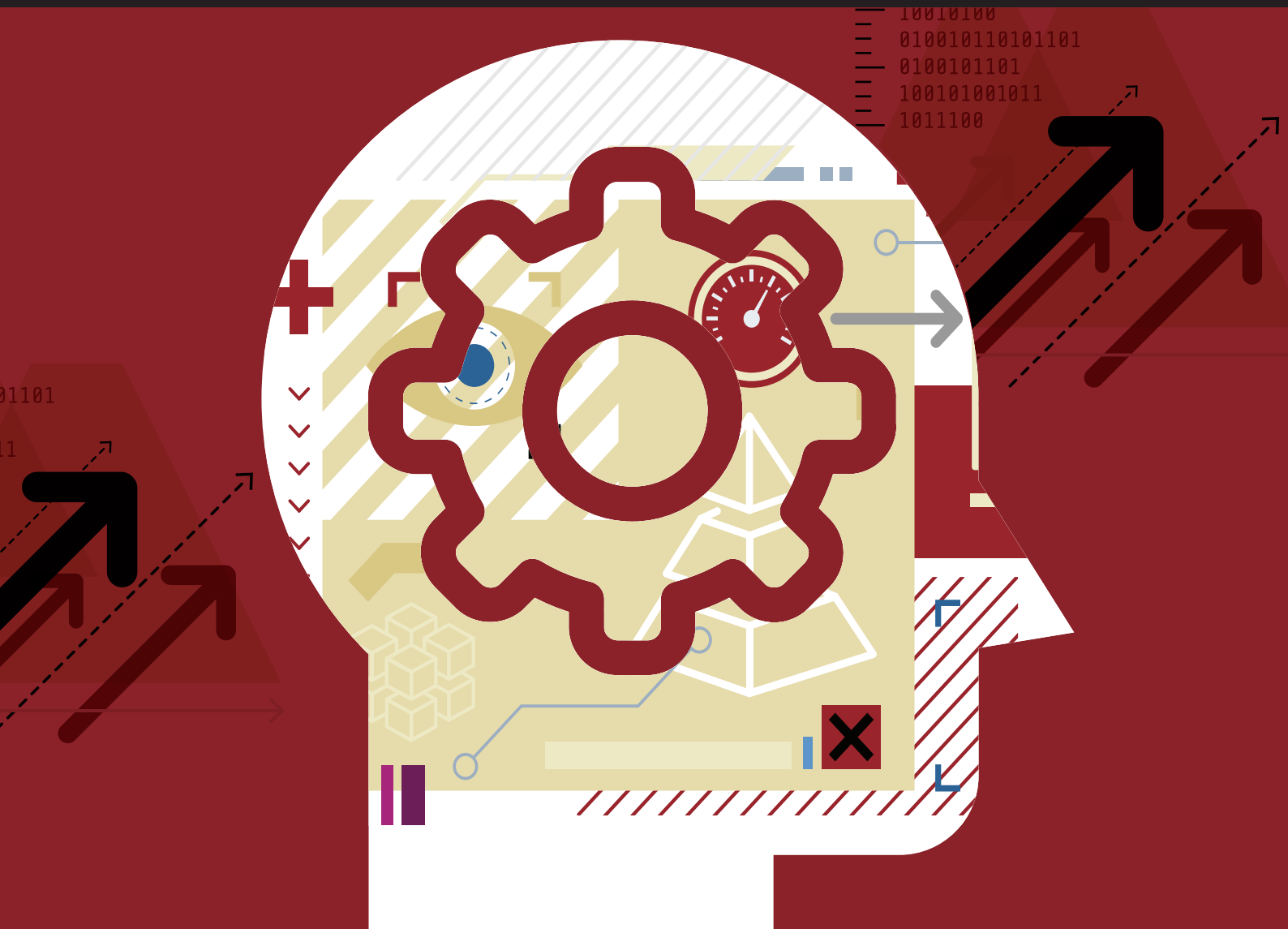
MIT Technology Review Insights

Produced in partnership with

databricks

How technology leaders are adopting emerging tools to deliver enterprise-wide AI.

The great acceleration: CIO perspectives on generative AI



Preface

"The great acceleration: CIO perspectives on generative AI" is an MIT Technology Review Insights report sponsored by Databricks. This report, based on seven in-depth interviews with senior executives and experts, conducted in April and May 2023, seeks to understand how technology leaders are adopting emerging generative AI tools as part of an enterprise-wide AI strategy. The report also draws on an MIT Technology Review Insights global survey of 600 senior data and technology executives, conducted in May and June 2022. Adam Green was the author of the report, Teresa Elsey was the editor, and Nico Crepaldi was the publisher. The research is editorially independent and the views expressed are those of MIT Technology Review Insights.

We would like to thank the following individuals for their time and insights:

Andrew Blyton, Vice President and Chief Information Officer, DuPont Water & Protection

Michael Carbin, Associate Professor, MIT, and Founding Advisor, MosaicML

Owen O'Connell, Senior Vice President and Chief Information Officer (Information Digital Services and Operations), Shell

Noriko Rzonca, Chief Digital Officer, Cosmo Energy Holdings

Richard Spencer Schaefer, Chief Health Informatics Officer, Kansas City VA Medical Center, and AI Solutions Architect, National Artificial Intelligence Institute, U.S. Department of Veterans Affairs

Cynthia Stoddard, Senior Vice President and Chief Information Officer, Adobe

Matei Zaharia, Cofounder and Chief Technology Officer, Databricks, and Associate Professor of Computer Science, University of California, Berkeley

CONTENTS

01 Executive summary 4

02 AI everywhere, all at once 6

03 Building for AI 10

04 Buy, build? Open, closed? 13

05 Workforce worries? 16

06 Risks and responsibilities 19

07 Conclusion 22

01 Executive summary

The emergence of consumer-facing generative AI tools in late 2022 and early 2023 radically shifted public conversation around the power and potential of AI. Though generative AI had been making waves among experts since the introduction of GPT-2 in 2019, it is just now that its revolutionary opportunities have become clear to enterprise. The weight of this moment – and the ripple effects it will inspire – will reverberate for decades to come.

The impact of generative AI on economies and enterprise will be revolutionary. McKinsey Global Institute estimates that generative AI will add between $2.6 and $4.4 trillion in annual value to the global economy, increasing the economic impact of AI as a whole by 15 to 40%.[1] The consultancy projects that AI will automate half of all work between 2040 and 2060, with generative AI pushing that window a decade earlier than previous estimates. Goldman Sachs predicts a 7% – or nearly $7 trillion – increase in global GDP attributable to generative AI, and the firm expects that two-thirds of U.S. occupations will be affected by AI-powered automation.[2]

Text-generating AI systems, such as the popular ChatGPT, are built on large language models (LLMs). LLMs train on a vast corpus of data to answer questions or perform tasks based on statistical likelihoods. Rather than searching and synthesizing answers, they use mathematical models to predict the most likely next

> "I can't think of anything that's been more powerful since the desktop computer."
>
> Michael Carbin, Associate Professor, MIT, and Founding Advisor, MosaicML

word or output.[3] "What was exciting to me, when I first interacted with ChatGPT, was how conversant it was," says Michael Carbin, associate professor at MIT and founding advisor at MosaicML. "I felt like, for the first time, I could communicate with a computer and it could interpret what I meant. We can now translate language into something that a machine can understand. I can't think of anything that's been more powerful since the desktop computer."

Although AI was recognized as strategically important before generative AI became prominent, our 2022 survey found CIOs' ambitions limited: while 94% of organizations were using AI in some way, only 14% were aiming to achieve "enterprise-wide" AI by 2025. By contrast, the power of generative AI tools to democratize AI – to spread it through every function of the enterprise, to support every employee, and to engage every customer – heralds an inflection point where AI can grow from a technology employed for particular use cases to one that truly defines the modern enterprise.

As such, chief information officers and technical leaders will have to act decisively: embracing generative AI to seize its opportunities and avoid ceding competitive ground, while also making strategic decisions about data infrastructure, model ownership, workforce structure, and AI governance that will have long-term consequences for organizational success.

This report explores the latest thinking of chief information officers at some of the world's largest and best-known companies, as well as experts from the public, private, and academic sectors. It presents their thoughts about AI against the backdrop of our global survey of 600 senior data and technology executives.[4]

Key findings include the following:

- **Generative AI and LLMs are democratizing access to artificial intelligence, finally sparking the beginnings of truly enterprise-wide AI.** Powered by the potential of newly emerging use cases, AI is finally moving from pilot projects and "islands of excellence" to a generalized capability integrated into the fabric of organizational workflows. Technology teams no longer have to "sell" AI to business units; there is now significant "demand pull" from the enterprise.

- **A trove of unstructured and buried data is now legible, unlocking business value.** Previous AI initiatives had to focus on use cases where structured data was ready and abundant; the complexity of collecting, annotating, and synthesizing heterogeneous datasets made wider AI initiatives unviable. By contrast, generative AI's new ability to surface and utilize once-hidden data will power extraordinary new advances across the organization.

