July 2020, <https://www.technologyreview.com/2020/07/10/1005048/ai-deep-learning-safe-from-hackers-adversarial-attacks> (accessed 17 November 2021).
120. Dickson, Ben, "TrojanNet – a simple yet effective attack on machine learning models", *The Daily Swig*, 30 June 2021 update, <https://portswigger.net/daily-swig/trojannet-a-simple-yet-effective-attack-on-machine-learning-models> (accessed 17 November 2021).
121. Strout, Nathan, "The 3 major security threats to AI", C4ISRNet, 10 September 2019, <https://www.c4isrnet.com/artificial-intelligence/2019/09/10/the-3-major-security-threats-to-ai> (accessed 17 November 2021).
122. Eadicicco, Lisa, "There's a terrifying trend on the internet that could be used to ruin your reputation, and no one knows how to stop it", *Business Insider India*, 10 July 2019, <https://www.businessinsider.in/tech/theres-a-terrifying-trend-on-the-internet-that-could-be-used-to-ruin-your-reputation-and-no-one-knows-how-to-stop-it/articleshow/70162319.cms> (accessed 16 November 2021).
123. Chokshi, Niraj, "How Surveillance Cameras Could Be Weaponized With A.I.", *The New York Times*, 13 June 2019, <https://www.nytimes.com/2019/06/13/us/aclu-surveillance-artificial-intelligence.html> (accessed 17 November 2021).
124. Larson, Jeff, Surya Mattu, Lauren Kirchner and Julia Angwin, "How We Analyzed the COMPAS Recidivism Algorithm", ProPublica, 23 May 2016, <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm> (accessed 17 November 2021).
125. Stupp, Catherine, "Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case", *The Wall Street Journal*, 30 August 2019, <https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402> (accessed 17 November 2021).
126. Ashman, Ben, "When AI goes wrong: What happens when machines go rogue?", TbTech, 24 July 2019, <https://tbttech.co/innovativetech/artificial-intelligence/what-happens-when-ai-goes-wrong> (accessed 17 November 2021).
127. Amidon, Alexandra, "How Concept Drift Ruins Your Model Performance", *Towards Data Science*, 29 June 2020, <https://towardsdatascience.com/concept-drift-can-ruin-your-model-performance-and-how-to-address-it-dff08f97e29b> (accessed 17 November 2021).
128. MacCarthy, Mark, "Fairness in algorithmic decision-making", Brookings, 6 December 2019, <https://www.brookings.edu/research/fairness-in-algorithmic-decision-making> (accessed 17 November 2021).
129. Datarama, "The Confluence of Artificial Intelligence and Political Risk", Medium, 19 April 2018, <https://medium.com/@dataramatech/the-confluence-of-artificial-intelligence-and-political-risk-d3ce33ec125f> (accessed 17 November 2021).
130. Muro, Mark, "Countering the geographical impacts of automation: Computers, AI, and place disparities", Brookings, 14 February 2019, <https://www.brookings.edu/research/countering-the-geographical-impacts-of-automation-computers-ai-and-place-disparities> (accessed 17 November 2021).
131. Hambling, David, "The US Army is creating robots that can follow orders", *MIT Technology Review*, 6 November 2019, <https://www.technologyreview.com/2019/11/06/132036/the-us-army-is-creating-robots-that-can-follow-ordersand-ask-if-they-dont-understand> (accessed 17 November 2021).
132. Ackerman, Evan, "Lethal Microdrones, Dystopian Futures, and the Autonomous Weapons Debate: The future of weaponized robots requires a reasoned discussion, not scary videos", Institute of Electrical and Electronics Engineers, *IEEE Spectrum*, 15 November 2017, <https://spectrum.ieee.org/lethal-microdrones-dystopian-futures-and-the-autonomous-weapons-debate> (accessed 17 November 2021).

133. Rao, "Five Views of AI Risk: Understanding the darker side of AI", op. cit.
134. Simonite, Tom, "New York City Proposes Regulating Algorithms Used in Hiring", Wired, 8 January 2021, <https://www.wired.com/story/new-york-city-proposes-regulating-algorithms-hiring> (accessed 17 November 2021).
