McKinsey & Company

McKinsey Explainers

What is AI?

Artificial intelligence is a machine's ability to perform the cognitive functions we usually associate with human minds.

April 2023

Humans and machines: a match made in productivity heaven. Our species wouldn't have gotten very far without our mechanized workhorses. From the wheel that revolutionized agriculture to the screw that held together increasingly complex construction projects to the robot-enabled assembly lines of today, machines have made life as we know it possible. And yet, despite their seemingly endless utility, humans have long feared machines—more specifically, the possibility that machines might someday acquire human intelligence and strike out on their own.

But we tend to view the possibility of sentient machines with fascination as well as fear. This curiosity has helped turn science fiction into actual science. Twentieth-century theoreticians, like computer scientist and mathematician Alan Turing, envisioned a future where machines could perform functions faster than humans. The work of Turing and others soon made this a reality. Personal calculators became widely available in the 1970s, and by 2016, the US census showed that 89 percent of American households had a computer. Machines—*smart* machines at that—are now just an ordinary part of our lives and culture.

Those smart machines are getting faster and more complex. Some computers have now crossed the exascale threshold, meaning that they can perform as many calculations in a single second as an individual could in 31,688,765,000 years. But it's not just about computation. Computers and other devices are now acquiring skills and perception that have previously been *our* sole purview.

AI is a machine's ability to perform the cognitive functions we associate with human minds, such as perceiving, reasoning, learning, interacting with an environment, problem solving, and even exercising creativity. You've probably interacted with AI even if you didn't realize it—voice assistants like Siri and Alexa are founded on AI technology, as are some customer service chatbots that pop up to help you navigate websites.

Applied AI—simply, artificial intelligence applied to real-world problems—has serious implications for the business world. By using artificial intelligence, companies have the potential to make business more efficient and profitable. But ultimately, the value of artificial intelligence isn't in the systems themselves but in how companies use those systems to assist humans—and their ability to explain to shareholders and the public what those systems do—in a way that builds and earns trust.

For more about AI, and how to apply it in business, read on.

What is machine learning?

Machine learning is a form of artificial intelligence based on algorithms that are trained on data. These algorithms can detect patterns and learn how to make predictions and recommendations by processing data and experiences, rather than by receiving explicit programming instruction. The algorithms also adapt in response to new data and experiences to improve their efficacy over time. The volume and complexity of data that is now being generated, too vast for humans to reasonably reckon with, has increased the potential of machine learning, as well as the need for it. In the years since its widespread deployment, which began in the 1970s, machine learning has had impact in a number of industries, including achievements in medical-imaging analysis and high-resolution weather forecasting.

What is deep learning?

Deep learning is a type of machine learning that can process a wider range of data resources (images, for instance, in addition to text), requires even less human intervention, and can often produce more accurate results than traditional machine learning. Deep learning uses neural networks—based on the ways neurons interact in the human brain—to ingest data and process it through multiple

The volume and complexity of data that is now being generated, too vast for humans to reasonably reckon with, has increased the potential of machine learning, as well as the need for it.

iterations that learn increasingly complex features of the data. The neural network can then make determinations about the data, learn whether a determination is correct, and use what it has learned to make determinations about new data. For example, once it "learns" what an object looks like, it can recognize the object in a new image.

Here are three types of artificial neural networks used in machine learning:

— **Feed-forward neural networks**
In this simple neural network, first proposed in 1958, information moves in only one direction: forward from the model's input layer to its output layer, without ever traveling backward to be reanalyzed by the model. That means you can feed, or input, data into the model, then "train" the model to predict something about different data sets. As just one example, feed-forward neural networks are used in banking, among other industries, to detect fraudulent financial transactions.

Here's how it works: first, you train a model to predict whether a transaction is fraudulent based on a data set you've used to manually label transactions as fraudulent or not. Then you can use the model to predict whether new, incoming transactions are fraudulent so you can flag them for closer study or block them outright.