- **The generative AI era requires a data infrastructure that is flexible, scalable, and efficient.** To power these new initiatives, chief information officers and technical leads are embracing next-generation data infrastructures. More advanced approaches, such as data lakehouses, can democratize access to data and analytics, enhance security, and combine low-cost storage with high-performance querying.

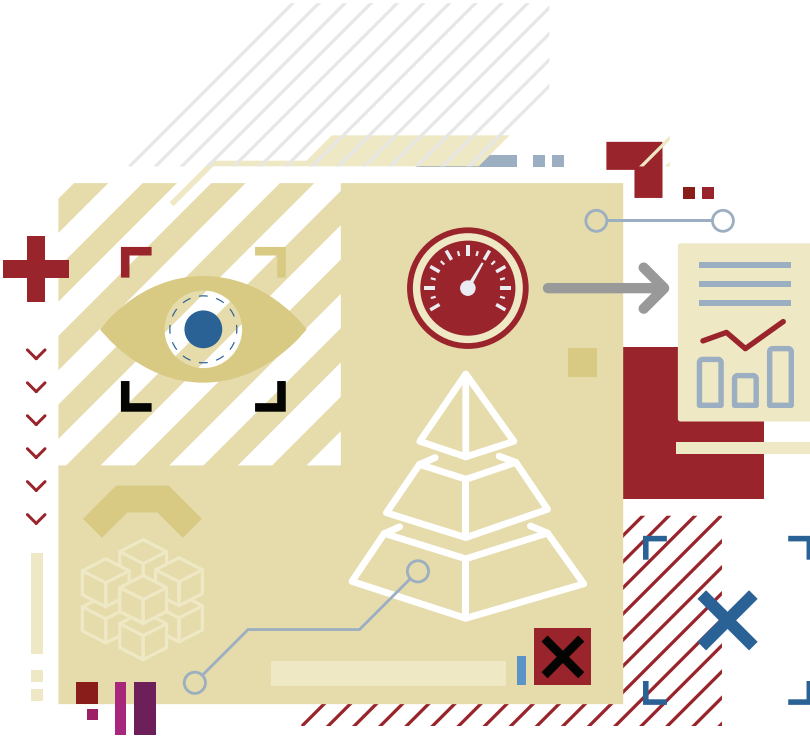
- **Some organizations seek to leverage open-source technology to build their own LLMs, capitalizing on and protecting their own data and IP.** CIOs are already cognizant of the limitations and risks of third-party services, including the release of sensitive intelligence and reliance on platforms they do not control or have visibility into. They also see opportunities around developing customized LLMs and realizing value from smaller models. The most successful organizations will strike the right strategic balance based on a careful calculation of risk, comparative advantage, and governance.

- **Automation anxiety should not be ignored, but dystopian forecasts are overblown.** Generative AI tools can already complete complex and varied workloads, but CIOs and academics interviewed for this report do not expect large-scale automation threats. Instead, they believe the broader workforce will be liberated from time-consuming work to focus on higher value areas of insight, strategy, and business value.

- **Unified and consistent governance are the rails on which AI can speed forward.** Generative AI brings commercial and societal risks, including protecting commercially sensitive IP, copyright infringement, unreliable or unexplainable results, and toxic content. To innovate quickly without breaking things or getting ahead of regulatory changes, diligent CIOs must address the unique governance challenges of generative AI, investing in technology, processes, and institutional structures.

02

AI everywhere, all at once

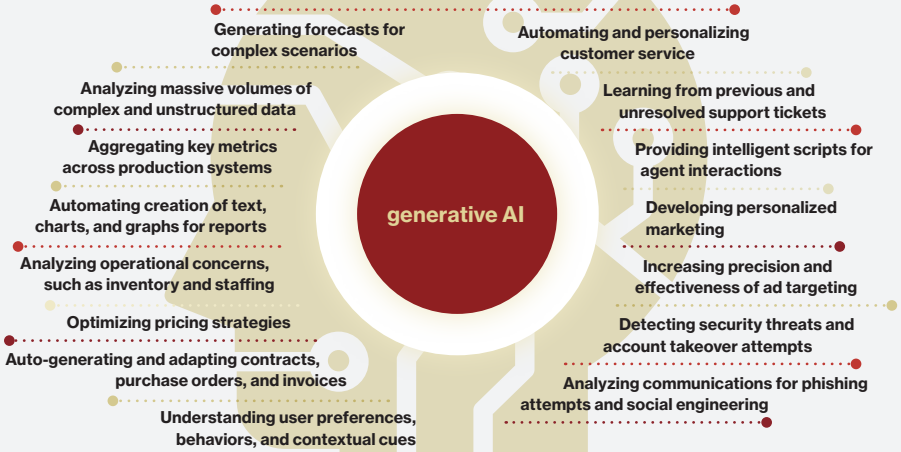


Thanks to its flexibility and range, and user-friendly, natural language–based interface, generative AI is showing its mettle everywhere from copywriting to coding. Its power and potential to revolutionize how work is done across industries and business functions suggests it will have a reverberating impact like that of the personal computer, the Internet, or the smartphone, launching entirely new business models, birthing new industry leaders, and making it impossible to remember how we worked before its spread (see Figures 1 and 2).

This sudden focus on generative AI's power and potential represents a marked shift in how enterprise thinks about AI: until recently, AI adoption was uneven across industries and between functions within companies.

Figure 1: Enterprise applications and use cases for generative AI

These are just a few of the business functions compellingly addressed by generative AI.



Source: Compiled by MIT Technology Review Insights, based on data from "Retail in the Age of Generative AI ,"[5] "The Great Unlock: Large Language Models in Manufacturing,"[6] "Generative AI Is Everything Everywhere, All at Once,"[7] and "Large Language Models in Media & Entertainment,"[8] Databricks, April–June 2023.

Our 2022 survey found just 8% of respondents saying that AI was a critical part of three or more business functions. Only in finance and IT did more than half of respondents say AI was a critical or widespread part of the function.

Before the advent of generative AI, only a rare few organizations had made AI a critical capability across the business, or even aspired to. While 94% of organizations were already using AI in some way, only 14% of them aimed to achieve "enterprise-wide AI," defined as having AI being a critical part of at least five core functions, by 2025 (see Figures 3 and 4).

But now generative AI is changing the conversation. Demonstrating applications in every business function, AI is poised to spread enterprise wide. AI is now even venturing assuredly into creative applications, once considered a uniquely human endeavor. Adobe, for example, has launched Firefly, a family of creative generative AI models that act as a copilot to creative and design workflows, according to Cynthia Stoddard, the company's senior vice president and chief information officer. Firefly tools can recolor existing images, generate new images, and edit new objects into and out of images, all based on a text description.

The energy and chemical industries are applying AI in domains that had previously been inaccessible. Multi-industrials giant DuPont, for instance, had worked on chatbots for both employee and consumer interfaces previously, but found their inaccuracy too frustrating. "LLMs now have the capability to achieve the necessary accuracy, and at a faster pace," says Andrew Blyton, vice president and chief information officer at DuPont Water & Protection. The company is now using AI in production scheduling, predictive reliability and maintenance, and sales price optimization applications.

Figure 2: Industry-specific applications and use cases
Every industry – and business – will find its own custom applications of generative AI technology.

Consumer goods and retail

- Providing virtual fitting rooms
- Scheduling delivery and installation
- Providing in-store product-finding assistance
- Optimizing demand prediction and inventory planning
- Generating novel product designs

Manufacturing

- Serving as expert copilot for technicians
- Allowing conversational interactions with machines
- Providing prescriptive and proactive field service
- Enabling natural language troubleshooting
- Assessing warranty status and documentation
- Understanding process bottlenecks and devising recovery strategies

Media and entertainment

- Providing intelligent search and tailored content discovery
- Writing engaging headlines and copy
- Providing real-time feedback on content quality
- Curating personalized playlists, news digests, and recommendations
- Enabling interactive storytelling, driven by viewer choices
- Delivering targeted offers and subscription plans

Financial services

- Uncovering potential trading signals and alerting traders to vulnerable positions
- Accelerating underwriting decisions
- Optimizing and rebuilding legacy systems
- Reverse-engineering banking and insurance models
- Monitoring for potential financial crimes and fraud
- Automating data gathering for regulatory compliance
- Extracting insights from corporate disclosures

Source: Compiled by MIT Technology Review Insights, based on data from "Retail in the Age of Generative AI,"[9] "The Great Unlock: Large Language Models in Manufacturing,"[10] "Generative AI Is Everything Everywhere, All at Once,"[11] and "Large Language Models in Media & Entertainment,"[12] Databricks, April–June 2023.

Figure 3: The great reconsideration

Before generative AI, few organizations had adopted AI as a critical part of any business function.