135. Rao, Anand, "Ten Principles of Responsible AI for Corporates", Towards Data Science, 18 December 2020, <https://towardsdatascience.com/ten-principles-of-responsible-ai-for-corporates-dd85ca509da9> (accessed 4 November 2021).
136. European Confederation of Institutes of Internal Auditing (ECIIA) and Federation of European Risk Management Associations (FERMA), "Guidance on the 8th EU Company Law Directive, article 41", 21 September 2010, <https://www.iia.nl/SiteFiles/ECIIA%20FERMA.pdf> (accessed 4 November 2021).
137. Institute of Internal Auditors (IIA), "Position Paper: The Three Lines of Defense in Effective Risk Management and Control", January 2013, <https://na.theiia.org/standards-guidance/Public%20Documents/PP%20The%20Three%20Lines%20of%20Defense%20in%20Effective%20Risk%20Management%20and%20Control.pdf> (accessed 4 November 2021).
138. Institute of Internal Auditors (IIA), "The IIA's Three Lines Model: An update of the Three Lines of Defense", July 2020, <https://global.theiia.org/about/about-internal-auditing/Public%20Documents/Three-Lines-Model-Updated.pdf> (accessed 4 November 2021).
139. Rao, Anand, "Data, Automation, Analytics, AI – The Unbeatable Quartet: Competitive Advantage from the Four Waves of Digital (4wD) (Part 1)", Towards AI, Medium, 18 September 2020, <https://pub.towardsai.net/data-automation-analytics-ai-the-unbeatable-quartet-509536b2f097> (accessed 17 November 2021).
140. Data Science Association, "Data Science Code of Professional Conduct", <https://www.datascienceassn.org/code-of-conduct.html> (accessed 17 November 2021).
141. Oxford-Munich Code of Conduct, "Code of Conduct for Professional Data Scientists", 2018, <http://www.code-of-ethics.org> (accessed 17 November 2021).
142. Institute of Electrical and Electronics Engineers (IEEE), "The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems", <https://standards.ieee.org/industry-connections/ec/autonomous-systems.html> (accessed 8 November 2021).
143. Rao, Anand, "Data Scientists are from Mars and Software Developers are from Venus (Part 1)", Towards Data Science, 29 August 2020, <https://towardsdatascience.com/data-scientists-are-from-mars-and-software-developers-are-from-venus-part-1-8dde19fb2eef> (accessed 17 November 2021).
144. Institute of Electrical and Electronics Engineers (IEEE), "The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems: Well-being", https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead_well_being_v2.pdf (accessed 17 November 2021).
145. World Economic Forum, "AI Procurement in a Box: Project overview", Toolkit, June 2020, https://www3.weforum.org/docs/WEF_AI_Procurement_in_a_Box_Project_Overview_2020.pdf (accessed 17 November 2021).
146. Rao, Anand, "Model Lifecycle: From ideas to value: Value scoping, discovery, delivery, and stewardship", Towards Data Science, 27 September 2020, <https://towardsdatascience.com/model-lifecycle-from-ideas-to-value-14e654b7d4a4> (accessed 17 November 2021).
147. Brubaker, Richard, "The Business Case for Ethics", Collective Responsibility, 15 October 2015, <https://www.coreponsibility.com/business-case-ethics> (accessed 4 November 2021).
148. Golbin, Ilana, and Maria Luciana Axente, "9 ethical AI principles for organizations to follow", World Economic Forum Agenda, 23 June 2021, <https://www.weforum.org/agenda/2021/06/ethical-principles-for-ai> (accessed 9 November 2021).
149. Institute of Electrical and Electronics Engineers (IEEE), "IEEE Ethics in Action: Ethically Aligned Design in Practice", <https://ethicsinaction.ieee.org> (accessed 9 November 2021).
