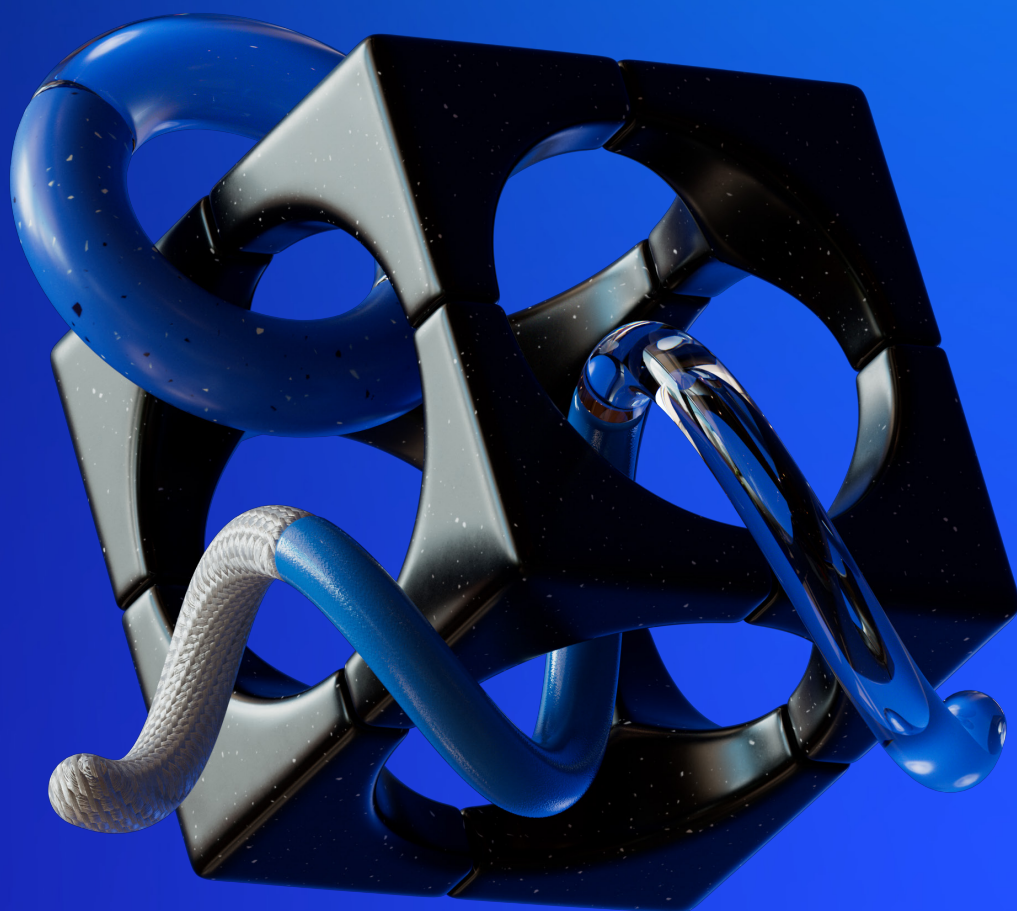# McKinsey & Company

Setting a Vision for Impact in Your AI and Tech Transformation
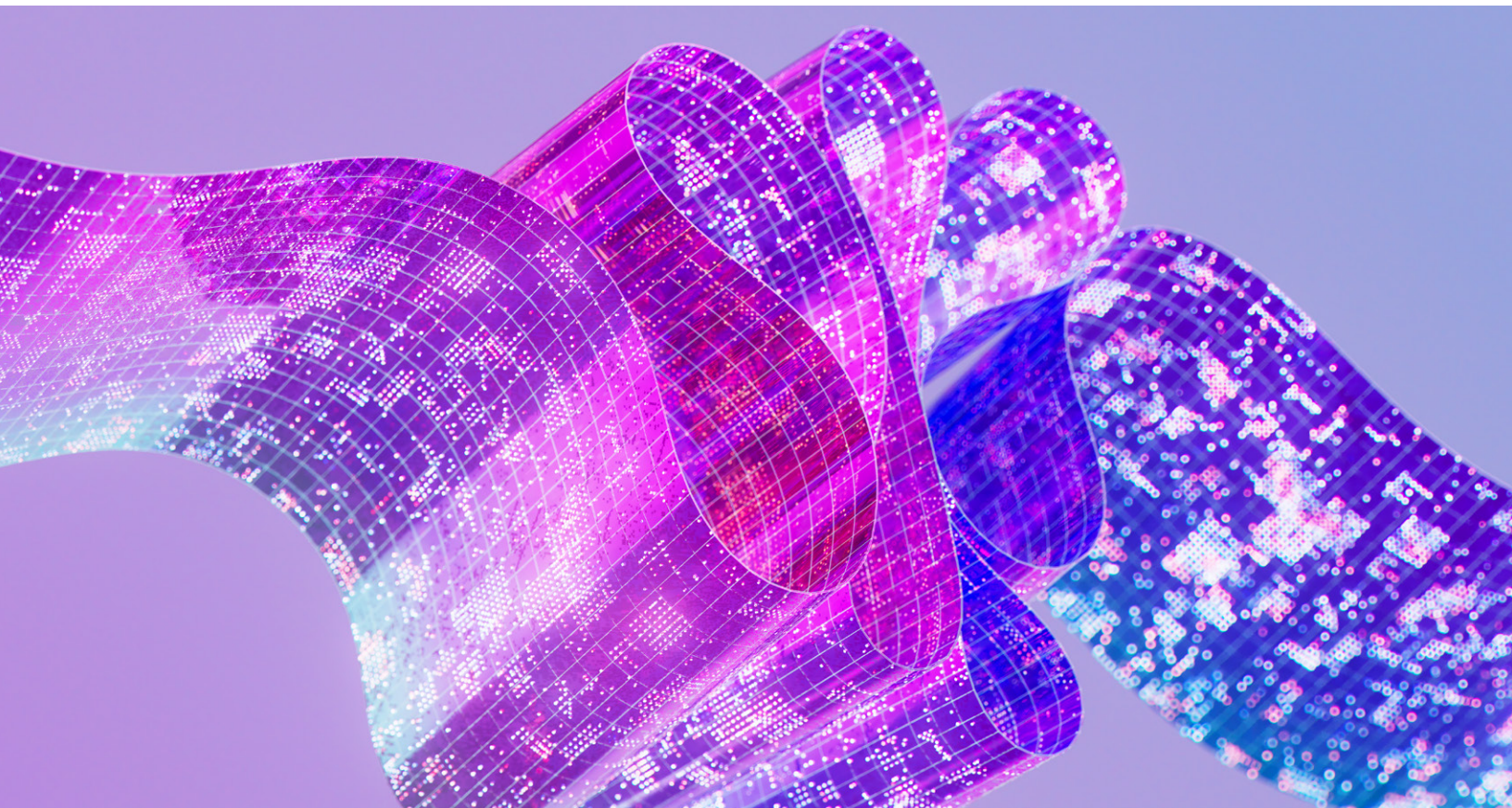
Browse McKinsey's related insights

Operations Practice

Generative AI in operations: Capturing the value

Uncover the transformative power of generative AI within the operations value chain—shaping C-suite discussions across industries.

In this episode of *McKinsey Talks Operations*, host Christian Johnson sits down with senior partner Nicolai Müller and partner Marie El Hoyek from McKinsey's Operations Practice. Together, they discuss the game-changing potential of generative AI. From automating complex processes to unprecedented opportunities across industries, discover insights on productivity boosts, system considerations, and the vital capabilities organizations need for successful integration.

Their conversation has been edited for clarity.

Christian Johnson: Your company's future demands agile, flexible, and resilient operations. I'm your host, Christian Johnson, and you're listening to *McKinsey Talks Operations*, a podcast where the world's C-suite leaders and McKinsey experts cut through the noise and uncover how to create a new operational reality. As we're recording this episode in late 2023, it's clear that generative AI, or gen AI, has become the topic in conversations about digital, analytics, and operations. This new deep learning technology is already making ripples with applications across the value chain.

For today's episode, I'm delighted to be joined by Marie El Hoyek, a partner based in London, and Nicolai Müller, a senior partner based in Cologne. Together, we'll be exploring what generative AI in operations is, how it's different from digital twins and other AI technologies, its potential, and its risks. We'll also look at what it takes to get started with these tools. Nicolai, great to have you here today. Welcome.

Nicolai Müller: Thank you. It's a pleasure to be here, Christian.

Christian Johnson: Marie, so pleased you're able to share your thoughts with us today. Thanks for joining.

Marie El Hoyek: Pleasure being here, Christian.

Christian Johnson: Great. So, Nicolai, can you tell us a bit about why you believe generative AI is worthy of discussion for operations leaders, especially now?

Nicolai Müller: In the past decades, there was this mantra of being faster, being more efficient, and pushing productivity. Tools we all know, such as lean, offshoring, reviewing make-or-buy decisions, and also through technology—but we see nowadays that this productivity improvement gets more complex.

In this scenario, we now have a new technology coming in: generative AI. It promises to automate processes that, in the past, were hard to automate— areas that are more in management collaboration, which currently humans are operating, and also in complex data that you have to manage. So, in this context, there's the question: how much will generative AI help in the search for productivity?