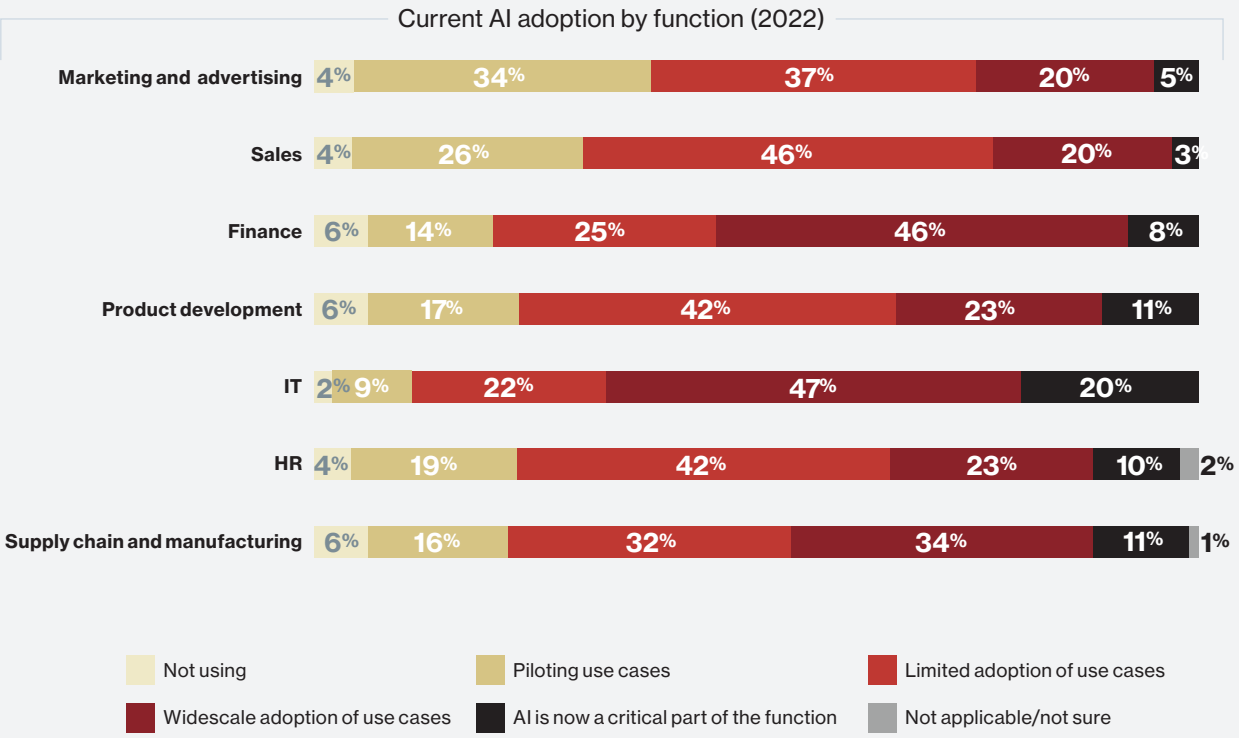
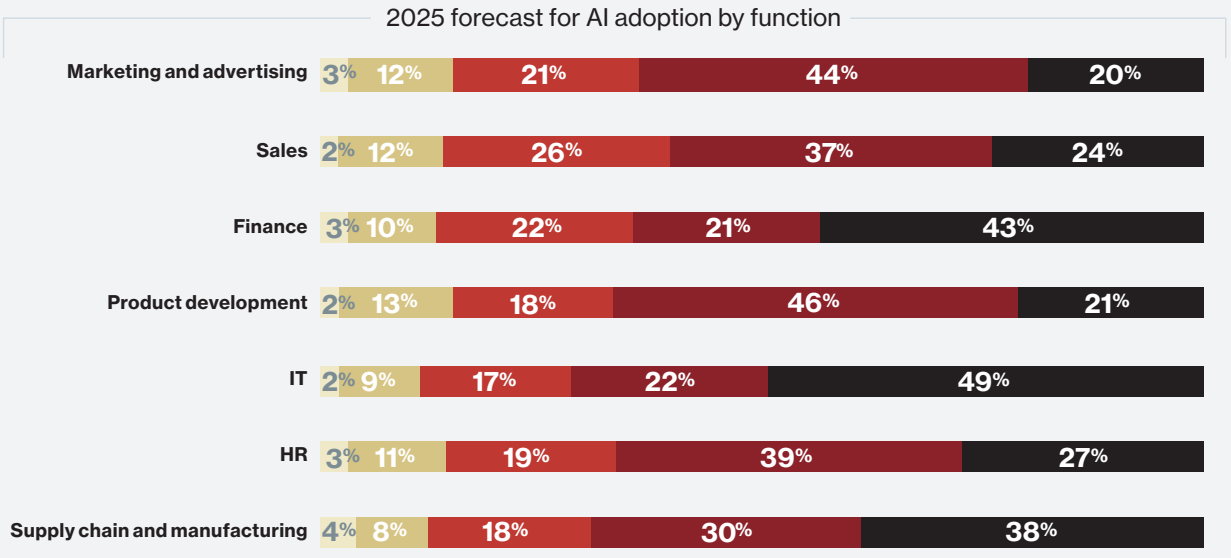


Figure 4: Limited AI ambitions

Before generative AI, few organizations aimed to make AI a critical capability across key functions by 2025.



Source: MIT Technology Review Insights survey, 2022.

Multinational organizations with assets stretching back decades have historically struggled to unify their digital infrastructure. Mergers and acquisitions have resulted in fragmented IT architectures. Important documents, from research and development intelligence to design instructions for plants, have been lost to view, locked in offline proprietary file types. “Could we interrogate these documents using LLMs? Can we train models to give us insights we’re not seeing in this vast world of documentation?” asks Blyton. “We think that’s an obvious use case.” Language models promise to make such unstructured data much more valuable.

Energy giant Shell concurs: “We’ve started to see benefits because a lot of documentation previously in repositories is now starting to come together,” says Owen O’Connell, senior vice president and chief information officer for information digital services and operations. The firm is also streamlining legal, regulatory, and human resources paperwork covering its many jurisdictions, and even surfacing insights from unstructured data in areas like recruitment and performance.

> “LLMs now have the capability to achieve the necessary accuracy, and at a faster pace.”
>
> Andrew Blyton, Vice President and Chief Information Officer, DuPont Water & Protection

High hopes for health care

Health and medicine, a deeply human-centered field, has also been a productive testing ground for AI. In the lab, AI-powered tools have proven their mettle at predicting protein structures, aiding in drug discovery, and tracking the spread of outbreaks, including the covid-19 pandemic. Generative AI is now showing its promise as a powerful assistant to front-line staff. Natural language processing tools, for instance, can transcribe and summarize medical notes, while chatbots might be trained to answer consumer medical questions.

Richard Spencer Schaefer, chief health informatics officer at the Kansas City VA Medical Center, which manages health care for American military veterans, and AI solutions architect at the VA’s National Artificial Intelligence Institute, is optimistic that automation and predictive analytics could help the sector. The VA recently validated a model that identified the 24-hour risk of a patient admitted to hospital subsequently being transferred to a higher level of care. “We found a very significant improvement in the accuracy and predictability of using that AI model, to a point where there could be significant reduction in mortality,” he reports.

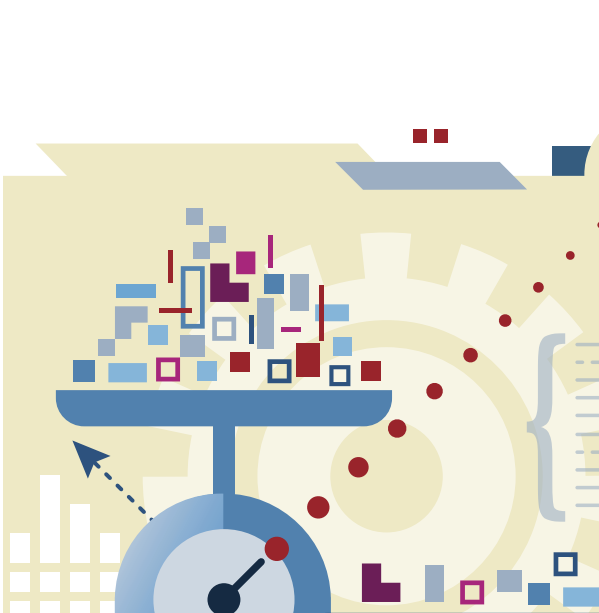
Schaefer is also excited about the power of AI to alleviate health-care employee fatigue while building their trust, via small, practical operational improvements. “I think where you’ll see much quicker adoption in health care,” he says, “is in very specific automated process-type improvements.” The VA, for example, is currently working on an AI-driven project to reduce unnecessary alarms and alerts in its medical centers. Machine learning models can help by defining dynamic and patient-centered thresholds for provider interventions, rather than relying on the alarm systems’ traditional and static guardrails and thresholds. “For this to work,” says Schaefer, “we have to build trust and help health-care workers understand, hey, this is what AI can do.”

03 Building for AI

AI applications rely on a solid data infrastructure that makes possible the collection, storage, and analysis of its vast data-verse. Even before the business applications of generative AI became apparent in late 2022, a unified data platform for analytics and AI was viewed as crucial by nearly 70% of our survey respondents (see Figure 5).