150. European Commission, "Ethics Guidelines for Trustworthy AI", op. cit.
151. Golbin, Ilana, Anand Rao, Ali Hadjarian and Daniel Krittman, "Responsible AI: A Primer for the Legal Community", 2020 Institute of Electrical and Electronics Engineers (IEEE) International Conference on Big Data (Big Data), 2020, pp. 2121-2126, <https://ieeexplore.ieee.org/document/9377738> (accessed 9 November 2021).
152. Ficklin, Patrice Alexander, Tom Pahl and Paul Watkins, "Innovation spotlight: Providing adverse action notices when using AI/ML models", Consumer Finance Protection Bureau (US), 7 July 2020, <https://www.consumerfinance.gov/about-us/blog/innovation-spotlight-providing-adverse-action-notices-when-using-ai-ml-models> (accessed 17 November 2021).
153. Flynn, Shannon, "13 Cities Where Police Are Banned From Using Facial Recognition Tech", Innovation & Tech Today, 18 November 2020, <https://innotechtoday.com/13-cities-where-police-are-banned-from-using-facial-recognition-tech> (accessed 18 November 2021).
154. Thieullent, Anne-Laure, et al., "AI and the Ethical Conundrum", Capgemini Research Institute, 2020, <https://www.capgemini.com/wp-content/uploads/2020/10/AI-and-the-Ethical-Conundrum-Report.pdf> (accessed 4 November 2021).
155. ESI ThoughtLab, "Driving ROI through AI", op. cit.
156. Thieullent, et al., "AI and the Ethical Conundrum", op. cit.
157. Ibid.
158. Economist Intelligence Unit (EIU), *Staying ahead of the curve: The business case for responsible AI*, October 2020, <https://pages.eiu.com/rs/753-RIQ-438/images/EIUSTayingAheadOfTheCurve.pdf> (accessed 18 November 2021).

159. Capgemini, "Retail in AI: An ethical dilemma. Consumers expect companies to respect their data when implementing AI", 31 July 2019, <https://www.capgemini.com/2019/07/retail-in-ai-an-ethical-dilemma> (accessed 4 November 2021).
160. Thieullent, et al., "AI and the Ethical Conundrum", op. cit.
161. Ibid.
162. Economist Intelligence Unit (EIU), *Staying ahead of the curve: The business case for responsible AI*, op. cit.
163. Ibid.
164. Capgemini, "Retail in AI: An ethical dilemma", op. cit.
165. Economist Intelligence Unit (EIU), *Staying ahead of the curve: The business case for responsible AI*, op. cit.
166. Volini, Erica, et al., "Ethics and the future of work", Deloitte, 15 May 2020, <https://www2.deloitte.com/us/en/insights/focus/human-capital-trends/2020/ethical-implications-of-ai.html> (accessed 4 November 2021).
167. Ibid.
168. Ibid.
169. Ibid.
170. Haider, Ailia, and Wade Islan, "Staying ahead of the curve – The business case for responsible AI", Economist Intelligence Unit, 8 October 2020, <https://www.eiu.com/n/staying-ahead-of-the-curve-the-business-case-for-responsible-ai> (accessed 4 November 2021).
171. Kelly, Richard, and Massimo Pellegrino, "Ethical AI: Tensions and trade-offs", PwC, 11 June 2019, <https://www.pwc.com.au/digitalpulse/ethical-artificial-intelligence-tensions-trade-offs.html> (accessed 18 November 2021).
172. Van Loon, Ronald, "How Corporate C-Levels Can Be the Guardians of Ethical AI," Simplilearn, 8 June 2020 update, <https://www.simplilearn.com/how-corporate-c-levels-can-be-the-guardians-of-ethical-ai-article> (accessed 18 November 2021).
173. Kelly and Pellegrino, "Ethical AI: Tensions and trade-offs", op. cit.
174. Blackman, Reid, "A Practical Guide to Building Ethical AI", *Harvard Business Review*, 15 October 2020, <https://hbr.org/2020/10/a-practical-guide-to-building-ethical-ai> (accessed 18 November 2021).
175. Tonkinwise, Cameron, "Ethics by Design, or the Ethos of Things", *Design Philosophy Papers*, vol. 2, no. 2, 2004, pp. 129-144, https://www.thestudioattheedgeoftheworld.com/uploads/4/7/4/0/47403357/05tonkinwiseethos_of_things.pdf (accessed 18 November 2021).