The McKinsey Global Institute has looked into this, and we discovered that, particularly in the areas of collaboration and management, around 50 percent of typical activities can now be automated by generative AI. Also, when it comes to handling complex data and synthesizing the essence of that, we believe there's a huge jump in automation. This may lead to value creation across industries and functions—from pharmaceuticals to automotive, to machinery and functions from engineering, procurement, and supply chain to customer operations—that can unleash tremendous value. We talk about $3.5 [trillion] to $4 trillion, which is approximately the GDP of the UK.

Christian Johnson: Nicolai, what are some of the more specific opportunities that your clients are focusing on, and that you're focusing on right now?

Nicolai Müller: Where I see our clients acting fast is in product development. And if you look deeper into product development, especially in software coding, we see up to a 50 percent productivity increase by having a machine produce code from the simple instruction, "Please give me the code for a program doing XY," and by using tools like ChatGPT and others, a solution is generated. This is one application area where we see generative AI becoming a copilot for humans, aiding in tasks ranging from program management to procurement, and assisting supply chain managers in performing their roles more effectively.

Christian Johnson: Thanks, Nicolai. That has given us a great idea of the why and some of the opportunities. Now, let's go into a little bit more detail about what generative AI is. Marie, what can you describe here for us?

Marie El Hoyek: Generative AI is a fascinating field, and just like the name suggests, it exists at the intersection of artificial intelligence and natural language processing. Essentially, it involves a machine that can analyze something, and this something can now be unstructured, like language or pictures. Similar to a person, generative AI is all about teaching machines to understand and generate text or content.

Now, to add a bit more flavor, let's discuss the different generations of large language models—LLMs. These models are the driving force behind what we refer to as generative AI. One of the first ones we commonly heard about is GPT-3, which stands for generative pretrained transformer 3. When it was introduced, it had 175 billion parameters. Think of parameters as the amount of information it had learned, allowing it to generate text ranging from writing letters to answering questions, primarily text-based. Soon after, GPT-4 was released, and we saw a leap from 175 billion to 170 trillion parameters. Consider how much more it had learned, making it more fluent and accurate, and now it could also be used for images and video.

This is the transformative possibility with generative AI. You can now generate new content in many different types of spaces. Now, that being said, generative AI comes with its own set of risks and challenges. If you imagine that it's based on logic or probabilities, very similar to the human brain, the answers come from what you've learned and your sources. Because of this fact, any generative AI can give you a convincingly wrong answer—and this is what we call hallucination.

Christian Johnson: I love that term. But what do you do about it? How do you mitigate?

Marie El Hoyek: If you had a person answering you based on wrong information, you would tell them, "I want your answer from this specific book." Similarly, you can prompt generative AI better by telling it, "I

want you to answer me from this data set or to tell me where you're guessing."

Another risk is model bias. Imagine that the model or the person has learned from the internet as its source, which is not the most respectful or kindest place. So, whenever you use a model, you need to be able to counter these biases and instruct it not to use inappropriate or flawed sources, or things you don't trust. Another risk that is top of mind is IP [intellectual property] risk. Now, if you imagine generative AI generating code for you, who owns the code? Is it the gen AI that generated it or the requester who wanted it? These details are something we will need to iron out soon.

Christian Johnson: What I'm appreciating here is the discussion of the very limits of the data sources. That's really critical, right?

Marie El Hoyek: It's critical. Additionally, the fact that you need to guide your own data means you have to take care of your data and ensure its safety. Otherwise, that is also an added risk. That being said, all of these risks can be mitigated. However, we need to be aware of them, plan for them, or approach them in a way that limits them so we can control them. By the way, we're witnessing regulations and offerings that are starting to adapt to these risks, and I expect we're going to see quite a few changes in the near future.

Christian Johnson: Just the evolution here—the rapid expansion from 100 billion with a "B" to 170 trillion with a "T" is really dramatic. I think one thing we would now like to turn to is how this is being used and where we are seeing use cases come to life in businesses today. What are some really good examples of that?

Nicolai Müller: I think it's a question that clients have to ask themselves: what impact do I want to achieve? In the end, we have to solve one big question and challenge: how to increase productivity, which involves efficiency and effectiveness.

When we look into use cases, we try to explore different angles. One is the question of automation. Things that currently take hours can be done in

seconds. But it's also about augmentation, where a human may only be able to work with a certain set of data. Imagine being able to access all the data in the world that exist. This was one of the big revolutions; the internet gave us access to all data. Now, with machines, we can use and synthesize that data. So we talk about augmentation. And then we see innovation.

Innovation is the capacity to come up with completely new solutions. Not just making an existing product cheaper or achieving faster product development, but now generating completely new ideas for features and services. So what have we seen? Automation. I talked about how I'm fascinated by what we can now do in software coding and the whole field of engineering. You also heard, for example, the CEO of Nvidia saying, "Hey, the era of software is over. I think we told all our kids to learn software; now you figure out software can be done by a machine." It's a huge evolution that we see, but not only in software.

Parts and hardware development. Synthesizing a huge amount of requirements that your customer gives to you, asking generative AI to understand what the requirements are and how the requirements differ from the last product. How do the requirements vary between products? Are they similar or different? It will help to come to a better synthesis, better understanding of the requirement, and develop faster and better products.

In augmentation in pharma and research, I think we'll see a humongous increase in effectiveness, output, and research. We have cases in pharma where you can imagine understanding each little molecule, what kind of effect it has, and how it reacts with other molecules. It's something that is instrumental. So we see vaccines or other pharma products being developed faster than traditionally was expected by using generative AI. This augmentation leads to a better kind of solution.

As for innovation, you may have also seen one famous German OEM in the US that has integrated ChatGPT into its products. So you can interact and speak with your car. This is innovation. But, Marie, you have also worked with me in this space. What have you seen?

Marie El Hoyek: My background is in industrials, very much deep in operations. Personally, I love all the copilot applications, especially in procurement. The idea that you can ask a friend who knows all your contracts and can answer any question by heart and in plain English is just mind-blowing to me. So, instead of analyzing old contracts, price history, and external trends, I can simply ask the questions. I'm sure there are many more cool applications in terms of content generation, et cetera, but this one, in particular, blew my mind.

Nicolai Müller: And Marie, what I observed are these humongous opportunities out there and the numerous use cases. I mean, we have been in workshops where we were sitting with our clients, and easily after an hour or two, we didn't end up with just five or six potential use cases across a whole different function, but rather 150 or more. I see here a huge opportunity, but the challenge that we're facing is, where do you start? What I call "happy generative AI," where a copilot can help you in your daily job, may become a commodity that everybody can do. Where is the truly transformative generative AI? Is it leading to a differentiating factor for your business? Is it really adding value and creating value for your customers?

I think this is the challenge we face. It's like what we say in Germany, you don't see the woods because of the amount of trees in front of you. So where do you start and where do you end?

Christian Johnson: Can't see the forest for the trees. That's exactly it. When I hear all of this excitement, I also think of the classic chart that we've seen for technologies in general, where you have this initial sharp upward curve as everybody gets very excited about it. Then it sounds like where you're moving is, we need to anticipate when organizations either find, as you've put it, that it's commoditized or that it's hard. And that gets us down then to value. How do companies think about long-term value and not just a set of very exciting use cases that may not build forward very much?

Nicolai Müller: This is a challenging question. If you look into the Google search index, which gives you a bit of a feeling of where we are on the curve, you'll find out that it's now googled more than any

traditional operational questions you have. You have seen all the digital manufacturing terms out there. We have cloud computing and the Internet of Things that we've now seen over the years, and it's a constant discussion.

Generative AI in operations has just started to pick up, I would say, in the first quarter of this year. And it has, in terms of the amount of searches people are doing, overtaken everything you can imagine. This may give you an indication that there is a huge hype out there. But has this hype and all the dreams come true yet? Indeed, people are now starting to recognize that things are easy, like the low-hanging fruits, but actually, the real core is still challenging to implement and also to make your company adaptive to changes. So we are still on the verge of answering one important question when it comes to generative AI: is it now just another tool kit in your operations, like lean or digital or any other artificial intelligence—that is, predictive maintenance—and enables levers you can pull? Or is it a disruption on its own? Is it changing the way you operate? I think these are two scenarios I can imagine.

I tend to believe that in the next two to three years, we'll see these two questions answered. And it may differ completely by player or by industry what the outcome is. Let's talk about disruption. Imagine that coding is now easy. Often, you have, for example, an automotive OEM defining requirements, and then you have a supplier more or less programming the code. If now that code can be programmed by machine, do you need a supplier anymore? It can be disruptive and threatening to say that the raison d'être, or the reason for the supplier to exist, is actually gone. So this is an extreme of a disruption.

For example, for a very research-heavy company, suddenly, if you tap into completely new sources of data, you come to a completely new set of products. And finding the language model that suits you by adopting generative AI in ways that are differentiating may help you to move faster and with better products. I think this is the most pressing question that clients have to answer.

Christian Johnson: I think one of the things we're struggling with and organizations seem to always

struggle with when it comes to a new technology or a new methodology is how do you scale? We talked years ago about pilot purgatory—this idea that you try a bunch of ideas, but then they're never really cohering in a way that creates lasting value. So how can organizations think about this in a way that they can minimize or even avoid that kind of stagnation with this idea?

Marie El Hoyek: This is a good question, Christian. Generative AI might be relatively new, but we have years of experience in scaling digital transformations. To your point, one of the biggest challenges is the pilot trap. Building a pilot or innovating with the technology is great, but transforming an organization is a whole different playing field.