— **Convolutional neural networks (CNNs)**
CNNs are a type of feed-forward neural network modeled on the makeup of the animal visual cortex, the part of the brain that processes images. As such, CNNs are well suited to perceptual tasks, like being able to identify bird or plant species based on photographs. Business use cases include diagnosing diseases from medical scans, or detecting a company logo in social media to manage a brand's reputation or to identify potential joint marketing opportunities.

Here's how CNNs work:

- First, the CNN receives an image—for example, of the letter "A"—that it processes as a collection of pixels.
- In the hidden layers, the CNN identifies unique features—for example, the individual lines that make up "A."
- The CNN can now classify a different image as the letter "A" if it finds that the image has the unique features previously identified as making up the letter.

— **Recurrent neural networks (RNNs)**
RNNs are artificial neural networks whose connections include loops, meaning the model both moves data forward and loops it backward

to run again through previous layers. RNNs are helpful for predicting a sentiment or an ending of a sequence, like a large sample of text, speech, or images. They can do this because each individual input is fed into the model by itself as well as in combination with the preceding input.

Continuing with the banking example, RNNs can help detect fraudulent financial transactions just as feed-forward neural networks can, but in a more complex way. Whereas feed-forward neural networks can help predict whether one individual transaction is likely to be fraudulent, recurrent neural networks can "learn" from the financial behavior of an individual—such as a sequence of transactions like a credit card history—and measure each transaction against the person's record as a whole. It can do this in addition to using the general learnings of the feed-forward neural-network model.

Which sectors can benefit from machine learning and deep learning?

McKinsey collated more than 400 use cases of machine and deep learning across 19 industries and nine business functions. Nearly all industries can benefit from machine and deep learning. Here are a few examples of use cases that cut across several sectors:

— **Predictive maintenance**
  Predictive maintenance is an important part of any industry or business relying on equipment. Rather than waiting until a piece of equipment breaks down, companies can use predictive maintenance to project when maintenance will be needed, thereby preventing downtime and reducing operating costs. Machine learning and deep learning have the capacity to analyze large amounts of multifaceted data, which can increase the precision of predictive maintenance. For example, AI practitioners can layer in data from new inputs, like audio and image data, which can add nuance to a neural network's analysis.

— **Logistics optimization**
  Using AI to optimize logistics can reduce costs through real-time forecasts and behavioral coaching. For example, AI can optimize routing of delivery traffic, improving fuel efficiency and reducing delivery times.

— **Customer service**
  AI techniques in call centers can enable a more seamless experience for customers and more efficient processing. The technology goes beyond understanding a caller's words: deep-learning analysis of audio can assess a customer's tone. If a caller is getting upset, the system can reroute to a human operator or manager.

What is generative AI?

Generative AI is an AI model that generates content in response to a prompt. It's clear that generative-AI tools like ChatGPT and DALL-E (a tool for making AI-generated art) have the potential to change how a range of jobs are performed. The full scope of that impact, though, is still unknown—as are the risks. But there are some questions we can answer—like how generative-AI models are built, what kinds of problems they are best suited to solve, and how they fit into the broader category of AI and machine learning.

How can businesses put generative AI to use?

You've probably seen that generative-AI tools like ChatGPT can generate endless hours of entertainment. The opportunity is clear for businesses as well. Generative-AI tools can produce a wide variety of credible writing in seconds, then respond to a user's critiques to make the writing more fit for purpose. This has implications for a broad range of industries, from IT and software organizations that can benefit from the instantaneous code generated by AI models to organizations in need of marketing copy. In short, any organization that needs to produce drafts of clearly written materials potentially stands to benefit. Organizations can also use generative AI to create more technical materials, such as higher-resolution versions of medical images. And with the time and resources saved, organizations can pursue new business opportunities and the chance to create more value.

But developing a proprietary generative-AI model is so resource intensive that it is out of reach for all but the biggest and best-resourced companies. To put generative AI to work, companies can either use generative-AI solutions out of the box or fine-tune them to perform a specific task. If you need to prepare slides according to a specific style, for example, you could ask the model to "learn" how headlines are normally written based on the data in the slides, then feed it slide data and ask it to write appropriate headlines.