176. Sandler, Ronald, and Johyn Basl, "Building Data and AI Ethics Committees", Accenture and Northeastern University Ethics Institute, 2019, https://www.accenture.com/_acnmedia/PDF-107/Accenture-AI-Data-Ethics-Committee-Report.pdf (accessed 18 November 2021).
177. See Karoff, Paul, "Embedding ethics in computer science curriculum", *The Harvard Gazette*, 25 January 2019, <https://news.harvard.edu/gazette/story/2019/01/harvard-works-to-embed-ethics-in-computer-science-curriculum> (accessed 8 November 2021).
178. Harvard University, "Embedded EthisCSTM @ Harvard", <https://embeddedethics.seas.harvard.edu> (accessed 19 November 2021).
179. See Karoff, "Embedding ethics in computer science curriculum", op. cit.
180. Gambelin, Olivia, "Brave: what it means to be an AI Ethicist", *AI and Ethics*, vol. 1, 2021, pp. 87-91, <https://link.springer.com/article/10.1007%2Fs43681-020-00020-5> (accessed 8 November 2021).
181. Ibid.
182. Dunham, Ann B., and Shawn Stout-Jough, "Are Ethics Hotlines Effective?", Society for Human Resource Management (SHRM), 26 February 2020, <https://www.shrm.org/hr-today/news/hr-magazine/spring2020/pages/are-ethics-hotlines-effective.aspx> (20 November 2021).
183. Timsit, Annabelle, "Alexa is very confused by little kids", Quartz Media, 12 August 2018, <https://qz.com/1352272/when-alexa-and-other-smart-speakers-misunderstand-little-kids> (accessed 20 November 2021).
184. Hassan, Aisha, "Workplace automation will hit women harder than men", Quartz Media, 1 April 2019, <https://qz.com/1583578/workplace-automation-will-hit-women-harder-than-men> (accessed 20 November 2021).
185. Saran, Cliff, "Collaboration tools must support the wellbeing of home workers", *Computer Weekly*, 5 October 2020, <https://www.computerweekly.com/news/252490012/Collaboration-tools-must-support-the-wellbeing-of-home-workers> (accessed 20 November 2021).
186. Royal Society of Arts (RSA), "Public services need citizen engagement on AI – report", Press Release, 24 October 2019, <https://www.thersa.org/press/releases/2019/rsa-public-services-need-citizen-engagement-on-ai-report> (accessed 20 November 2021).
187. Mökander, Jakob, Jessica Morley, Mariarosaria Taddeo and Luciano Floridi, "Ethics-Based Auditing of Automated Decision-Making Systems: Nature, Scope, and Limitations", *Science and Engineering Ethics*, vol. 27, no. 44, 2021, <https://link.springer.com/article/10.1007%2Fs11948-021-00319-4> (accessed 8 November 2021).
188. Carrier, Ryan, and Shea Brown, "Taxonomy: AI Audit, Assurance & Assessment", ForHumanity, February 2021, <https://static1.squarespace.com/static/5ff3865d3fe4fe33db92ffdc/t/60329e0a4cfbaa172691f7e6/161392999802/Taxonomy+of+AI+Audit+%282%29.pdf> (accessed 20 November 2021).

189. Excerpt from Institute of Electrical and Electronics Engineers (IEEE), “IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems: Methods to Guide Ethical Research and Design”, in *Ethically Aligned Design*, First Edition, reproduced in IEEE, “A Call to Action for Businesses Using AI”, <https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead/ead-for-business.pdf> (accessed 8 November 2021).
190. Microsoft, “Microsoft AI Principles”, <https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1:primary6> (accessed 9 November 2021).
191. Crampton, Natasha, “The building blocks of Microsoft’s responsible AI program”, Microsoft, 19 January 2021, <https://blogs.microsoft.com/on-the-issues/2021/01/19/microsoft-responsible-ai-program> (accessed 9 November 2021).