Nicolai talked about the business-led mindset to prioritize applications that are useful with real business ROI. Beyond that, getting a real impact out of any digital change, and for generative AI in particular, will always be both a human and systems question. The way I'd summarize it is, without people, the best technology has no impact. We need to take our people on a real change journey to build the capabilities to use this technology, develop this technology, but also just to know what you can ask of this technology. And by the way, in terms of developing it, there are new skills that are needed here.

Christian Johnson: So what sort of capabilities do organizations really need now?

Marie El Hoyek: I'm thinking about prompt engineering, for example, which is the ability to ask a question really, really well. Now, number two is in terms of systems. There are fundamental questions that businesses should consider early to ensure that whatever they decide leads to capable, consistent, and safe technology usage. You don't want to end up with ten different decisions on the technology because pilots are going left and right.

So you're going to be wondering, do we build our own language models? Do we work with partners? Do we get off-the-shelf solutions? Where do we put our data? How do we process it? These questions

are better learned early, and you need to make a conscious decision about them, to ensure that later on, as you use generative AI more and more, your solution is safe, scalable, and consistent. So, yes, for me, it's both the people element and the systems element that will enable us to go through to the finish line.

Christian Johnson: Excellent. Thank you very much. We're now nearing the end of our discussion. But before you go, I'd like to ask one final question, which is, what should our audience be doing now to bring generative AI to their organizations? There's so much noise out there. We've got a strong idea of that with the Google searches. So how do you start to cut through and make a solid start?

Nicolai Müller: I would recommend two things. First is to start with a pilot, and I would even use the term "play" with generative AI. The cost of doing nothing is just too high because everybody has this at the top of their agenda. I think it's the one topic that every management board has looked into, that every CEO has explored across all regions and industries. So it's important that you start and see what generative AI can do.

In parallel, you need to really think about your strategy. When I talk about strategy, it includes a couple of elements. It's a question of how will this impact my business? Where will it lead to improvements? Where will it not lead to improvement? Should I go fast? Should I not go fast? Do I have solutions out there? Do I need partners? Can I rely on existing LLMs out there, or should I build my own? I think this is the whole question of truly understanding what generative AI in three to five years means for us.

Then there's a layer in the strategy, which is about getting the data technology right. It's understanding how you want to put governance and organization in place, which can build solutions. And there's the question, where do the competencies in my company actually come from? Can I build them? Do I need to acquire them? So you need to be thoughtful about the whole question of competencies needed.

And then there's the question of actually making the change. We often hear that this is the most important thing. You need to make people work with generative AI. You need to capture the early wins, but also things that are more challenging.

Christian Johnson: Excellent. And Marie, anything you'd like to add?

Marie El Hoyek: Yes, Christian. Nicolai, last time we spoke, you talked about this fresh breath of innovation in our companies, and I love to repeat this. You can see it in our discussion even. This gives us the ability to dream again, to come up with new things, and to hope for more impact. And I think, to some extent, we just need to learn, and start doing it, and start capturing it.

Christian Johnson: That's a lovely ending, Marie. Thank you both, Nicolai and Marie, for sharing your expertise and experiences of generative AI with us today. It's a topic that we don't see going away anytime soon. So, your advice on diving in, but with both eyes open to risk mitigation and value creation, is a great note to end on.

Marie El Hoyek is a partner in McKinsey's London office, and **Nicolai Müller** is a senior partner in the Cologne office. **Christian Johnson** is an executive editor and is based in McKinsey's Washington, DC, office.

McKinsey & Company

From AI to Impact: Capabilities powering Lighthouses' 4IR adoption

When technologies reach a certain level of maturity, the next challenge is speed and scale. Factories and entire supply networks—not use cases—become the pilots.

This article is a collaborative effort by Dinu de Kroon, Enno de Boer, Federico Torti, Forest Hou, and Rahul Shahani, with Henry Bristol, representing views from McKinsey's Operations Practice and the World Economic Forum.

This article is one of McKinsey's contributions to the Global Lighthouse Network's latest white paper, published on December 14, 2023. The first article in this series explored the evolution of AI and how leading manufacturers have harnessed it to propel leading-edge manufacturing from the learning phase to the doing phase of the Fourth Industrial Revolution (4IR). The second article examined what artificial intelligence looks like among today's leading manufacturers. This final installment shows how the latest Lighthouses have built the capabilities to deploy AI and other 4IR technologies with both speed and scale.

In the mid-to-late 2010s, the Fourth Industrial Revolution—a group of cutting-edge technologies that could eventually redefine manufacturing—started to emerge. But this industry-shaping impact did not materialize, at least not at first. Even in 2018, when the Global Lighthouse Network was founded, over 70 percent of manufacturers reported being in "pilot purgatory"—their use case pilots failed to have a significant business impact.

Nonetheless, the first Lighthouses began to crack the code, proving the potential impact of data-informed solutions such as automated guided vehicles (AGVs) and digital dashboards. But these early use cases were significantly less advanced than today's—only 10 percent of top use cases in early cohorts relied on AI; by contrast, over 60 percent of today's top use cases do. Moreover, early use cases took significantly longer to implement—averaging ten to 20 months, compared with the under six months that more than three-quarters of today's Lighthouses have achieved.

Efficient implementation at scale is what sets today's industry leaders apart. They no longer focus on piloting individual use cases; instead, they've built the capabilities to get new use cases right quickly, without trials. For the 25 companies with more than one Lighthouse site—and especially the ten advanced scalers, with more than four—entire factories serve as pilots for network scale deployment. Leaders now capture the value of 4IR technologies in ten to 50 factories at a go and have two or three times the ROI of organizations that are still working to find value in a single factory.

Scaling is a big challenge, and even leading companies struggle to achieve it. In a 2022 survey, only 11 percent of companies with a Lighthouse site said they had successfully scaled 4IR technologies across their production networks. The most critical bottlenecks they cited were fragmented data landscapes, legacy IT infrastructures, and a shortage of in-house talent.

As our colleagues Eric Lamarre, Kate Smaje, and Rodney Zemmel wrote in their recent book *Rewired: The McKinsey guide to outcompeting in the age of digital and AI*, achieving scale means "getting thousands of people across different units of the organization working together and working differently. It means bringing on new talent and developing accelerated learning loops that harness their skills and help them grow." This is no small task. Yet Lighthouses, accelerating at just the right pace, have found that sweet spot. Some, in fact, have transformed sites in a mere matter of months. Let's see how.

Lessons from leaders: Six steps to powering scale

Lighthouses, like digital leaders in other industries, have driven themselves up the adoption curve with a six-part approach to site transformation. (1) They first set the strategic road map. That's like setting the GPS to guide the transformation toward an organization reimagined with technology and determining the route to prioritize and sequence the value at stake—including use cases (at the site level) and factories (at the network level).

These companies then build their delivery capabilities: the engine that powers the transformation. (2) The engine parts—its pistons, crankshafts, drivetrain, and timing belt—are programs to hire, train, and retain digital talent. (3) An agile operating model (often including digital studios) fosters speed, quality, and collaboration. (4) A technology backbone affords a clear, scalable, and distributed archi-

Efficient implementation at scale is what sets today's industry leaders apart. They've built the capabilities to get new use cases right quickly, without trials.

tecture for providing digital services and solutions easily. (5) The data architecture and governance enable critical decisions and ensure quality, easy consumption, and reuse.

But capabilities alone are not enough. (6) The final enabler—change management for adoption and scaling—is like the driver's hands on the steering wheel. That is what maximizes value by ensuring the adoption and scaling of digital and analytics solutions. It involves building new skills and leadership characteristics and tightly manages the transformation's progress and risks. And, of course, it validates the program's impact by tracking key metrics in a standardized format. This enabler often includes a transformation office, which nearly 70 percent of Lighthouses cite as the most critical of the six to get right.

No less critical to the transformation's functioning is the motor oil: ecosystem collaborations that keep the engine humming as universities, technology providers, innovation incubators, public entities, and many others shape best-in-class capabilities. Each Lighthouse in this recent cohort awarded in December 2023 has listed an ecosystem collaboration as a key enabler of the 4IR journey.

In all, the capabilities that pilot a factory transformation are the same ones that Lighthouses use to unlock network-level impacts at speed and scale. Johnson & Johnson's latest Lighthouse (in Thailand), for example, implemented an energy management use case that was not only piloted but also scaled to 16 other sites within a year.

Setting your GPS with a strategic road map

The first enabler, the strategic road map, is like the GPS navigation system that plans a journey. Aligning senior leaders on the transformation's vision and value, it involves a shared reimagining of business domains to deliver exceptional customer experiences and create competitive distance. We've observed that leading companies tend to fit one of three strategic archetypes as they plot their scaling course: IT-led, center of excellence (COE)—led, or "build and replicate" (Exhibit 1). Choosing the right scaling strategy depends on three factors: the diversity of your production processes, the staffing models of your sites, and the maturity of your enterprise information technology and operational-technology (IT/OT) stack.

For companies, such as Tata Steel, that have only a few but very large sites, it may make sense to swarm one factory at a time. That's exactly what Tata did, first creating a Lighthouse factory in the Netherlands and then replicating the changes, one at a time, at two factories in India. Now, 80 percent of Tata Steel's global production comes from one of the company's three Lighthouse sites.

If a company has a large footprint, with diverse production processes and very site- or business unit—oriented staffing models, a capability-led approach to scale may be superior. Siemens, with hundreds of sites spanning dozens of manufacturing processes, tools, and systems, took this approach. Sites in Germany (Amberg) and in China (Chengdu) were chosen as digital-innovation hubs for

Exhibit 1

There are three strategic archetypes for how companies scale digital use cases across their production networks.