Generative AI is not without its risks. Generative-AI models will confidently produce inaccurate, plagiarized, or biased results, without any indication that its outputs may be problematic. That's because the models have been trained on the internet, which is hardly a universally reliable source. Leaders should be aware of these risks before turning to generative AI as a business solution. For more on the risks of generative AI, and how businesses can mitigate them, see the section below called "What are the limitations of AI models, and how can they be overcome?"

What are some specific business use cases for generative AI?

Generative-AI models are in the very early days of scaling, but we've started to see the first batch of applications across functions:

— *Marketing and sales.* Generative AI can craft personalized marketing, social-media, and technical-sales content, including text, images, and video.

— *Operations.* AI models can generate task lists for efficient execution of a specific activity.

— *IT/engineering.* Generative AI can write, document, and review code.

Case study: Vistra Corp. and the Martin Lake Power Plant

Vistra is a large power producer in the United States, operating plants in 12 states with a capacity to power nearly 20 million homes. Vistra has committed to achieving net-zero emissions by 2050. In support of this goal, as well as to improve overall efficiency, QuantumBlack, AI by McKinsey worked with Vistra to build and deploy an AI-powered heat rate optimizer (HRO).

"Heat rate" is a measure of the thermal efficiency of the plant; in other words, it's the amount of fuel required to produce each unit of electricity. To reach the optimal heat rate, plant operators continuously monitor and tune hundreds of variables, such as steam temperatures, pressures, oxygen levels, and fan speeds.

Vistra and a McKinsey team, including data scientists and machine-learning engineers, built a multilayered neural-network model. The model combed through two years' worth of data at the plant and learned which combination of factors would optimize the algorithm and attain the most efficient heat rate at any point in time. When the models were accurate to 99 percent or higher and run through a rigorous set of real-world tests, the team converted them into an AI-powered engine that generates recommendations every 30 minutes for operators to improve the plant's heat-rate efficiency. One seasoned operations manager at the company's plant in Odessa, Texas, said, "There are things that took me 20 years to learn about these power plants. This model learned them in an afternoon."

Overall, the AI-powered HRO helped Vistra achieve the following metrics:

— approximately 1.6 million tons of carbon abated annually

— 67 power generators optimized

— $60 million saved in about a year

— *Risk and legal.* AI models can answer complex questions, based on vast amounts of legal documentation, and draft and review annual reports.

— *R&D.* Generative AI can help accelerate drug discovery through better understanding of diseases and discovery of chemical structures.

While generative AI on its own has a great deal of potential, it's likely to be most powerful in combination with humans, who can help it achieve faster and better work.

How is the use of AI expanding?

AI is a big story for all kinds of businesses, but some companies are clearly moving ahead of the pack. McKinsey's state of AI in 2022 survey showed that adoption of AI models has more than doubled since 2017—and investment has increased apace. What's more, the specific areas in which companies see value from AI have evolved, from manufacturing and risk to these:

— marketing and sales

— product and service development

— strategy and corporate finance

And one set of companies continues to pull ahead of its competitors, by making larger investments in AI, leveling up its practices to scale faster, and hiring and upskilling the best AI talent. More specifically, this group of leaders is more likely to link AI strategy to business outcomes and "industrialize" AI operations by designing modular data architecture that can quickly accommodate new applications.

What are the limitations of AI models, and how can they be overcome?

Since they are so new, we have yet to see the long-tail effect of AI models. This means there are some inherent risks involved in using them—some known and some unknown.

The outputs AI models produce may often sound extremely convincing. This is by design. But sometimes the information they generate is just plain wrong. Worse, sometimes it's biased (because it's built on the gender, racial, and myriad other biases of the internet and society more generally) and can even be manipulated to enable unethical or criminal activity. For example, ChatGPT won't give you instructions on how to hotwire a car, but if you tell it you need to hotwire a car to save a child, the algorithm will instantly comply. Organizations that rely on generative-AI models should reckon with reputational and legal risks involved in unintentionally publishing biased, offensive, or copyrighted content.

While generative AI on its own has a great deal of potential, it's likely to be most powerful in combination with humans, who can help it achieve faster and better work.