192. Microsoft, “Putting principles to practice”, <https://www.microsoft.com/en-us/ai/our-approach?activetab=pivot1:primary7> (accessed 20 November 2021).
193. Microsoft, “Transparency Note: Azure Cognitive Services Face API”, 6 May 2019, <https://azure.microsoft.com/en-us/resources/transparency-note-azure-cognitive-services-face-api> (accessed 20 November 2021).
194. Roach, John, “Using AI, people who are blind are able to find familiar faces in a room”, Microsoft, <https://news.microsoft.com/innovation-stories/project-tokyo> (accessed 20 November 2021).
195. Brill, Julie, “GDPR’s first anniversary: A year of progress in privacy protection”, Microsoft, 20 May 2019, <https://blogs.microsoft.com/on-the-issues/2019/05/20/gdpr-first-anniversary-a-year-of-progress-in-privacy-protection> (accessed 20 November 2021).
196. Bird, Sarah, et al., “A toolkit for assessing and improving fairness in AI”, Microsoft, May 2020, <https://www.microsoft.com/en-us/research/publication/fairlearn-a-toolkit-for-assessing-and-improving-fairness-in-ai/> (accessed 20 November 2021).
197. Sameki, Mehrnoosh, Sarah Bird and Kathleen Walker, “InterpretML: A toolkit for understanding machine learning models”, Microsoft, 18 May 2020, <https://www.microsoft.com/en-us/research/uploads/prod/2020/05/InterpretML-Whitepaper.pdf> (accessed 20 November 2021).
198. This case study was provided by Best Practice AI in collaboration with Deutsche Telekom.
199. The materials and facts were provided by Deutsche Telekom.
200. Fulde, Verena, “Guidelines for Artificial Intelligence”, Deutsche Telekom, 2018, <https://www.telekom.com/en/company/digital-responsibility/details/artificial-intelligence-ai-guideline-524366> (accessed 20 November 2021).
201. Davenport, Tom, Theodoros Evgeniou and Thomas C. Redman, “Your Data Supply Chains Are Probably a Mess. Here’s How to Fix Them”, *Harvard Business Review*, 24 June 2021, <https://hbr.org/2021/06/data-management-is-a-supply-chain-problem> (accessed 21 November 2021).
202. U.S. Department of Health & Human Services, Centers for Disease Control and Prevention (CDC), “The Six Dimensions of [Early Hearing Detection and Intervention] EHDI Data Quality Assessment”, 2020, <https://www.cdc.gov/ncbddd/hearingloss/documents/dataqualityworksheet.pdf> (accessed 21 November 2021).
203. Bhatt, Yogesh, and Subhashish Prasad, “OmniaAI: Eliminate data chaos and ambiguity”, Deloitte, <https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/deloitte-analytics/ca-omnia-datagovernance-digital-en.pdf?elqTrackId=5120ab63c9974ab1b5c937b420b48781&elqaid=1199&elqat=2> (accessed 21 November 2021).
204. Deloitte, “Learning to drive your Ferrari: The Data Agenda empowers your most valuable asset”, <https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/insights-and-issues/ca-en-insights-issues-the-data-agenda.pdf> (accessed 21 November 2021).
205. Microsoft, Azure, “What is cloud computing? A beginner’s guide: Types of cloud computing”, <https://azure.microsoft.com/en-us/overview/what-is-cloud-computing/#cloud-deployment-types> (accessed 21 November 2021).
206. Chui, Michael, Vishnu Kamalnath and Brian McCarthy, “An executive’s guide to AI”, McKinsey & Company, 2020 update, <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai> (accessed 21 November 2021).
207. Deloitte, “Building trust in AI: How to overcome risk and operationalize AI governance”, <https://www2.deloitte.com/ca/en/pages/financial-services/articles/bridging-operationalizing-trust-in-ai.html> (accessed 22 November 2021).
208. Abdulrahman, Sawsan, et al., “A Survey on Federated Learning: The Journey From Centralized to Distributed On-Site Learning and Beyond”, *IEEE Internet of Things Journal*, vol. 8, no. 7, 2021, pp. 5476-5497, <https://ieeexplore.ieee.org/document/9220780> (accessed 8 November 2021).