Archetypes of digital use cases

1 Technology-led

Use cases deployed at sites at once: *Resourcing primarily from IT/corporate (CATL)*

- Deploys single use cases across hundreds of production lines simultaneously (eg, deployment of new AI models for asset-level energy optimization)

- Rollout driven by central technical organization (eg, intelligent manufacturing department)

- Approach enabled by strong central IT network and similarity of production lines across dozens of factories

2 Capability-led

One lead site that 'trickles out' use cases: *Resourcing primarily from innovation site (eg, Siemens)*

- Locates talent at single site designated as a center of excellence (eg, a blockchain-based software to see supplier emissions and products' carbon footprint was built first in Amberg, Germany)

- Allows concentration of resources to accelerate development, after which technologies are scaled to other sites

- Approach enabled by attractiveness of Amberg site (eg, co-located with Siemens' Impulse Visitor Center)

- A reference design enables site-specific customizations when scaling to meet needs of diverse IT systems across sites

3 Site-led

One site fully transformed at a time: *Resourcing primarily from the "active" site (Tata Steel)*

- Netherlands location designated as a champion site, where a complete digital transformation was executed to resolve implementation challenges before replicating 1 at a time to 2 more plants in India; all 3 are now lighthouses

- This approach was chosen because each plant is large and distinct—today, nearly 80% of total steel production comes from 1 of their 3 lighthouse sites

- Other reasons to select this approach include when the optimal approach is not initially clear or when time isn't a critical factor

McKinsey & Company

---

disseminating use cases to other factories. COE teams across the network include leaders from multiple sites and are designed around specific technological capabilities, including a next-gen manufacturing execution system (MES) and digital twins.

Finally, companies with highly replicable processes across their production base and a strong, universal IT backbone may be best off scaling one use case or technology at a time. Contemporary Amperex Technology Co. Limited (CATL) is a prime example. Its dozens of factories have hundreds of battery production lines that all look and feel similar. The company's IT and Ops departments, not its sites, host hundreds of data scientists, AI engineers, and

other kinds of tech talent. CATL's MES, enterprise resource planning (ERP), and other systems are fairly standardized, and it has invested in developing its own universal data profiles and name spaces to help standardize and accelerate rollouts. As a result, a new use case (such as a specific algorithm for set point optimization) can reach hundreds of production lines in just weeks—and sometimes less.

Building your 4IR capability engine

Today's Lighthouses teach us that the principles explained in *Rewired* work for manufacturers. Much hinges on the capability engine, guided by an effective strategy: enablers two to five. Let's open up the hood and check things out.

Digital talent

Leading organizations know that talent is just as vital as technology. Moreover, they know that each company tends to have unique talent needs: technology can typically be replicated once it becomes established, but skills and knowledge must be tailored. Learning must therefore be not only customized but also continually renewed and maintained. Front-runners know there's never a time to rest. Companies can't take the talent pipeline or their understanding of their skills and talent gaps at any given moment for granted.

LONGi, a Chinese solar-technology manufacturer, exemplifies a talent strategy focused on reskilling the existing workforce. To tailor the training of almost 1,000 employees, the company implemented an evaluation-training-certification method involving the identification of skill gaps, personalized learning, and value-based certification. A diagnostic informed a comprehensive talent road map. The 3F model—forums (learning), field (practice), and feedback (evaluation)—grounded all learning, especially for new digital roles, such as agile coaches and data scientists. A closed-loop, impact-driven certification process mapped the impact of the work of employees while recognizing and supplementing their capabilities (Exhibit 2).

Agile operating model

An effective capability engine must realize a challenging goal: bringing business and technology together to throttle up. This is a leadership challenge, and success depends on alignment across functions. An appropriate allocation of resources, clear incentives, and intelligently built

Exhibit 2

Successfully reskilling an existing workforce requires a multilevel strategy.

Reskilling in a 3F (fieldwork, feedback, forum) structure

- Impact — New technologies, behavioral science advances
- Certifications — Value-based certifications
- **3** Forums to learn — Workshops, Toolboxes — Mobile learning and microlearning
- **2** Feedback to grow — Coaching, Boards, Groups — Measurement and behavioral nudges
- **1** Fieldwork to apply — Reflection, Growth planning, Learning — Personalization

Diagnose → Intervention execution

Impact

>30	workers certified in lean and digital proficiencies	>980	workers trained	80–150%	improvement on assessments	>1,500	workers evaluated

McKinsey & Company

teams are essential components of the agile operating model, which couples people resources with tech and data resources, setting up regular and consistent forward momentum (Exhibit 3).

ReNew is a prime example: its ReNew Digital (ReD) team of nearly 30 engineers and data scientists runs five or six agile pods, each focused on discrete use cases, located in its innovation studio. The team is responsible for leadership buy-in, including active sponsorship by the chief d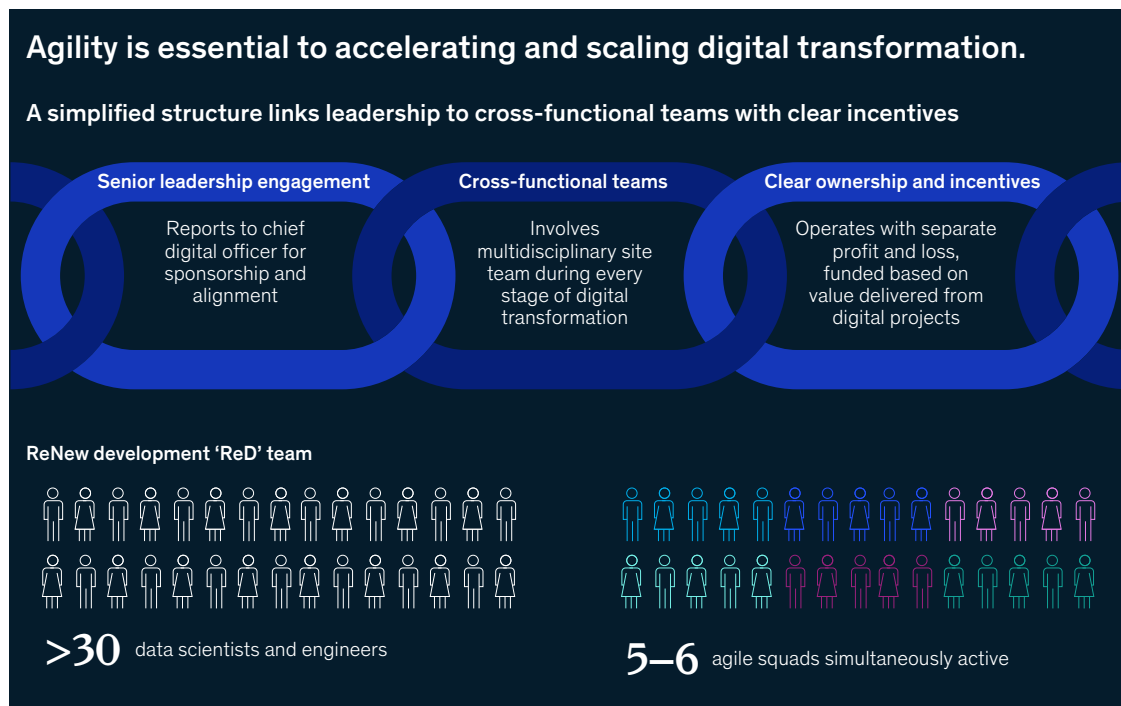igital officer (CDO); aligns incentives with an independent profit and loss, funded in part by the impact of new use cases; and uses a multistakeholder staffing model that includes end users, internal business owners, data scientists, engineers, and process owners from all relevant departments.

Technology backbone

To achieve an at-scale transformation and unlock business value, a strong technology backbone is vital. It's important to understand that this means more than just one or even a handful of substantial new technology investments. Those tend to be localized, which can promote fragmentation in the long run—creating, ironically, a force antithetical to transformation at scale. Rather, the technology backbone must be just that: the infrastructure needed to convey signals across the organization for at-scale deployment. It's therefore necessary to think about at-scale deployment from day one, so leading organizations prioritize the accessibility and adaptability of their data environments by using decoupled architecture solutions (such as microservices), along with advanced development environments and tools.

Exhibit 3

Agility is essential to accelerating and scaling digital transformation.