These risks can be mitigated, however, in a few ways. For one, it's crucial to carefully select the initial data used to train these models to avoid including toxic or biased content. Next, rather than deploying an off-the-shelf generative-AI model, organizations could consider using smaller, specialized models. Organizations with more resources could also customize a general model based on their own data to fit their needs and minimize biases. Organizations should also keep a human in the loop (that is, make sure a real human checks the output of a generative-AI model before it is published or used) and avoid using generative-AI models for critical decisions, such as those involving significant resources or human welfare.

How can organizations scale up their AI efforts from ad hoc projects to full integration?

Most organizations are dipping a toe into the AI pool—not cannonballing. Slow progress toward widespread adoption is likely due to cultural and organizational barriers. But leaders who effectively break down these barriers will be best placed to capture the opportunity of the AI era. And—crucially—companies that are not making the most of AI are being overtaken by those that are, in industries such as auto manufacturing and financial services.

To scale up AI, organizations can make three major shifts:

1. **Move from siloed work to interdisciplinary collaboration.** AI projects shouldn't be limited to discrete pockets of organizations. Rather, AI is most effective when it's being used by different teams with a range of varied talents to help ensure that AI addresses broad business priorities.

2. **Empower frontline data-based decision making.** AI has the potential to enable faster, better decisions at all levels of an organization. To put this into practice, employees must be able to trust what the algorithm suggests and feel empowered to act accordingly.

3. **Adopt and bolster an agile mindset.** The agile test-and-learn mindset can help employees view errors as inspiration, allaying the fear of failure and speeding up development.

Learn more about McKinsey's Digital Practice, and check out AI-related job opportunities if you're interested in working at McKinsey.

Articles referenced:

— "What is agile?," March 27, 2023

— "What is generative AI?," January 19, 2023

— "Tech highlights from 2022—in eight charts," December 22, 2022

— "Generative AI is here: How tools like ChatGPT could change your business," December 20, 2022, Michael Chui, Roger Roberts, and Lareina Yee

— "The state of AI in 2022—and a half decade in review," December 6, 2022, Michael Chui, Bryce Hall, Helen Mayhew, Alex Singla, and Alex Sukharevsky

— "Why businesses need explainable AI—and how to deliver it," September 29, 2022, Liz Grennan, Andreas Kremer, Alex Singla, and Peter Zipparo

— "Why digital trust truly matters," September 12, 2022, Jim Boehm, Liz Grennan, Alex Singla, and Kate Smaje

— "McKinsey Technology Trends Outlook 2022," August 24, 2022, Michael Chui, Roger Roberts, and Lareina Yee

— "An AI power play: Fueling the next wave of innovation in the energy sector," May 12, 2022, Barry Boswell, Sean Buckley, Ben Elliott, Matias Melero, and Micah Smith

- “Scaling AI like a tech native: The CEO’s role,” October 13, 2021, Jacomo Corbo, David Harvey, Nicolas Hohn, Kia Javanmardian, and Nayur Khan
- “Winning with AI is a state of mind,” April 30, 2021, Thomas Meakin, Jeremy Palmer, Valentina Sartori, and Jamie Vickers
- “A smarter way to digitize maintenance and reliability,” April 23, 2021, Guillaume Decaix, Matthew Gentzel, Andy Luse, Patrick Neise, and Joel Thibert
- “Breaking through data-architecture gridlock to scale AI,” January 26, 2021, Sven Blumberg, Jorge Machado, Henning Soller, and Asin Tavakoli
- “An executive’s guide to AI,” November 17, 2020, Michael Chui, Brian McCarthy, and Vishnu Kamalnath
- “Executive’s guide to developing AI at scale,” October 28, 2020, Nayur Khan, Brian McCarthy, and Adi Pradhan
- “The analytics academy: Bridging the gap between human and artificial intelligence,” *McKinsey Quarterly*, September 25, 2019, Solly Brown, Darshit Gandhi, Louise Herring, and Ankur Puri
- “Notes from the AI frontier: Applications and value of deep learning,” April 17, 2018, Michael Chui, James Manyika, Mehdi Miremadi, Nicolaus Henke, Rita Chung, Pieter Nel, and Sankalp Malhotra

Designed by McKinsey Global Publishing