209. Racharak, Teeradaj, “On Approximation of Concept Similarity Measure in Description Logic *ELH* with Pre-Trained Word Embedding”, *IEEE Access*, vol. 9, 2021, pp. 61429-61443, <https://ieeexplore.ieee.org/document/9406016/authors#authors> (accessed 8 November 2021).
210. Preece, Alun, et al., “Stakeholders in Explainable AI”, 29 September 2018, <https://arxiv.org/abs/1810.00184> (accessed 21 November 2021).
211. DataRobot, “Machine Learning Model Deployment”, 2021, <https://www.datarobot.com/wiki/machine-learning-model-deployment> (accessed 8 November 2021).
212. Ibid.
213. Ibid.
214. Vonau, Manuel, “Xayn wants to be your privacy-first Google replacement, complete with its own Discover feed”, Android Police, 8 December 2020, <https://www.androidpolice.com/2020/12/08/xayn-wants-to-be-your-privacy-first-google-replacement-complete-with-its-own-discover-feed> (accessed 8 November 2021).

215. Ibid.
216. Ibid.
217. Ibid.
218. Raynor, Michael, and Derek Pankratz, "A new business paradigm to address climate change", Deloitte, 16 October 2020, <https://www2.deloitte.com/us/en/insights/topics/strategy/corporate-climate-change-sustainability.html> (accessed 8 November 2021).
219. Leonard N. Stern School of Business, New York University, "Latest Research From NYU Stern Center for Sustainable Business and IRI Shows That Sustainability Is Surviving COVID-19", 16 July 2020, <https://www.stern.nyu.edu/experience-stern/faculty-research/latest-research-nyu-stern-center-sustainable-business-and-iri-shows-sustainability-surviving-covid> (accessed 21 November 2021).
220. Kerrigan, Sarah, and Duleesha Kulasooriya, "The sustainability transformation: Look ahead, look inside, and look around", Deloitte, 31 July 2020, <https://www2.deloitte.com/global/en/insights/topics/strategy/sustainable-transformation-in-business.html> (accessed 8 November 2021).
221. World Economic Forum, "Measuring Stakeholder Capitalism: Towards Common Metrics and Consistent Reporting of Sustainable Value Creation", White Paper, 2020, https://www3.weforum.org/docs/WEF_IBC_Measuring_Stakeholder_Capitalism_Report_2020.pdf (accessed 21 November 2021).
222. Kerrigan and Kulasooriya, "The sustainability transformation: Look ahead, look inside, and look around", op. cit.
223. Ibid.
224. Raynor and Pankratz, "A new business paradigm to address climate change", op. cit.
225. Trinks, Arjan, Machiel Mulder and Bert Scholtens, "An efficiency perspective on carbon emissions and financial performance", *Ecological Economics*, 2020.
226. Holbrook, Emily, "New Initiative Helps Farmers Generate and Sell Carbon Credits", Environment + Energy Leader, 3 November 2020, <https://www.environmentalleader.com/2020/11/new-initiative-helps-farmers-generate-and-sell-carbon-credits> (accessed 21 November 2021).
227. Faull, Jennifer, "Ad Association launches Climate Action group to coordinate industry efforts", The Drum, 28 January 2020, <https://www.thedrum.com/news/2020/01/28/ad-association-launches-climate-action-group-coordinate-industry-efforts> (accessed 22 November 2021).
228. Raynor and Pankratz, "A new business paradigm to address climate change", op. cit.
229. AIAnsari, Mansour, "4 steps to using AI in an environmentally responsible way", World Economic Forum Agenda, 1 April 2021, <https://www.weforum.org/agenda/2021/04/4-steps-to-using-ai-in-an-environmentally-responsible-way-artificial-intelligence-bcg-code-carbon> (accessed 8 November 2021).