A simplified structure links leadership to cross-functional teams with clear incentives

Senior leadership engagement	Cross-functional teams	Clear ownership and incentives
Reports to chief digital officer for sponsorship and alignment	Involves multidisciplinary site team during every stage of digital transformation	Operates with separate profit and loss, funded based on value delivered from digital projects

ReNew development 'ReD' team

>30 data scientists and engineers

5—6 agile squads simultaneously active

Use cases from ReNew Ratlam (India) cluster, deployed rapidly and with scale

Impact

80—90% decrease in unplanned maintenance

60—70% decrease in major field failure

Scale

>70 wind farms
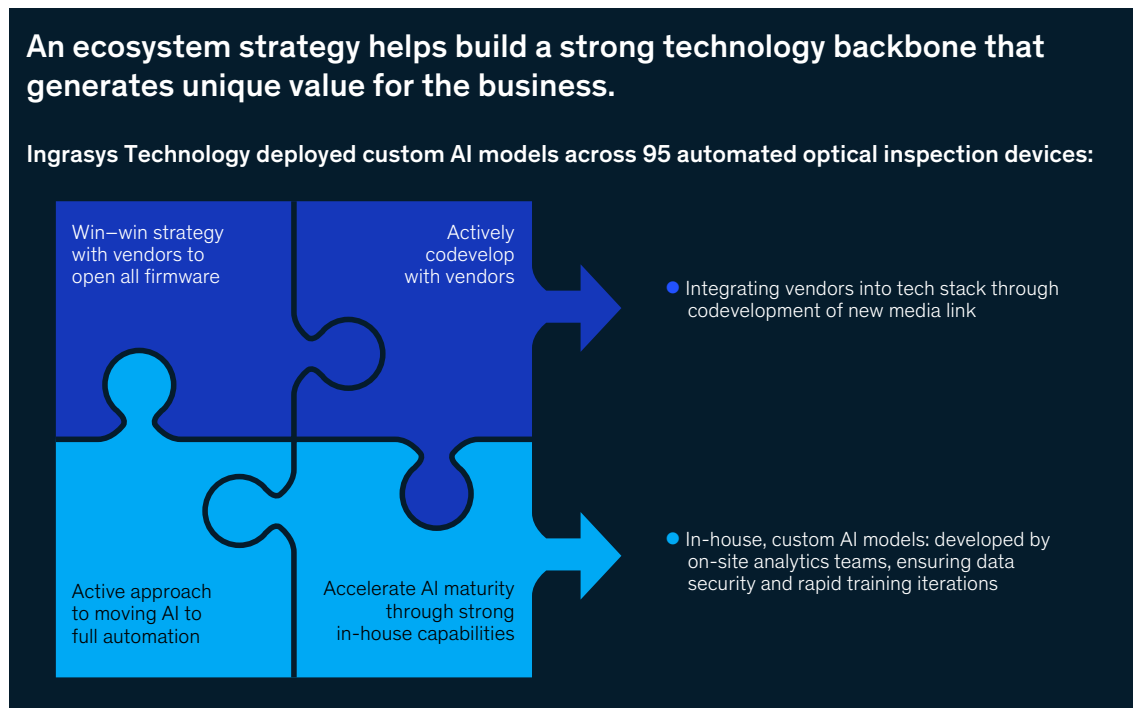
>1,500 turbines

McKinsey & Company

Ingrasys demonstrates the benefits of starting small, but with a clear growth plan that intelligently integrates an ecosystem approach (Exhibit 4). Manufacturing servers require a significant number of black box vendor technologies, and Ingrasys addresses this need with clear design principles and vendor requirements that provide for security and scale. As a result, most (if not all) of the company's vendors must open up their firmware to make data and controls accessible, design data outputs to mesh with the Ingrasys technology infrastructure, and collaborate for customized development and win–win solutions. In one case, an automated optical inspection (AOI) vendor collaborated on a new physical IT/OT plug-in or "media link" on the machine, so Ingrasys could deploy an in-house–developed AI inspection model that augmented the preexisting vendor solution. The vendor now offers this "bring your own AI" as a service solution to customers.

Data architecture
A solid data architecture works hand in hand with the technology backbone to empower talented people and help them work together in an agile way. The most advanced AI and analytics technologies emerging today thrive on massive, ever-expanding data sets. Front-runner organizations must therefore develop clear reference architectures and robust data pipelines to enable both business intelligence and machine learning solutions. Furthermore, they must have automated tools that actively support data quality and maintenance routines.

Exhibit 4

An ecosystem strategy helps build a strong technology backbone that generates unique value for the business.

Ingrasys Technology deployed custom AI models across 95 automated optical inspection devices:

Win–win strategy with vendors to open all firmware

Actively codevelop with vendors

Active approach to moving AI to full automation

Accelerate AI maturity through strong in-house capabilities

- Integrating vendors into tech stack through codevelopment of new media link

- In-house, custom AI models: developed by on-site analytics teams, ensuring data security and rapid training iterations

Impact

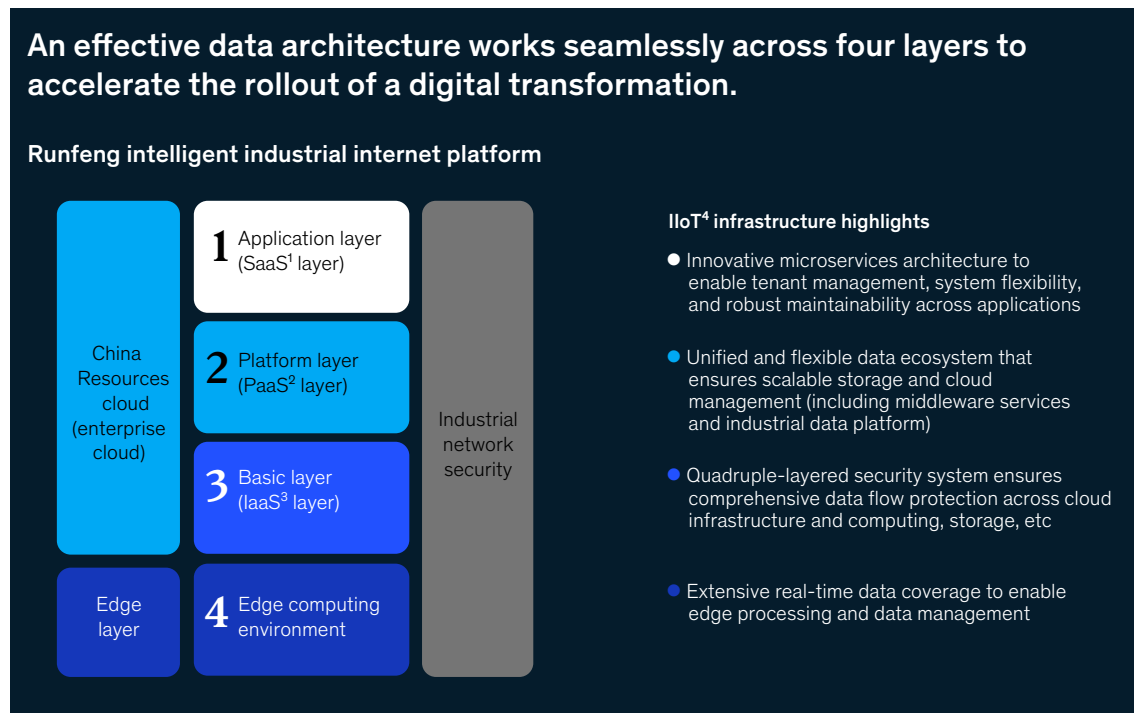| 90–100% | SMT[1] defective rate decrease | 80–150% | Root cause analysis time reduction | 40–50% | Labor productivity in quality testing |

[1]Surface-mount technology.

McKinsey & Company

At CR Building Materials Tech, the data and analytics across 35 sites are powered by the Runfeng intelligent industrial internet platform, which sped up the deployment of digital use cases by 50 percent. The platform has four service layers, each making data accessible, accurate, and efficient. Within the application layer, an innovative microservices architecture improves tenant management, bolsters the system's flexibility, and ensures maintainability. To ensure scalable storage and cloud management, the platform layer merges flexible configuration, enhanced deployment, and AI-driven analytics with a unified data ecosystem. The edge computing environment, which excels at managing industrial operations in real time, runs applications deployable across all 35 Runfeng sites (Exhibit 5).

Scale up with effective change management

Capabilities alone can't drive a transformation: effective change management is needed to steer around potholes and roadblocks. Often, Lighthouses use transformation offices for this purpose; in fact, nearly 70 percent cite them as the most critical enabler for transformations. Their role is to track the impact of transformation programs, to provide appropriate financial incentives and proactive risk management, and to build critical digital skills and a strong culture of change across the organization. Many also support the active "assetization" of solutions for easy replication.

Exhibit 5

An effective data architecture works seamlessly across four layers to accelerate the rollout of a digital transformation.

Runfeng intelligent industrial internet platform

China Resources cloud (enterprise cloud)

1 Application layer (SaaS[1] layer)

2 Platform layer (PaaS[2] layer)

3 Basic layer (IaaS[3] layer)

Edge layer

4 Edge computing environment

Industrial network security

IIoT[4] infrastructure highlights

- Innovative microservices architecture to enable tenant management, system flexibility, and robust maintainability across applications

- Unified and flexible data ecosystem that ensures scalable storage and cloud management (including middleware services and industrial data platform)

- Quadruple-layered security system ensures comprehensive data flow protection across cloud infrastructure and computing, storage, etc

- Extensive real-time data coverage to enable edge processing and data management

Impact

35 sites and multiple business units with standardized, accessible, secure data

24/7 continuous highly reliable operating environment[5]

50% Labor productivity in quality testing

[1]Software as a service. [2]Platform as a service. [3]Infrastructure as a service. [4]Industrial Internet of Things. [5]For critical functions (eg, energy management, predictive-maintenance planning).

McKinsey & Company

A networkwide outlook

Lighthouses have shown that to succeed with transformation efforts, companies must adopt a networkwide outlook right from the outset. Thinking too small won't cut it. A pilot designed for site-level implementation might yield some benefits. But for a transformation to have a true business impact, companies must think, from day one, about at-scale deployment: building a strategy around a clear destination and then implementing a talent development approach that combines upskilling and new hiring to provide the capabilities needed. Coca-Cola, for example, used go-and-see programs, boot camp–style sessions, and training academies in its highly effective change management plan.

The collaborative ecosystem

The six enablers—including the four of the capability engine—work best when allowed to flourish in a collaborative environment. Innovation can and does happen in-house, but like a superior vehicle design, the best-performing capabilities harness the benefits of strategic partnerships. For industrial front-runners, this means building meaningful connections with universities, technology partners, and innovation centers. These not only are important sources of talent but also enable front-runners to share the knowledge and innovations that drive entire industries up the 4IR adoption curve.