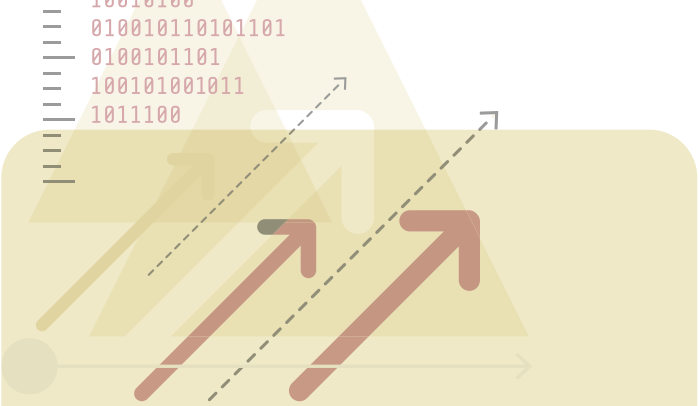
Data infrastructure and architecture covers software and network-related infrastructure, notably cloud or hybrid cloud, and hardware like high-performance GPUs. Enterprises need an infrastructure that maximizes the value of data without compromising safety and security, especially at a time when the rulebook for data protection and AI is thickening. To truly democratize AI, the infrastructure must support a simple interface that allows users to query data and run complex tasks via natural language. "The architecture is moving in a way that supports democratization of analytics," says Schaefer.

Data lakehouses have become a popular infrastructure choice. They are a hybrid of two historical approaches – data warehouses and data lakes. Data warehouses came to prominence in the 1980s to systematize business intelligence and enterprise reporting. However, warehouses do not offer real-time services, operating on a batch processing basis, and cannot accommodate emerging and non-traditional data formats. Data lakes, favored for their ability to support more AI and data science tasks, emerged more recently. But they are complex to construct, slow, and suffer from inferior data quality controls. The lakehouse combines the best of both, offering an open architecture that combines the flexibility and scale of data lakes with the management and data quality of warehouses.



"We have aggregated data across a lot of different technologies over time, and I think what we're finding now is that the lakehouse has the best cost performance straight off."

Andrew Blyton, Vice President and Chief Information Officer, DuPont Water & Protection



Controlling the data faucet

Training Open AI's GPT-3 cost more than $4 million, while the training of Meta's LLaMA model took 1 million GPU hours and cost over $2.4 million.[13] Training an AI model is energy-consuming; it can expend more electricity than 100 typical U.S. homes over an entire year.[14] The training of Open AI's GPT-3 consumed 1,287 MWh of electricity and emitted over 550 tons of carbon dioxide.[15] "These models are some of the most compute-heavy models you can train," says Shell's Owen O'Connell.

Larger models are more expensive to train and run, with the cost of each interaction roughly related to the model's size, says Michael Carbin, of MIT and MosaicML. OpenAI is reportedly spending $40 million a month to process the queries entered by users, while Microsoft's Bing chatbot reportedly needs $4 billion worth of infrastructure to serve its potential user base.[16] More efficient smaller models (see next chapter) may be a promising alternative.

Companies will need to ramp up their attention to efficiency while startups and researchers are pioneering new ways to improve AI model efficiency. "As our costs grow, we're hiring engineers who know how to optimize and control costs," says DuPont's Andrew Blyton. "At some point, we get much more efficient on a query or compute basis because we can hire experts to optimize." DuPont also tracks and analyzes cloud computing costs at a granular level. "We have some talented internal capability, and we rely on our vendors to ensure we're optimizing everything. The wrong approach here can have a significant impact on both time and performance," says Blyton.

For the VA, the lakehouse is appealing because it minimizes the need to move data, which creates privacy and governance risks. "We face a significant challenge in data movement, which is why we are dedicating substantial resources and expertise to address it," says Schaefer.

DuPont has invested in multiple technology infrastructures and has found the data lakehouse promising. "We have aggregated data across a lot of different technologies over time," says Blyton, "and I think what we're finding now is that the lakehouse has the best cost performance straight off, and hence we've started investing more heavily in scaling." For DuPont, a data lakehouse provides visibility into a complex organizational portfolio that includes years of M&A activity and technical debt. "Making sense of the data of our business has been the main reason to invest in the data lakehouse approach," says Blyton. "How do I take multiple enterprise resource planning systems (ERPs) of data, merge them together, and give people almost real-time access to information that previously was being done manually?"

The lakehouse abstracts complexity in a way that allows users to perform advanced activities regardless of technical competency. Shell has built an "enterprise layer" that allows users to interact dynamically. "Previously, you had to go to data stores, extract the data, cleanse it, and do multiple transform activities," says O'Connell. The lakehouse approach allows users to run workloads and do historical analysis and trend-matching themselves, while cloud services provide computational elasticity.

The data lakehouse supports both structured and unstructured data, says O'Connell. "Typical databases are designed for only one type of data," he says. Lakehouse "allows us to move much quicker. We have all that data coming together into a common architecture that people can mix together and find business value from." This is crucial for a data-intensive company like Shell. For predictive maintenance alone, the company has 17,000 models and 4 trillion rows of data – adding another 20 billion per day from 3 million sensors across facilities around the world.

Data lakehouse adopters are using the technology as part of a data strategy that enables practical and impactful data and AI use cases. Says Schaefer,

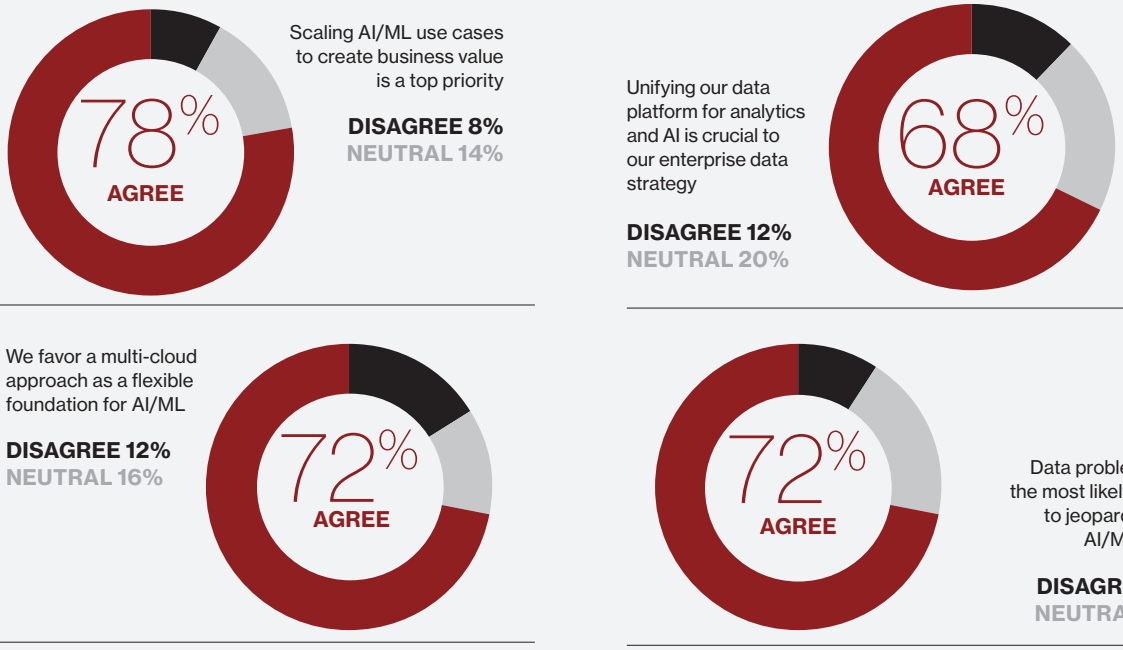
"Democratizing the data is pretty expensive to do, and historically there weren't good resources for that. But we're rapidly entering the era of lakehouse and hybrid lake transactional models and software tools that are finally playing together well enough. I think in health care particularly we'll see a pretty rapid adoption of tools that improve operational efficiencies because we have a lot of opportunity to improve the way that we do things and a lot of opportunity for process automation."

Blyton adds that, at DuPont, "we are using the lakehouse and tools like the data lake to build a data foundation and apply ML algorithms to it. We are expanding predictive maintenance using algorithms with third-party sensors to figure out when a machine's going to break before it breaks, as opposed to after it's broken. These are good rock-solid business cases, real value being generated from the investment that's producing a tangible and a measurable return for us."

"Typical databases are designed for only one type of data. Lakehouse allows us to move much quicker."

Owen O'Connell, Senior Vice President and Chief Information Officer (Information Digital Services and Operations), Shell

Figure 5: AI priorities for executives



Source: MIT Technology Review Insights survey, 2022.

04 Buy, build? Open, closed?

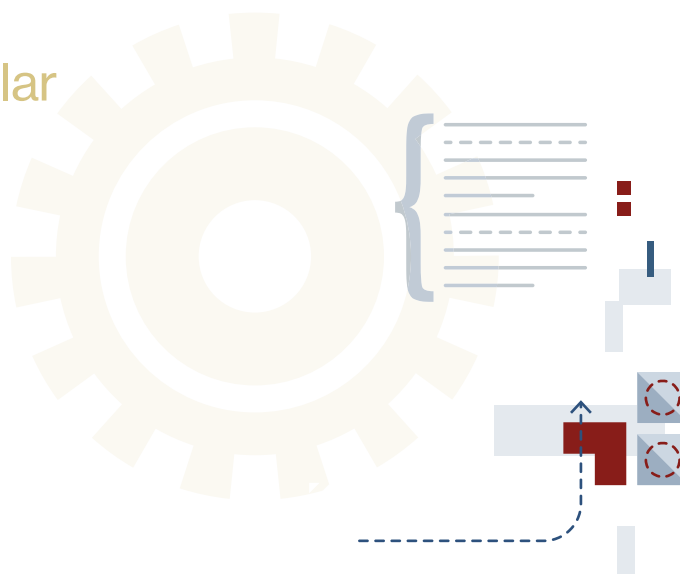
Today's CIOs and leadership teams are re-evaluating their assumptions on ownership, partnership, and control as they consider how to build on the capabilities of third-party generative AI platforms.