230. PwC, "Using AI to better manage the environment could reduce greenhouse gas emissions, boost global GDP by up to US \$5 trillion and create up to 38m jobs by 2030", 16 April 2019, <https://www.pwc.com/gx/en/news-room/press-releases/2019/ai-realise-gains-environment.html> (accessed 8 November 2021).
231. AIAnsari, "4 steps to using AI in an environmentally responsible way", op. cit.
232. Makala, Baloko, and Tonci Bakovic, "Artificial Intelligence in the Power Sector", International Finance Corporation, April 2020, https://www.ifc.org/wps/wcm/connect/bd3a196d-a88f-45af-bbc6-e0b00790fba8/EMCompass_Note_81-05-web.pdf?MOD=AJPERES&CVID=n72pj5g#:~:text=Artificial%20intelligence%2C%20or%20AI%2C%20has.and%20control%20of%20power%20systems (accessed 8 November 2021).
233. Association for Computing Machinery (ACM), Communications of the ACM Technology Review, "Training A Single AI Model Can Emit As Much Carbon As Five Cars In Their Lifetimes", 7 June 2019, <https://cacm.acm.org/careers/237345-training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes/fulltext> (accessed 22 November 2021).
234. AIAnsari, "4 steps to using AI in an environmentally responsible way", op. cit.
235. Malliaraki, Eirini, "AI and climate change: The promise, the perils and pillars for action", Climate-KIC (Knowledge and Innovation Community), 5 November 2020, <https://www.climate-kic.org/opinion/ai-and-climate-change-the-promise-the-perils-and-pillars-for-action> (accessed 22 November 2021).
236. Mila, "Top AI Experts Create CodeCarbon, a Tool to Track and Reduce Computing's CO2 Emissions", 1 December 2020, <https://mila.quebec/en/top-ai-experts-create-codecarbon-a-tool-to-track-and-reduce-computings-co2-emissions> (accessed 8 November 2021).
237. Gerdeman, Dina, "Who Is the Chief Sustainability Officer?", Harvard Business School, Working Knowledge, 8 October 2014, <https://hbswk.hbs.edu/item/who-is-the-chief-sustainability-officer> (accessed 22 November 2021).
238. Microsoft, "Shell and Microsoft form alliance to help address carbon emissions", News Center, 22 September 2020, <https://news.microsoft.com/2020/09/22/shell-and-microsoft-form-alliance-to-help-address-carbon-emissions> (accessed 22 November 2021).
239. Mcfall-Johnsen, Morgan, "The fashion industry emits more carbon than international flights and maritime shipping combined. Here are the biggest ways it impacts the planet", *Business Insider India*, 18 October 2019, <https://www.businessinsider.in/science/news/the-fashion-industry-emits-more-carbon-than-international-flights-and-maritime-shipping-combined-here-are-the-biggest-ways-it-impacts-the-planet-/articleshow/71640863.cms> (accessed 22 November 2021).

240. H&M, *H&M Group Sustainability Report 2018*, 2019, https://about.hm.com/content/dam/hmgroup/groupsite/documents/masterlanguage/CSR/reports/2018_Sustainability_report/HM_Group_SustainabilityReport_2018_%20FullReport.pdf (accessed 22 November 2021).
241. European Federation for Transport and Environment, "What is the impact of shipping on climate change?", <https://www.transportenvironment.org/challenges/ships/greenhouse-gases> (accessed 22 November 2021).
242. Viu-Roig, Marta, and Eduard J. Alvarez-Palau, "The Impact of E-Commerce-Related Last-Mile Logistics on Cities: A Systematic Literature Review", *Sustainability*, vol. 12, 6492, 12 August 2020, <https://www.mdpi.com/2071-1050/12/16/6492/pdf> (accessed 23 November 2021).
243. Nichols, Megan Ray, "10 Ways AI Improves Distribution and the Supply Chain", *Towards Data Science*, 22 June 2020, <https://towardsdatascience.com/10-ways-ai-improves-distribution-and-the-supply-chain-bbd2dc600965> (accessed 8 November 2021).
244. Deloitte, "Jespers Torvekokken will use new technology to ensure that the right amount of food is produced at the right time", 2021, <https://www2.deloitte.com/dk/d/a/pages/climate-and-sustainability/jespers-torvekokken.html> (accessed 8 November 2021).