——————

This article series has focused on what the most advanced industrial manufacturers have done to take the world to the current inflection point of the Fourth Industrial Revolution—and where they are poised to take it next. By implementing the bold steps needed to innovate and climb the learning curve, Lighthouse organizations have shown what's possible as they move beyond using 4IR innovations for site-level transformations and progress toward a network-level impact. Through a fervent commitment to capability building and a disciplined embrace of emerging machine intelligence technologies, Lighthouses continue to power up the doing part of the 4IR adoption curve, leaving the learning stage far behind. Next, they are poised to enter the optimizing stage and usher in radical levels of industry transformation at scale, effectively rewiring manufacturing across the value chain.

Dinu de Kroon is a partner in McKinsey's Zurich office, **Enno de Boer** is a senior partner in the New Jersey office, **Forest Hou** is a partner in the Shanghai office, and **Rahul Shahani** is a partner in the New York office, where **Henry Bristol** is a consultant. **Federico Torti** is the World Economic Forum's lead for its advanced manufacturing and supply chain initiatives.

McKinsey & Company | QuantumBlack
AI by McKinsey

From promising to productive: Real results from gen AI in services

Service organizations that are early generative AI adopters are finding that to capture more value, they need to get more disciplined.

by Jorge Amar and Oana Cheta
with Ivan Huang and Stephen Xu

August 2024

Generative AI (gen AI) could provide the productivity boost operations leaders have hoped for, as well as a means to fight cost pressures—if only leaders could get going. McKinsey's latest tech trends research finds that only 11 percent of companies worldwide are using gen AI at scale.[1]

Operations is a major gap: in a February 2024 survey of 150 executives at large North American and European companies, only 3 percent of respondents said their organization has scaled a gen AI use case in an operations-related domain. A separate survey, conducted in April 2024 of more than 250 corporate-function leaders worldwide, found that service operations is faring only slightly better. In finance functions, for example, about

45 percent of organizations are now piloting gen AI solutions, compared with 11 percent in 2023—but only 6 percent have achieved scale.
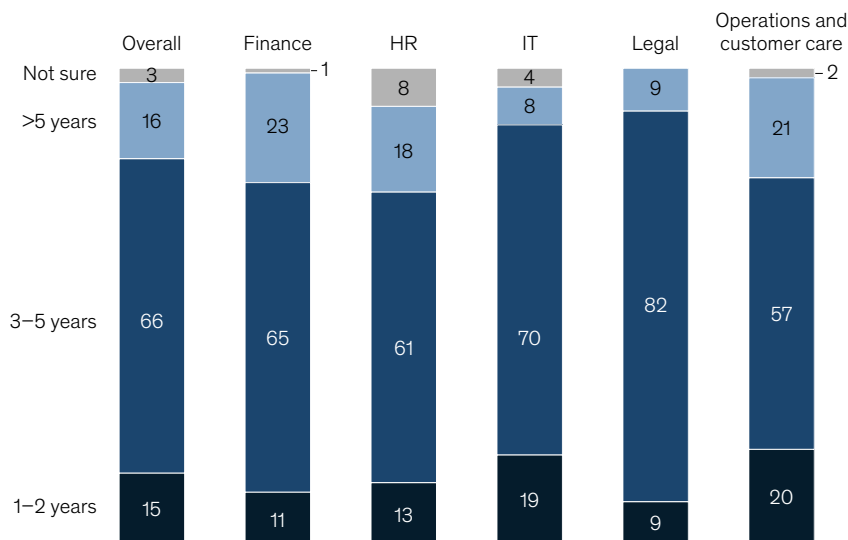
The results reflect uncertainty among operations leaders about which of the many use cases they have deployed will yield real competitive advantage. Executives understand that realizing full value from their gen AI investments won't be instantaneous: two-thirds of the April survey respondents set a three- to five-year timeline (Exhibit 1).

But many also said that they wanted to be more confident that their commitments would pay off. One CEO recently told us, "We've already spent about $100 million funding hundreds of gen AI

Exhibit 1

Most executives expect that it will take three to five years to capture the full value from their generative AI investments.

Expected time to realize value of current generative AI plans, by function, % of respondents

	Overall	Finance	HR	IT	Legal	Operations and customer care
Not sure	3	1	8	4		2
>5 years	16	23	18	8	9	21
3–5 years	66	65	61	70	82	57
1–2 years	15	11	13	19	9	20

Note: Figures may not sum to 100%, because of rounding.
Source: 2024 McKinsey Corporate and Business Functions CXO Survey, conducted Apr 10–May 30, 2024, n = 276

McKinsey & Company

---

[1] "Moving past gen AI's honeymoon phase: Seven hard truths for CIOs to get from pilot to scale," McKinsey, May 13, 2024.

experiments; harvesting at least some of the value will help us see where additional investment will be worthwhile." Companies also cited unclear road maps, talent shortages, and immature governance as further impediments to scale.

A few companies, however, are capturing real value already, attributing more than 10 percent of their EBIT to their use of gen AI.[2] Early successes like these reveal three critical tasks in setting gen AI up for scaling across an organization. The first is to design a cohesive, disciplined operational strategy for deploying gen AI. That means prioritizing use cases for long-term value by focusing on their potential to not only transform specific process points or domains but also reimagine complete workflows.

Second, to sustain broad impact over time, companies will need to focus on the enablers supporting the humans who make gen AI work—providing the necessary governance and performance infrastructure while also investing in change management and a continuous innovation culture.

The third task is the culmination of the first two: thoughtfully integrating gen AI tools with human capabilities to create the most advanced solutions, such as autonomous gen AI agents or copilots. The most successful can tackle every step of a complex workflow. At one bank, for instance, a gen AI agent now drafts credit-risk memos, increasing revenue per relationship manager by 20 percent. And a copilot in the finance department of a consumer goods maker is reducing operating expenses relating to financial planning and analysis by between $6 million and $10 million.

Deploying operational gen AI strategically

As with earlier waves of technological change, gen AI raises the specter of pilot purgatory, in which dozens of experiments fail to amount to sustained impact. Organizations that have already built up their capabilities in deploying gen AI tend to see better returns on their gen AI investments over both the short and longer term. They especially excel at thinking through sequencing, with a focus on

Organizations that have already built up their capabilities in deploying gen AI tend to see better returns on their gen AI investments over both the short and longer term.

[2] "The state of AI in early 2024: Gen AI adoption spikes and starts to generate value," McKinsey, May 30, 2024. Out of 876 survey respondents who estimated the proportion of their organization's EBIT that was attributable to gen AI, 46 gave a figure of more than 10 percent.

scalability and reusability so that they can reimagine entire chains of value creation.

Building this sort of maturity in gen AI transformation is now essential, leaving companies little time to waste. Ideally, lessons from lower-risk, earlier applications of gen AI build critical capabilities that help higher-risk (and higher-reward) later applications succeed.

Prioritizing use cases

The experience of a global bank illustrates the benefits of deploying gen AI strategically. First, based on a detailed assessment of business impact and technical feasibility, it winnowed 23 potential domains for use of gen AI to just two: the contact center in its consumer banking unit and the know-your-customer (KYC) function for corporate
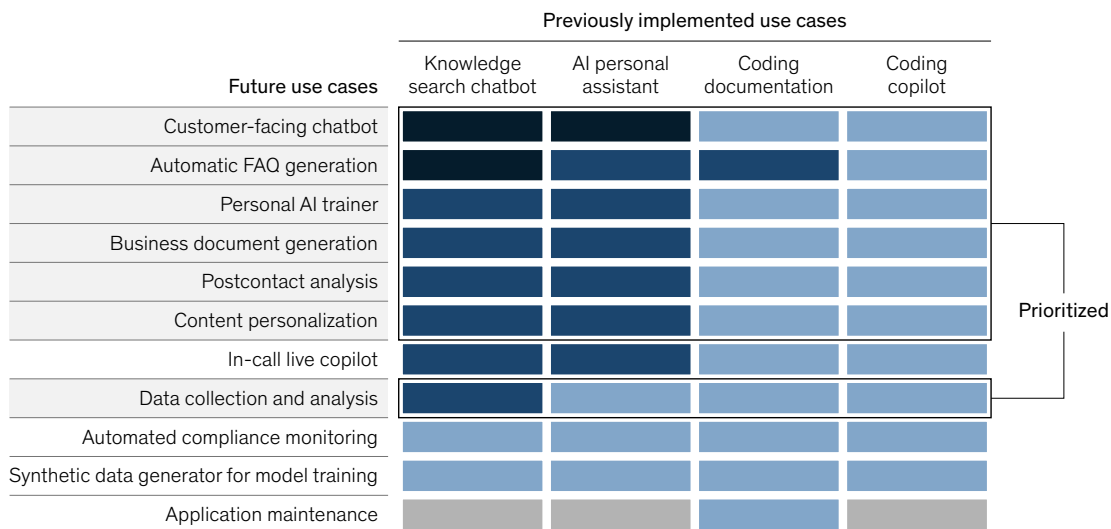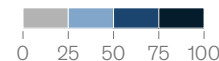
and investment banking. Despite the apparent differences, the two domains not only showed high potential for gen AI impact but also shared a few commonalities, particularly for gen-AI-based knowledge extraction and synthesis. The same technologies could support customers looking for information and employees looking for internal documents, so that the underlying technology could be reused and scaled more effectively (Exhibit 2).