Over-leveraging a general-purpose AI platform is unlikely to confer competitive advantage. Says Carbin, "If you care deeply about a particular problem or you're going to build a system that is very core for your business, it's a question of who owns your IP." DuPont is "a science and innovation company," adds Blyton, "and there is a need to keep LLM models internal to our organization, to protect and secure our intellectual property – this is a critical need."

Creating competitive risk is one worry. "You don't necessarily want to build off an existing model where the data that you're putting in could be used by that company to compete against your own core products," says Carbin. Additionally, users lack visibility into the data, weightings, and algorithms that power closed models, and the product and its training data can change at any time. This is particularly a concern in scientific R&D where reproducibility is critical.[17]

Some CIOs are taking steps to limit company use of external generative AI platforms. Samsung banned ChatGPT after employees used it to work on commercially sensitive code. A raft of other companies, including JP Morgan Chase, Amazon, and Verizon,

> "If you care deeply about a particular problem or you're going to build a system that is very core for your business, it's a question of who owns your IP."
>
> Michael Carbin, Associate Professor, MIT, and Founding Advisor, MosaicML



have enacted restrictions or bans. "We can't allow things like Shell's corporate strategy to be flowing through ChatGPT," says O'Connell. And since LLMs harness the totality of the online universe, he believes that companies will in the future be more cautious about what they put online in the first place: "They are realizing, hang on, someone else is going to derive a lot of value out of my data."

Inaccurate and unreliable outputs are a further worry. The largest LLMs are, by dint of their size, tainted by false information online. That lends strength to the argument for more focused approaches, according to Matei Zaharia, cofounder and chief technology officer at Databricks and associate professor of computer science at the University of California, Berkeley. "If you're doing something in a more focused domain," he says, "you can avoid all the random junk and unwanted information from the web."

Companies cannot simply produce their own versions of these extremely large models. The scale and costs put this kind of computational work beyond the reach of all but the largest organizations: OpenAI reportedly used 10,000 GPUs to train ChatGPT.[18] In the current moment, building large-scale models is an endeavor for only the best-resourced technology firms.

Smaller models, however, provide a viable alternative. "I believe we're going to move away from 'I need half a trillion parameters in a model' to 'maybe I need 7, 10, 30, 50 billion parameters on the data that I actually have,'" says Carbin. "The reduction in complexity comes by narrowing your focus from an all-purpose model that knows all of human knowledge to very high-quality knowledge just for you, because this is what individuals in businesses actually really need."

Thankfully, smaller does not mean weaker. Generative models have been fine-tuned for domains requiring less data, as evidenced through models like BERT—for biomedical content (BioBERT), legal content (Legal-BERT), and French text (the delightfully named CamemBERT).[19] For particular business use cases, organizations may choose to trade off broad knowledge for specificity in their business area. "People are really looking for models that are conversant in their domain," says Carbin. "And once you make that pivot, you start to realize there's a different way that you can operate and be successful."

Dolly: The $30 conversational LLM

In March 2023, Databricks released Dolly, an open-source LLM aimed at democratizing the power of generative AI.[20] Trained for less than $30, it demonstrated ChatGPT-like conversational ability (namely, the ability to follow user instructions). It is based on the LLaMA tool from Meta, fine-tuned on high-quality inputs crowdsourced from Databricks employees. Dolly, named after the world's first cloned mammal, has just 6 billion parameters—less than 3.5% of the 175 billion used in GPT-3.

"The highest-end models, with hundreds of billions of parameters, can be very expensive to train and run," notes Databricks's Matei Zaharia. These costs leave many companies relying on proprietary tools like ChatGPT or spending huge sums and resources to build their own. While improved hardware and software efficiencies will lower costs over time, building models with trillions of parameters will still be beyond reach for many, he observes.

To provide an alternative to AI tools becoming centralized in just a few large companies, Databricks sought to develop an open-source chat model that used open-source code and data but also allowed commercial use. "The main thing we wanted to investigate," says Zaharia, "is how expensive is it to create one using just public data sets that are out there, using open-source code and using or collecting training data that is open-source and is runnable by everyone."

Databricks open-sourced the training code, data set, and model weights for Dolly. Its April 2023 release, Dolly 2.0, is the first open-source LLM licensed for commercial use.[21] This allows companies to combine their own data with the Databricks data set to build personalized applications that will not compromise their IP or corporate information.

By enabling companies to cheaply build and customize their own tools, platforms like Dolly can democratize access to generative AI.

> “All the large models that you can get from third-party providers are trained on data from the web. But within your organization, you have a lot of internal concepts and data that these models won’t know about.”
>
> Matei Zaharia, Cofounder and Chief Technology Officer, Databricks, and Associate Professor of Computer Science, University of California, Berkeley

“People are starting to think a lot more about data as a competitive moat,” says Zaharia. “Examples like BloombergGPT [a purpose-built LLM for finance] indicate that companies are thinking about what they can do with their own data, and they are commercializing their own models.”

“Companies are going to extend and customize these models with their own data, and to integrate them into their own applications that make sense for their business,” predicts Zaharia. “All the large models that you can get from third-party providers are trained on data from the web. But within your organization, you have a lot of internal concepts and data that these models won’t know about. And the interesting thing is the model doesn’t need a huge amount of additional data or training time to learn something about a new domain,” he says.

Smaller open-source models, like Meta’s LLaMA, could rival the performance of large models and allow practitioners to innovate, share, and collaborate. One team built an LLM using the weights from LLaMA at a cost of less than $600, compared to the $100 million involved in training GPT-4. The model, called Alpaca, performs as well as the original on certain tasks.[22] Open source’s greater transparency also means researchers and users can more easily identify biases and flaws in these models.

“Much of this technology can be within the reach of many more organizations,” says Carbin. “It’s not just the OpenAIs and the Googles and the Microsofts of the world, but more average-size businesses, even startups.”

05 Workforce worries

From the telephone to the desktop computer, every leap in everyday technology sparks worries about unemployment and threats to human craft and skill. But generative AI appears to mark a new phase given the unprecedented range of tasks that can now feasibly be automated. Uniquely, some of the most automation-vulnerable sectors today are in high-end technical and creative fields.

An Accenture analysis determined that 40% of working hours across industries could be automated or augmented by generative AI, with banking, insurance, capital markets, and software showing the highest potential.[23] McKinsey projects that generative AI and related technologies could automate activities that currently take up 60 to 70% of worker time – up from 50% before the advent of generative AI.[24]

Goldman Sachs predicts that two-thirds of U.S. occupations will be affected by AI-powered automation, but does not expect this to lead to widespread job loss: "Most jobs and industries," the firm says, "are only partially exposed to automation, and are thus more likely to be complemented rather than substituted by AI." It goes on to note that technology-related job losses tend to be balanced – and exceeded – by technology-related job growth: "more than 85% of employment growth over the last 80 years is explained by the technology-driven creation of new positions," the authors write.[25]

The CIOs and academics interviewed for this report offer an optimistic view on generative AI's effects for enterprises and society at large. They believe AI could help sectors like health care where the workforce is stretched, and that human experts will remain essential. "People are worried that they will lose their jobs because of AI, but that's not the primary concern in health care," says Schaefer. "The technologies that we're putting in place are enabling physicians to be a part of the

> "The technologies that we're putting in place are enabling physicians to be a part of the development of AI, and because of the level of validation involved, I think there will be more trust in the models that we develop."
>
> Richard Spencer Schaefer,
> Chief Health Informatics Officer,
> Kansas City VA Medical Center

development of AI, and because of the level of validation involved, I think there will be more trust in the models that we develop." Humans are still necessary to annotate, curate, expose models to data sets, and provision for responsible AI. "We've got a way to go where we can unleash AI to do our research because there is so much hand-holding that happens at the moment," adds O'Connell.

Generative AI will also empower by democratizing access to technical capabilities once confined to a slice of the workforce. Companies are already taking steps to encourage engagement. "We have created communities of practitioners internally who do not necessarily have mainstream IT backgrounds," says Blyton. "We want to empower our business users to craft their own dashboards and drive their own insights from data." In fact, this type of data democratization was already a top way that companies were generating tangible benefits from AI in 2022, according to our survey (see Figure 6).

Blyton also predicts that ideas for AI will start coming from the workforce, marking the start of a more self-service and entrepreneurial era within organizations. According to O'Connell, business unit leaders, aware of AI thanks to the blizzard of coverage over recent months, are already creating a "demand pull" for AI applications, rather than technology teams having to sell AI ideas to the organization.