245. Ibid.
246. SMS, "Climate change targets would be put in jeopardy without smart meters", 20 May 2019, <https://www.sms-plc.com/insights/blogs-news/climate-change-targets-would-be-put-in-jeopardy-without-smart-meters> (accessed 22 November 2021).
247. Mariaw, "Florida Power & Light: Making a Smart (Meter) Bet on the Power Grid of the Future", Harvard Business School, Technology and Operations Management, 18 November 2016, <https://digital.hbs.edu/platform-rctom/submission/florida-power-light-making-a-smart-meter-bet-on-the-power-grid-of-the-future> (accessed 22 November 2021).
248. Deloitte, "The energy company Brunata can now identify over-consumption in individual radiators that return water with an excessive temperature", 2021, <https://www2.deloitte.com/dk/d/a/pages/climate-and-sustainability/brunata.html> (accessed 8 November 2021).
249. Knutson, Ted, "AI Is Mining ESG Data Companies Aren't Disclosing, Says BlackRock Exec", *Forbes*, 18 June 2020, <https://www.forbes.com/sites/tedknutson/2020/06/18/ai-mining-esg-data-companies-not-disclosing-says-blackrock-exec/?sh=491862737d97> (accessed 23 November 2021).
250. McKinsey & Company, "The state of AI in 2020", Survey, 17 November 2020, <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/global-survey-the-state-of-ai-in-2020> (accessed 8 November 2021).
251. Lee, Jay, *Industrial AI: Applications with Sustainable Performance*, Springer, 2020.
252. Peres, Ricardo Silva, et al., "Industrial Artificial Intelligence in Industry 4.0 – Systematic Review, Challenges and Outlook", *IEEE Access*, vol. 8, 2020, https://www.researchgate.net/publication/346915033_Industrial_Artificial_Intelligence_in_Industry_40_-Systematic_Review_Challenges_and_Outlook (accessed 8 November 2021).
253. Ibid.
254. Lee, Jay, Jaskaran Singh and Moslem Azamfar, "Industrial Artificial Intelligence", 22 October 2019 update, <https://arxiv.org/abs/1908.02150v3> (accessed 23 November 2021).
255. Jaiswal, Akanksha, C. Joe Arun and Arup Varma, "Rebooting employees: Upskilling for artificial intelligence in multinational corporations", *The International Journal of Human Resource Management*, 16 March 2021, <https://www.tandfonline.com/doi/full/10.1080/09585192.2021.1891114> (accessed 23 November 2021).
256. MEXT [website], Turkey, <https://www.mext.org.tr> (accessed 22 November 2021).
257. Kushner, David, "The real story of Stuxnet", Institute of Electrical and Electronics Engineers, *IEEE Spectrum*, 25 February 2013, <https://spectrum.ieee.org/telecom/security/the-real-story-of-stuxnet> (accessed 17 November 2021).
258. Peres, et al., "Industrial Artificial Intelligence in Industry 4.0 – Systematic Review, Challenges and Outlook", op. cit.
259. Ibid.
260. Thangaraj, Jeevitha, and Ramya Lakshmi Narayanan, "Industry 1.0 to 4.0: The Evolution of Smart Factories", October 2018, <https://www.coursehero.com/file/77572037/INDUSTRY1TO40THEEVOLUTIONOFSMARTFACTORIESpdf> (accessed 8 November 2021).
261. Singh, Alok, et al., "Future of Work", McKinsey & Company Turkey and McKinsey Global Institute, <https://www.mckinsey.com/tr/~media/mckinsey/locations/europe%20and%20middle%20east/turkey/our%20insights/future%20of%20work%20turkey/future-of-work-mckinsey-turkey-full-report.pdf> (accessed 25 November 2021).



COMMITTED TO
IMPROVING THE STATE
OF THE WORLD

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

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

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

Tel.: +41 (0) 22 869 1212
Fax: +41 (0) 22 786 2744
contact@weforum.org
www.weforum.org