To determine which of the two finalists would go first, the company applied an additional screen: risk. The confidential nature of the KYC function's data made it a higher-risk target, so the bank started instead with the contact center. The final strategic decision concerned which use cases to deploy within customer care. Keeping "ability to scale" and "reusability" top of mind, the chatbot came out on

Exhibit 2

Companies can prioritize generative AI use cases based on potential synergies between underlying modules.

Overlap in module usage between future and previously implemented use cases, illustrative, %

	Previously implemented use cases			
Future use cases	Knowledge search chatbot	AI personal assistant	Coding documentation	Coding copilot
Customer-facing chatbot	██	██	██	██
Automatic FAQ generation	██	██	██	██
Personal AI trainer	██	██	██	██
Business document generation	██	██	██	██
Postcontact analysis	██	██	██	██
Content personalization	██	██	██	██
In-call live copilot	██	██	██	██
Data collection and analysis	██	██	██	██
Automated compliance monitoring	██	██	██	██
Synthetic data generator for model training	██	██	██	██
Application maintenance	██	██	██	██

Legend: 0 25 50 75 100

Prioritized

Source: Expert interviews

McKinsey & Company

The core question is: "How could gen AI help me rethink my operations?" Answering it means reexamining each process as part of a larger workflow.

top: it's comparatively easy to implement, generates measurable outcomes, and helps build a foundation for similar use cases that extract and synthesize complex data.

Within just a few weeks, the center's fully designed use cases included a customer-facing chatbot. In just seven weeks of use, the new chatbot offered an improved customer experience, eliminating wait times for about 20 percent of contact center requests.

Moreover, lessons from the contact center have formed a reusable foundation that the bank can adapt for the KYC function. Chatbots are now a component of a "smart, virtual agent" that guides relationship managers through a far more automated KYC process. The virtual agent can prepopulate client information into forms, determine which documents are required, validate data uploads, and follow up on any missing information.

From point solutions to complete workflows
As the bank example illustrates, however, the core operational question for generative AI isn't "How could gen AI help me improve my current processes?" To "improve" a process often means addressing only a symptom rather than the underlying condition—for example, using gen AI to automate note-taking and action-item generation for meetings without considering why there were so many meetings in the first place.

The core question is therefore much broader: "How could gen AI help me rethink my operations?" Answering it means reexamining each process as part of a larger workflow—and, in many cases, as part of a user or customer journey.

Breaking barriers to better service. To illustrate the difference, consider the case of a leading North American telecommunications provider, whose use case prioritization exercise led it to focus on customer care. Rather than start by exploring how gen AI tools could improve particular process steps in care, the company stepped back, asking instead how gen AI could combine with traditional process improvement techniques and new talent to raise productivity within the customer care function overall.

That shift in perspective led the company to reevaluate its customer journeys, starting with a traditional mapping of every touchpoint, from initial contact to final resolution. With the resulting flowcharts in hand, company leaders questioned each process step to see if it was overengineered or unnecessary. It considered the step's effect on customer experience (such as increasing complexity or wait times) versus the potential risks from its elimination (such as increased fraud or security lapses).

For example, after the company mapped out the journey of changing a phone number, one particular

step surfaced as so complex and painful that the company gave customers the option to delegate it to staff in exchange for a fee. But because customers were reluctant to pay, staff would often guide them through the step—a costly alternative for the contact center. Once the company understood the reasons customers got stuck, it could design a self-service solution. In combination with other technologies, gen AI's capabilities to provide detailed, automated guidance meant that the company could reduce average call length (and cost) while eliminating the fee entirely, improving customer experience (Exhibit 3).

Deeper analysis into the root causes of customer pain points also revealed internal misalignments that the company needed to address before gen AI

could provide a solution—such as when price changes set by the marketing team led to surges in customer calls that the care team couldn't handle. Unaware of the changes, agents would transfer customers to other departments, often in multiple loops, leading the care team to provide deep discounts in hopes of retaining the frustrated customers.

Accordingly, the company revamped its cross-functional workflows so that the customer care team could work with marketing to anticipate potential customer concerns and develop appropriate responses in advance. Leaders also reexamined the skills that service teams would need, developing new talent profiles (and associated capability building modules) that
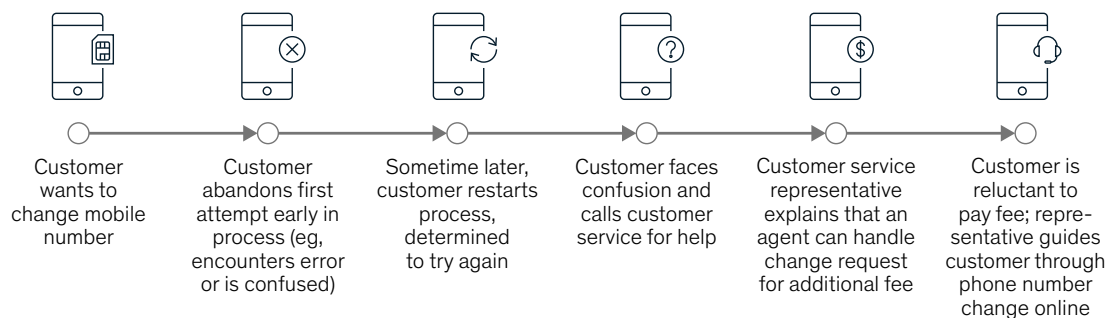
Exhibit 3

Generative-AI-based assistants can proactively message and help customers who get stuck in administrative tasks.
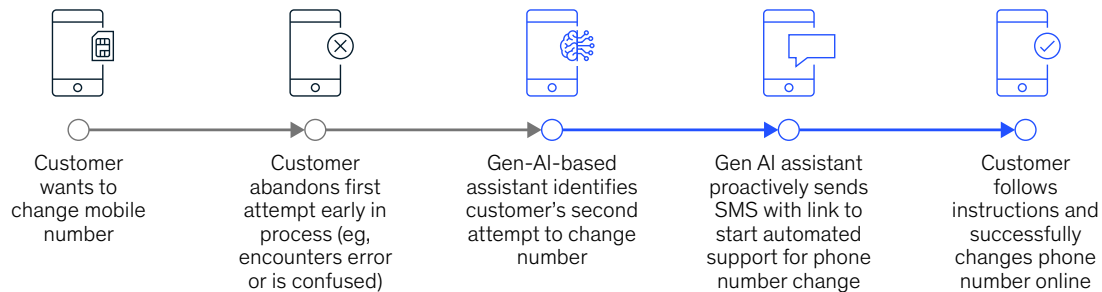
Customer journey for changing mobile number, illustrative   ○ Generative AI (gen AI) solution

Current process

| Customer wants to change mobile number | Customer abandons first attempt early in process (eg, encounters error or is confused) | Sometime later, customer restarts process, determined to try again | Customer faces confusion and calls customer service for help | Customer service representative explains that an agent can handle change request for additional fee | Customer is reluctant to pay fee; representative guides customer through phone number change online |

Improved process with gen AI

| Customer wants to change mobile number | Customer abandons first attempt early in process (eg, encounters error or is confused) | Gen-AI-based assistant identifies customer's second attempt to change number | Gen AI assistant proactively sends SMS with link to start automated support for phone number change | Customer follows instructions and successfully changes phone number online |

McKinsey & Company

could evolve with the workflows. Changing the internal collaboration model set the foundation for a later, gen-AI-based self-service option, while analytic AI tools could optimize staff allocation to provide additional call center coverage.

Freeing up employees' capacity. Employee journeys were the final piece of the puzzle. The company analyzed every step of the agent experience, from logging in to resolving customer inquiries and completing tasks. This analysis involved streamlining processes and reducing the complexity of technology systems that agents had to interact with. The telecommunications provider also identified potential misalignments between agent incentives and customer needs, ensuring that agents were given incentives to prioritize customer satisfaction and resolution rather than simply handle a high volume of calls.

By taking an integrated approach to workflow optimization as a part of critical journeys, the telecommunications provider achieved significant and lasting improvement in its customer care function, with gen AI building on a range of analog and tech-based improvements. Total call volume fell by about 30 percent, and average handle time by more than one-quarter, even as service quality improved: first-call resolution rates rose by ten to 20 percentage points.

The (human) secret to scale

As with previous technologies, gen AI's full potential depends on its reaching scale throughout an organization. Few companies have reached this point. Their experience underscores the importance of four elements, all of which center on humans rather than technology. The first two elements provide critical guidance; the second two more directly change the way people work, with a particular focus on change management.

Governance
Successful deployment of gen AI can't be ad hoc. This is due to not only gen AI's well-publicized risks—from inaccurate training data for gen-AI-based tools to "hallucinations" that produce incorrect results—but also the tendency of the most advanced organizations (the ones generating more than 10 percent of EBIT from gen AI) to centralize their gen AI initiatives. Almost half of these high performers report centralizing compared with only 35 percent of other companies.

As with previous technologies, gen AI's full potential depends on its reaching scale throughout an organization. Few companies have reached this point.