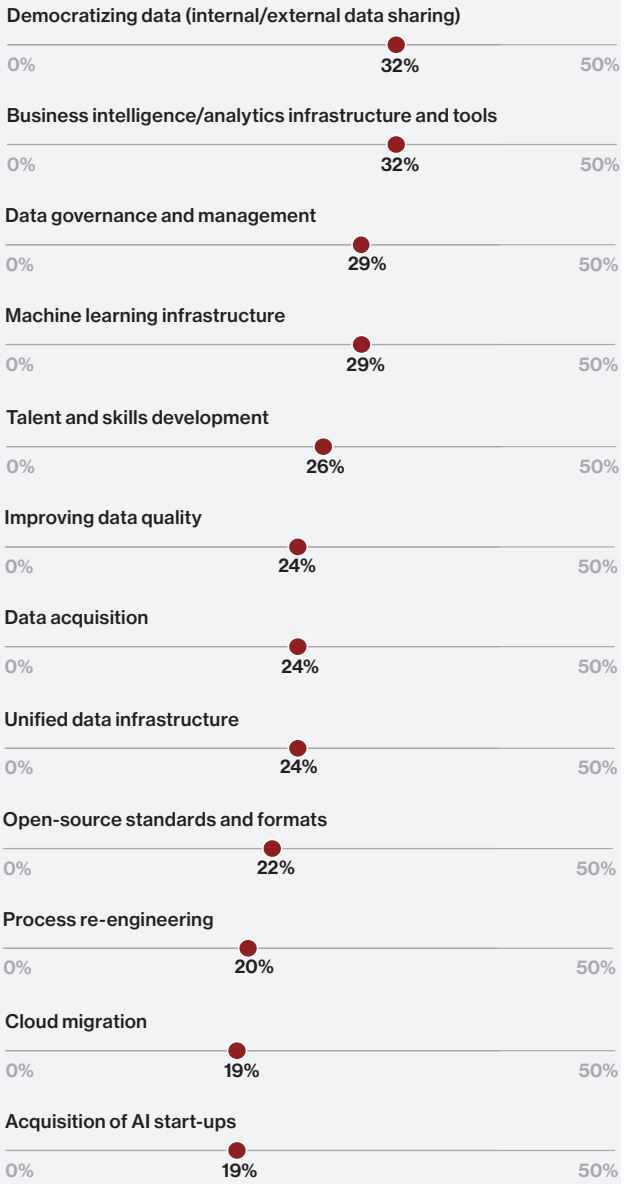
Over the last decade, software programming has been among the most sought-after, highly paid roles in business. Yet this is one domain where generative AI tools are demonstrating utility, prompting concerns. CIOs and technical leads interviewed for this report offer a more optimistic view. They expect rising demand for specialties such as operational efficiency and responsible AI, and they believe programmers will shift their attention to higher value and less tedious work. Blyton asks, "Do I need someone to be laboriously typing six hours' worth of code into an engine and trying to debug it for three days? I can see some huge efficiency by not doing that."

Some expect programmers shift to chaperoning raw initial ideas into more mature and business-ready outputs. "It's not hard to imagine that in the near future, you could use the English language to generate code," predicts Blyton. "In a year or so, my workflow could involve business users doing parts of the analysis in English language, and then I take the output code and

Figure 6: Most instrumental investment areas for data and AI

When asked which investments have been instrumental to creating benefits from AI/ML, business leaders are most likely to say democratizing data access and building business intelligence infrastructure have already paid off.

Percentage of respondents who ranked each option among their top three investments.



Source: MIT Technology Review Insights survey, 2022.

> “In the next five to ten years we will see how quickly we can adapt, and companies that fail to adapt, no matter how big, are going to disappear.”
>
> Noriko Rzonca, Chief Digital Officer, Cosmo Energy Holdings

incorporate it into something more robust.” Stoddard says AI is accelerating a trend already underway: “You need people who understand math and science to work with the models, and you need people who have deep analytic ways of thinking. In my area, the roles have been migrating towards analytics, working with models, understanding deviations and coding.”

CIOs largely frame AI as a copilot rather than a rival to human employees. Zaharia predicts that LLMs will allow executives to acquire high-quality data for decisions, saving time otherwise spent by analysts responding to their on-demand report requests. But he adds, “I don’t think they’ll be able to do extremely complicated work that requires a lot of planning fully automatically. I don’t think there’s anyone whose job is just the super-simple stuff that can be done by a language model.” Stoddard concurs: “Generative AI lets creators at all levels use their own words to generate content,” she says. “But I don’t think it will ever replace humans. Instead, it’s going to be an assistant to humans.”

She adds, “We internally view AI/ML as being a helper, truly helping our people, and then allowing them to spend more time on other value-added activities.” At Adobe, a change management process was necessary to win over the workforce. Says Stoddard, “Back when we started this, I would say people were skeptical, but when they saw that they were truly able to spend more time on value added, and in some cases, their jobs were upskilled, people became believers, and they still are believers.”

The democratization of access to technical tools could be society-wide. “All these people who couldn’t program or manipulate design software, what are the new opportunities for them?” asks Carbin. “This is going to lower the barrier for people accessing technology and programming or manipulating software. Now you have this natural-language interface for design software allowing you to navigate what was otherwise an extremely difficult tool used only by the best creative professionals.”

Risk aversion and cultural factors, like fear of failure, also need to be addressed to drive AI adoption in the workforce. “Sometimes those who have been in or have managed organizations in Japan for a very long time and who are used to a certain style, don’t welcome risk-taking as much as the newer members of the business and commercial realm who constantly offer unique ideas and are eager for change,” says Noriko Rzonca, chief digital officer, Cosmo Energy Holdings. “As a business, it may take a longer time to adapt to and use new initiatives, but we are fortunate in the fact that improvement is something the Japanese are very good at. As we adjust to using AI in business, I am certain we will see great improvement and change to the way businesses in Japan operate. I believe there is huge potential in AI for Japan.”

Rzonca says that organizations will have to embrace data and AI if they are to thrive: “In the next five to ten years we will see how quickly we can adapt, and companies that fail to adapt, no matter how big, are going to disappear.”

06

Risks and responsibilities

AI, and particularly generative AI, comes with governance challenges that may exceed the capabilities of existing data governance frameworks. When working with generative models that absorb and regurgitate all the data they are exposed to, without regard for its sensitivity, organizations must attend to security and privacy in a new way. Enterprises also now must manage exponentially growing data sources and data that is machine-generated or of questionable provenance, requiring a unified and consistent governance approach. And lawmakers and regulators have grown conscious of generative AI's risks, as well, leading to legal cases, usage restrictions, and new regulations, such as the European Union's AI Act.

As such, CIOs would be reckless to adopt AI tools without managing their risks – ranging from bias to copyright infringement to privacy and security breaches. At Shell, "we've been spending time across legal, finance, data privacy, ethics, and compliance to review our policies and frameworks to ensure that they're ready for this and adapted," says O'Connell.

> "If your entire business model is based on the IP you own, protection is everything."
>
> Andrew Blyton, Vice President and Chief Information Officer, DuPont Water & Protection

One concern is protecting privacy at a time when reams of new data are becoming visible and usable. "Because the technology is at an early stage, there is a greater need for large data sets for training, validation, verification, and drift analysis," observes Schaefer. At the VA, "that opens up a Pandora's box in terms of ensuring protected patient health information is not exposed. We have invested heavily in federally governed and secured high-compute cloud resources."

Commercial privacy and IP protection is a related concern. "If your entire business model is based on the IP you own, protection is everything," says Blyton of DuPont. "There are many bad actors who want to get their hands on our internal documentation, and the creation of new avenues for the loss of IP is always a concern."

Another data governance concern is reliability. LLMs are learning engines whose novel content is synthesized from vast troves of content and they do not differentiate true from false. "If there's an inaccurate or out-of-date piece of information, the model still memorizes it. This is a problem in an enterprise setting, and it means that companies have to be very careful about what goes into the model," says Zaharia. Blyton adds, "It's an interesting thing with ChatGPT: you ask it the same question twice and you get two subtly different answers; this does make science and innovation company people raise their eyebrows."

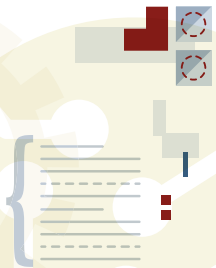
Model explainability is key to earning the trust of stakeholders for AI adoption and for proving the technology's business value. This type of explainability is already a priority in AI governance and regulation, especially as algorithms have begun to play a role in life-changing decisions like credit scoring or reoffending risk in the criminal justice system. Critics argue that AI

systems in sensitive or public interest domains should not be a black box, and "algorithmic auditing" has received growing attention.

"It's pretty obvious there's a lot riding on health-care model explainability," says Schaefer. "We've been working on model cards" – a type of governance documentation that provides standardized info about a model's training, strengths, and weaknesses – "as well as tools for model registries that provide some explainability. Obviously algorithm selection is an important consideration in model explainability, as well."