The components of the governance structure help support rapid implementation and common standards (Exhibit 4). Clear decision rights are especially important for assessing gen AI proposals, supported by a transparent vetting process with well-articulated standards for each stage gate.
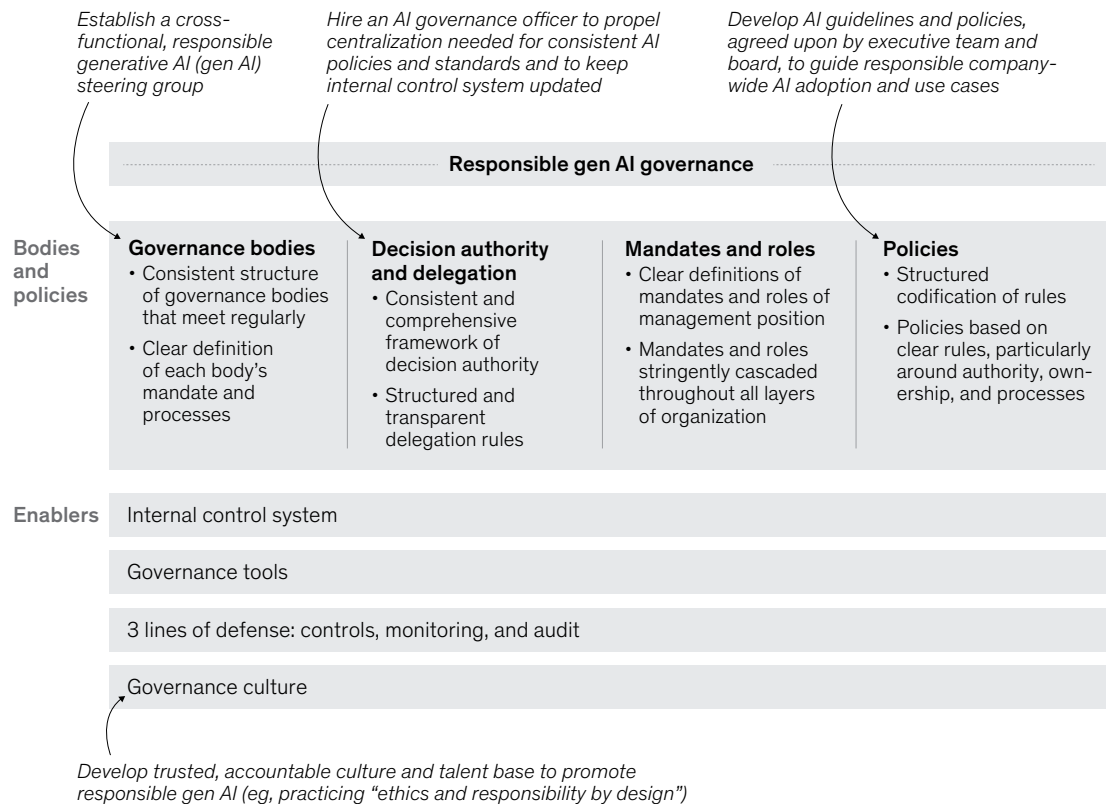
Performance infrastructure, data, and analytics
Modernizing performance infrastructure is crucial to accommodate gen AI's changes to the work landscape. The first step is redefining metrics to reflect the company's new operational strategy—and to allow leaders to see how gen AI itself is progressing across the organization. Such metrics can help the organization generate and sustain positive results. Next, a disciplined, stage-gated review process with clear go/no-go criteria separates the merely promising deployments from the ones most likely to be productive. Finally, with better measurement of productivity gains, customer experience improvements, and related outputs, companies can tailor coaching and training programs for human workers and interventions when gen AI's performance lags.

Exhibit 4

A revised governance approach to generative AI can help companies move at speed while mitigating risk.

Establish a cross-functional, responsible generative AI (gen AI) steering group

Hire an AI governance officer to propel centralization needed for consistent AI policies and standards and to keep internal control system updated

Develop AI guidelines and policies, agreed upon by executive team and board, to guide responsible company-wide AI adoption and use cases

Responsible gen AI governance

Bodies and policies

Governance bodies	**Decision authority and delegation**	**Mandates and roles**	**Policies**
• Consistent structure of governance bodies that meet regularly	• Consistent and comprehensive framework of decision authority	• Clear definitions of mandates and roles of management position	• Structured codification of rules
• Clear definition of each body's mandate and processes	• Structured and transparent delegation rules	• Mandates and roles stringently cascaded throughout all layers of organization	• Policies based on clear rules, particularly around authority, own-ership, and processes

Enablers

Internal control system

Governance tools

3 lines of defense: controls, monitoring, and audit

Governance culture

Develop trusted, accountable culture and talent base to promote responsible gen AI (eg, practicing "ethics and responsibility by design")

McKinsey & Company

Change management

It's a truism that changing technology isn't the hard part of transforming an organization—it's changing how people work that's hard. Early experience seems to show that this is even more true for gen AI, for which a good rule of thumb is "for every dollar spent on model development, a company should plan to spend three dollars on change management."

Communication is the starting point. By providing updates on what to expect and addressing potential anxieties, organizations can promote future adoption and create a culture of understanding among employees. But even better than just speaking to staff is to listen: the expertise and knowledge they contribute can make the difference between robust, cost-effective gen-AI-based solutions and gen AI gimmicks with little impact. In parallel, upskilling and reskilling initiatives can help smooth the transition.

Continuous innovation culture

Celebrating successes and sharing best practices is especially vital with a new technology such as gen AI, where innovation cycles are short. Simply keeping abreast of the latest opportunity requires both effort and openness: it's not a question of "buy versus build" but "buy *and* build," continually reviewing what the market offers.

Organizations can foster an environment where frontline workers feel empowered to contribute ideas, whatever their source—and where they feel free to reexamine assumptions about the potential role of partners and vendors in sourcing innovation. By encouraging continuous improvement through feedback and innovation, organizations can optimize the agent and customer experience while maximizing the value of gen AI.

To illustrate, consider the case of a leading European media and telecommunications company. This organization embarked on a mission to industrialize and scale gen AI by 2024, with tangible benefits expected within another year. The company's approach was not merely about chasing the latest tech trend; it was about empowering its workforce and transforming the customer

experience. To bring its vision to reality, the company identified a high-impact use case: a gen-AI-powered copilot designed to equip customer service agents with faster and more effective knowledge retrieval during calls.

Keeping agents informed and engaged was a top priority for the company, which hosted weekly working groups to gather qualitative feedback on usability and design. Additionally, quantitative feedback was collected through agent ratings of the AI-generated responses. "Office hours" provided a forum for questions and project updates, fostering a sense of ownership among agents. This transparency helped mitigate potential frustrations and ensured that agents felt invested in the success of the copilot—and led to substantial changes in design.

The user-centric approach proved instrumental not only in refining the copilot but also in encouraging successful scaling. By including frontline agents early in the process, the company made sure that the solution solved real problems in current processes and improved customer service and agent experience. The end result was a 65 percent reduction in average handle time for agents in finding relevant knowledge.

Honing gen AI's cutting edge

The most advanced companies are already combining tools, which can help overcome some of the limitations of the large language models (LLMs) and retrieval-augmented-generation (RAG) technologies at the core of gen AI's initial wave. In particular, LLMs and RAG struggle with complex processes—but automating only part of a process, even at a high level of reliability, often doesn't free enough worker time to create much benefit. Lack of cross-verification can leave LLMs and RAG prone to error. LLMs are also limited to text applications, while both LLMs and RAG using multiple data sources are expensive to build and scale. Finally, LLMs have only limited capabilities with quantitative analysis, making entire streams of value unreachable.

By contrast, a multiagent system helps humans coach machines to perform complex workflows, often by augmenting models with human expertise. By recursively breaking down a process into smaller tasks until each task can be executed reliably, the multiagent corrects itself, improving the quality of the outcome.

This approach has allowed a North American bank to transform its workflows for writing credit risk memos, a tedious and time-consuming task with little room for error. Relationship managers (RMs) would spend one to three days gathering data from a dozen sources (or more), analyzing multiple interdependencies, and, finally, writing a 20-page memo providing nuanced reasoning to support a lending decision.

A multiagent system now automatically identifies the correct data sources, ingests up-to-date data, and integrates qualitative and quantitative insights that reflect the latest business rules and products. It cites the data sources for each assumption and provides the key rationale behind quantitative trends, while also generating insightful comments based on the data it has integrated, content from prior memos, the RM's feedback, and human expertise.

Credit decisions are now 30 percent faster, while RM productivity has more than doubled. Most important, revenue per RM has risen by 20 percent.

Renewing gen AI potential

With so many companies already investing in gen AI, the question isn't where to start but how to find gen AI's potential quickly—and start earning the rewards. For the 97 percent of companies that haven't yet scaled gen AI in their operations, there's an urgent need for focus. At the same time, gen AI technology is moving so quickly that companies can often find opportunity simply by reassessing where they stand on a two- to three-month cadence. Issues to consider include the following:

— What is our current backlog of gen AI ideas? The first step is to review the ideas that haven't yet moved forward.

— What's feasible now? New combinations of gen AI tools mean that problems too difficult to solve a quarter ago may now be within reach.

— Which pilots aren't showing value? Rationalizing gen AI efforts means letting go of gen AI ideas that aren't meeting their promise, no matter how tantalizing. They might come back later, though, as tools evolve.

— Is our gen AI transformation approach still working? As the organization gains maturity in deploying gen AI, its transformation approach will likely need to evolve—becoming more agile as technologies become more familiar or more robust as risks come into focus. Applying a continuous-improvement mindset to the transformation itself helps ensure that it keeps finding more value from gen AI.

———————

With better performance, greater specialization, and increased accessibility, gen AI can revolutionize service operations across industries. A few judicious decisions now could help a company leapfrog its competitors in creating new and sustained value.