At Cosmo Energy Holdings, Rzonca considers how to balance strong AI governance with empowerment. "We are developing governance rules, setting them up, and training people," she says. "In the meantime, we are trying to get something easy to adapt to, so people can feel and see the results." To strike the balance, she pairs democratized access to data and AI with centralized governance: "Instead of me doing everything directly, I'm trying to empower everyone around and help them realize they have the individual ability to get things done. Thanks to an empowered team, I can be freed up to

> "Generative AI evolves the possibility and promise of AI exponentially. You can transform the conversation between the creator and the computer."
>
> Cynthia Stoddard, Senior Vice President and Chief Information Officer, Adobe



The creative copilot: AI at Adobe

The creative industries are embracing generative AI to empower design professionals with fascinating new abilities. The software firm Adobe is now offering a suite of generative AI design models called Adobe Firefly, offering creative capabilities that extend the powers of Photoshop, Illustrator, Express, and its enterprise offerings to new heights.

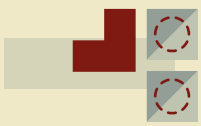
The company has also developed Adobe Sensei GenAI to deliver more speed and productivity in Adobe Experience Cloud, as a copilot for customer experience workflows. This can improve enterprise productivity and efficiency in areas like personalized content, editing, and conversational experience.

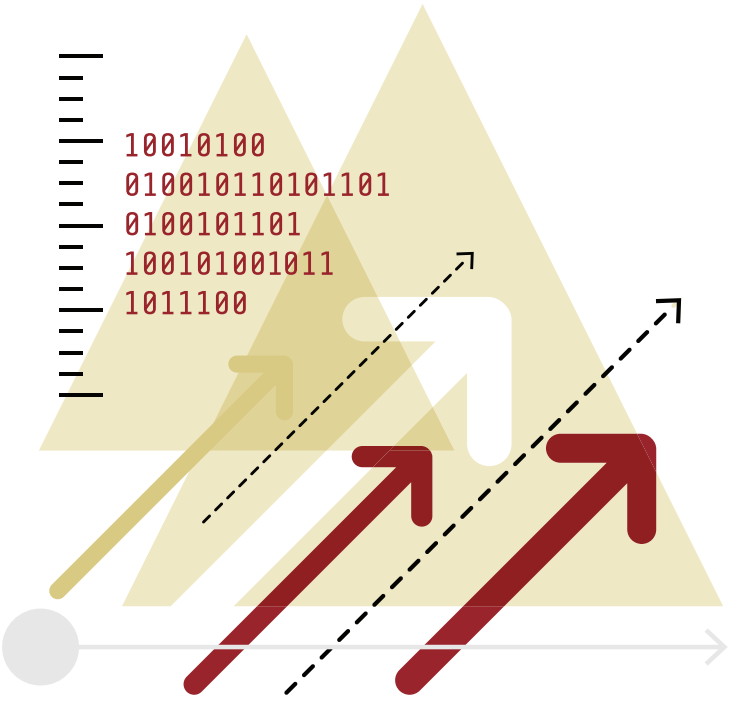
"Generative AI evolves the possibility and promise of AI exponentially," says Adobe senior vice president and CIO Cynthia Stoddard. "The tools are becoming a lot stronger, and you can do more, so you can transform the conversation between the creator and the computer into something that's easy, more natural to use, intuitive."

Adobe is also using AI to optimize its own back-office processes. Internally, the company is deploying automation to sift customer query tickets that can be solved without human intervention, reducing workloads and giving customers a speedier resolution. It has employed AI to catalog the sprawling and often overlapping software solutions used across departments, rationalizing spending and simplifying the technology stack.

Adobe has also built a "self-healing" platform that identifies real or emerging technical issues – the kinds of faults a system administrator might get called out in the middle of the night to fix – and fixing them automatically.

Stoddard says, "as a company, AI is core to delivering experiences to our customers. Then as the IT organization, we're following those same principles to inject capabilities into what we do to make it easier for our constituents."





> "We are seeing the need to have very integrated governance models, integrated governance structures for all data and all models."
>
> Richard Spencer Schaefer,
> Chief Health Informatics Officer,
> Kansas City VA Medical Center

focus on security and ensure data governance to avoid any type of unfortunate data-related issue."

The importance of unified governance to manage the risks of generative AI was a common theme in our interviews. Schaefer says, at the VA, "we are seeing the need to have very integrated governance models, integrated governance structures for all data and all models. We're putting a strong increased focus on very centralized tools and processes that allow us to have an enterprise data structure." And while unified governance was always a need, generative AI has raised the stakes. "The risk of having non-standardized non-well-defined data running through a model, and how that could lead to bias and to model drift, has made that a much more important aspect," he says.

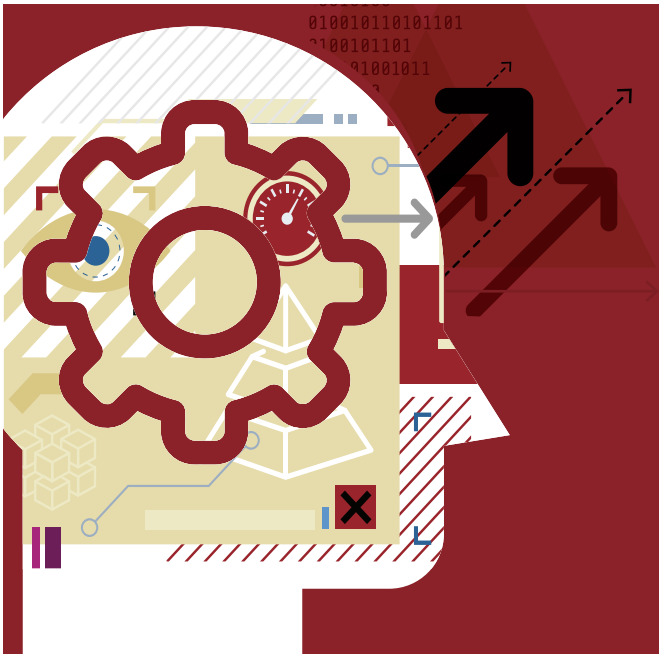
Stoddard notes the importance of including a wide range of voices throughout the AI oversight process at Adobe: "It's important to have the element of diverse oversight through the whole process and to make sure that we have diversity not only of things like ethnicity, gender, and sexual orientation, but also diversity of thought and professional experience mixed into the process and the AI impact assessment." And organization-wide visibility matters, too. Schaefer adds: "High on our list is getting governance tools in place that provide a visual overview of models that are in development, so that they can be spoken to by leadership or reviewed by stakeholders at any time."

Constitutional AI, an approach currently advocated by the startup Anthropic, provides LLMs with specific values and principles to adhere to rather than relying on human feedback to guide content production.[26] A constitutional approach guides a model to enact the norms outlined in the constitution by, for example, avoiding any outputs that are toxic or discriminatory.

And though the risks they bring to the enterprise are substantial, on the flip side, AI technologies also offer great power in reducing some business risks. Zaharia notes, "analyzing the results of models or analyzing feedback from people's comments does become easier with language models. So it's actually a bit easier to detect if your system is doing something wrong and we're seeing a little bit of that." Our 2022 survey found that security and risk management (31%) was the top tangible benefit to AI executives had noted to date, while fraud detection (27%), cyber security (27%), and risk management (26%) were the top three positive impacts anticipated by 2025.

A powerful new technology like generative AI brings with it numerous risks and responsibilities. Our interviews suggest that a motivated AI community of practitioners, startups and companies will increasingly attend to the governance risks of AI, just as they do with environmental sustainability, through a mixture of public interest concern, good governance, and brand protection.

07 Conclusion



AI has been through many springs and winters over the last half century. Until recently, it had made only limited inroads into the enterprise, beyond pilot projects or more advanced functions like IT and finance. That is set to change in the generative AI era, which is sparking the transition to truly enterprise-wide AI.

Few practitioners dismiss generative AI as hype. If anything, its critics are more impressed by its power than anyone. But executives and experts believe that organizations can reap the fruits of the new harvest and also attend to its risks. Commercial organizations and governments alike have to tread a fine line between embracing AI to accelerate innovation and productivity while creating guardrails to mitigate risk and anticipate the inevitable accidents and mishaps ahead.

With these precautions in mind, the most forward-looking CIOs are moving decisively into this AI era. "People who went through the computer and Internet revolution talk about when computers first came online," says Blyton. "If you were one of those people who learned how to work with computers, you had a very good career. This is a similar turning point: as long as you embrace the technology, you will benefit from it."

> "If you were one of those people who learned how to work with computers, you had a very good career. This is a similar turning point: as long as you embrace the technology, you will benefit from it."
>
> Andrew Blyton, Vice President and Chief Information Officer, DuPont Water & Protection

About MIT Technology Review Insights

MIT Technology Review Insights is the custom publishing division of *MIT Technology Review,* the world's longest-running technology magazine, backed by the world's foremost technology institution—producing live events and research on the leading technology and business challenges of the day. Insights conducts qualitative and quantitative research and analysis in the U.S. and abroad and publishes a wide variety of content, including articles, reports, infographics, videos, and podcasts. And through its growing MIT Technology Review Global Insights Panel, Insights has unparalleled access to senior-level executives, innovators, and thought leaders worldwide for surveys and in-depth interviews.

About Databricks

Databricks is the Data and AI company. More than 10,000 organizations worldwide — including Comcast, Condé Nast, and over 50% of the Fortune 500 — rely on the Databricks Lakehouse Platform to unify their data, analytics and AI. Databricks is headquartered in San Francisco, with offices around the globe. Founded by the original creators of Apache Spark™, Delta Lake and MLflow, Databricks is on a mission to help data teams solve the world's toughest problems. In April 2023, Databricks released Dolly, the world's first open source, instruction-following LLM, fine-tuned on a human-generated instruction dataset licensed for research and commercial use. To learn more, follow Databricks on **Twitter**, **LinkedIn**, and **Facebook**.

Endnotes

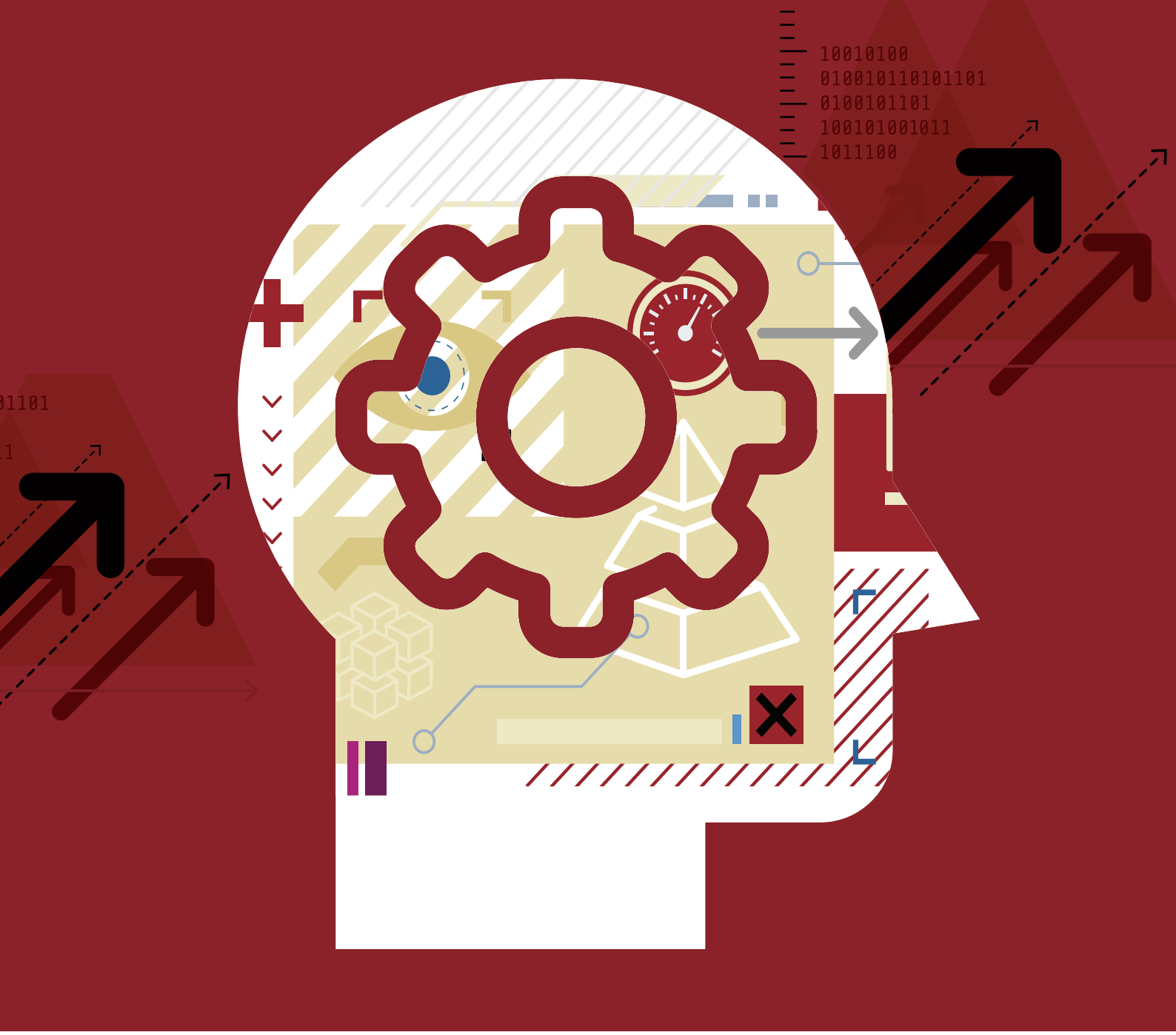
1. "The economic potential of generative AI," McKinsey & Company, June 14, 2023, https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#/.
2. "Generative AI could raise global GDP by 7%," Goldman Sachs, April 5, 2023, https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html.
3. Jim Euchner, "Generative AI," Research-Technology Management, April 20, 2023, https://doi.org/10.1080/08956308.2023.2188861.
4. "CIO vision 2025: Bridging the gap between BI and AI," MIT Technology Review Insights, September 2022, https://www.technologyreview.com/2022/09/20/1059630/cio-vision-2025-bridging-the-gap-between-bi-and-ai/.
5. "Retail in the Age of Generative AI," Databricks, April 13, 2023, https://www.databricks.com/blog/2023/04/13/retail-age-generative-ai.html.
6. "The Great Unlock: Large Language Models in Manufacturing," Databricks, May 30, 2023, https://www.databricks.com/blog/great-unlock-large-language-models-manufacturing.
7. "Generative AI Is Everything Everywhere, All at Once," Databricks, June 7, 2023, https://www.databricks.com/blog/generative-ai-everything-everywhere-all-once.
8. "Large Language Models in Media & Entertainment," Databricks, June 6, 2023, https://www.databricks.com/blog/large-language-models-media-entertainment.
9. Ibid.
10. Ibid.
11. Ibid.
12. Ibid.
13. Jonathan Vanian and Kif Leswing, "ChatGPT and generative AI are booming, but the costs can be extraordinary," CNBC, March 13, 2023, https://www.cnbc.com/2023/03/13/chatgpt-and-generative-ai-are-booming-but-at-a-very-expensive-price.html.
14. Josh Saul and Dina Bass, "Artificial Intelligence Is Booming—So Is Its Carbon Footprint," Bloomberg, March 9, 2023, https://www.bloomberg.com/news/articles/2023-03-09/how-much-energy-do-ai-and-chatgpt-use-no-one-knows-for-sure#xj4y7vzkg.
15. Chris Stokel-Walker, "The Generative AI Race Has a Dirty Secret," Wired, February 10, 2023, https://www.wired.co.uk/article/the-generative-ai-search-race-has-a-dirty-secret.
16. Vanian and Leswing, "ChatGPT and generative AI are booming," https://www.cnbc.com/2023/03/13/chatgpt-and-generative-ai-are-booming-but-at-a-very-expensive-price.html.
17. Arthur Spirling, "Why open-source generative AI models are an ethical way forward for science," Nature, April 18, 2023, https://www.nature.com/articles/d41586-023-01295-4.
18. Sharon Goldman, "With a wave of new LLMs, open-source AI is having a moment—and a red-hot debate," VentureBeat, April 10, 2023, https://venturebeat.com/ai/with-a-wave-of-new-llms-open-source-ai-is-having-a-moment-and-a-red-hot-debate/.
19. Thomas H. Davenport and Nitin Mittal, "How Generative AI Is Changing Creative Work," Harvard Business Review, November 14, 2022, https://hbr.org/2022/11/how-generative-ai-is-changing-creative-work.
20. "Hello Dolly: Democratizing the magic of ChatGPT with open models," Databricks, March 24, 2023, https://www.databricks.com/blog/2023/03/24/hello-dolly-democratizing-magic-chatgpt-open-models.html.
21. "Introducing the World's First Truly Open Instruction-Tuned LLM," Databricks, April 12, 2023, https://www.databricks.com/blog/2023/04/12/dolly-first-open-commercially-viable-instruction-tuned-llm.
22. "Just how good can China get at generative AI?" The Economist, May 9, 2023, https://www.economist.com/business/2023/05/09/just-how-good-can-china-get-at-generative-ai.
23. "Gen AI LLM—A new era of generative AI for everyone," Accenture, April 17, 2023, https://www.accenture.com/content/dam/accenture/final/accenture-com/document/Accenture-A-New-Era-of-Generative-AI-for-Everyone.pdf.
24. "The economic potential of generative AI," McKinsey & Company, June 14, 2023, https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#/.
25. "Generative AI could raise global GDP by 7%," Goldman Sachs, April 5, 2023, https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html.
26. "Claude's Constitution," Anthropic, May 9, 2023, https://www.anthropic.com/index/claudes-constitution.

Illustrations

All illustrations assembled from Adobe Stock and The Noun Project icons.

While every effort has been taken to verify the accuracy of this information, MIT Technology Review Insights cannot accept any responsibility or liability for reliance by any person on this report or any of the information, opinions, or conclusions set out in this report.

© Copyright MIT Technology Review Insights, 2023. All rights reserved.



MIT Technology Review Insights

www.technologyreview.com

insights@technologyreview.com