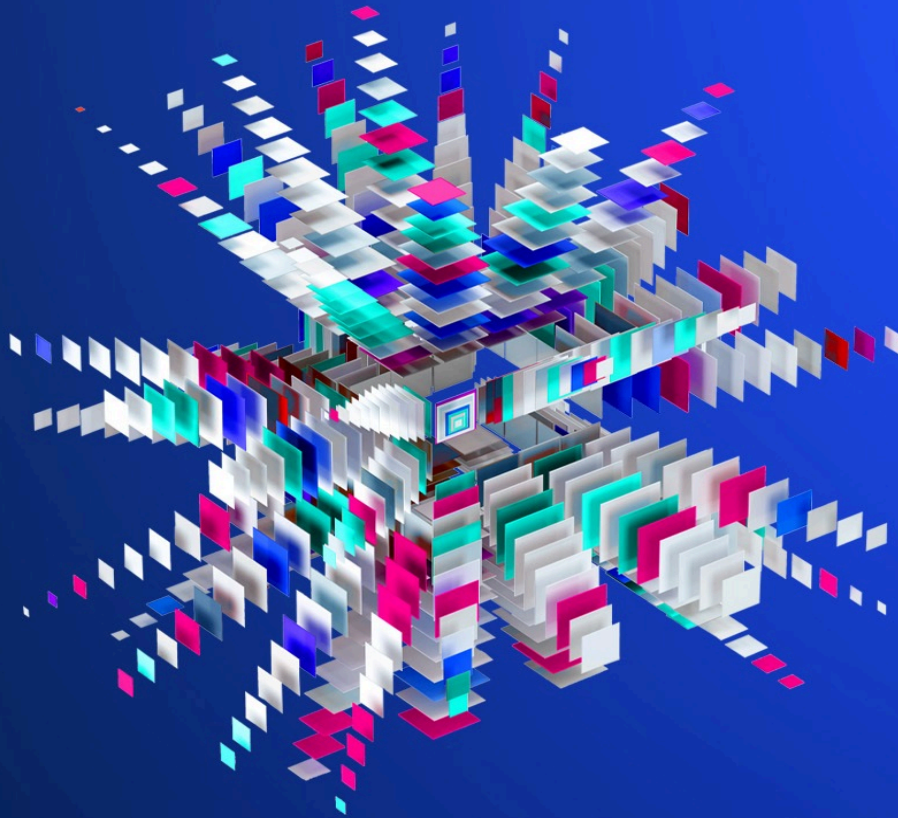# McKinsey & Company

AI for IT modernization: Faster, cheaper, better

Gen AI agents are starting to deliver breakthrough value, but only when companies figure out how to build and orchestrate hundreds of them.

This article is a collaborative effort by Aaron Bawcom and Matt Fitzpatrick, with Chi Wai Cheung, Dan Collins, and Dante Gabrielli, representing views from McKinsey Technology and QuantumBlack, AI by McKinsey.

December 2024

At the heart of virtually every large organization is a massive anchor slowing a business down: the tech debt found in legacy IT systems. Often built decades ago, these large systems form the technical backbone of companies and functions across almost every sector. As much as 70 percent of the software used by Fortune 500 companies was developed 20 or more years ago[1] (see sidebar "What are legacy systems, and how do they hold organizations back?").

Modernizing these aging systems and paying down tech debt have traditionally been considered an "IT problem," and business leaders have been content to more or less kick the problem down the road. The reasons are familiar: it's too expensive (often hundreds of millions of dollars), it takes too long (five to seven years), it's too disruptive, the return on the investment is unclear, and the current systems basically work.

But as technology infiltrates every nook of the business and becomes central to a company's ability to generate value, modernizing IT systems must become a CEO priority. The opportunities, and risks, generated from advances in technology—from generative AI (gen AI) to cloud to robotics—require modern technology foundations. In fact, technology enables about 71 percent of the value derived from business transformations.[2] The fact that the programmers who built and maintain these aging enterprise systems are reaching retirement age lends an even greater urgency to the need for modernization.

New developments in AI, particularly in gen AI, are radically recalibrating the costs and benefits of modernizing legacy tech and reducing tech debt as part of a larger set of changes in how IT operates. Consider a transaction processing system for a leading financial institution, which three years ago would have cost much more than $100 million to

modernize and today is well less than half of that when using gen AI. This shift makes many modernization efforts that were once too expensive or time-consuming suddenly viable. And with the ability to measure and track the direct cost of technology debt and its effect on P&L outcomes (in many cases up to 40 to 50 percent of total investment spend), companies can track the value they're generating.

While these are still early days, our experience indicates that harnessing gen AI can eliminate much of the manual work, leading to a 40 to 50 percent acceleration in tech modernization timelines and a 40 percent reduction in costs derived from technology debt while also improving the quality of the outputs. That value, however, is less tied to the technology itself and more to how it's used, with a particular focus on the following:

— *Improving business outcomes.* Converting old code into modern tech languages simply transports your tech debt from a legacy system into a modern one. Avoiding this "code and load" issue requires using gen AI to help make better business decisions and modernize what matters.

— *Enabling autonomous gen AI agents.* Building and training an army of gen AI agents that can work independently and collaboratively with human oversight on a range of end-to-end processes is proving to deliver significant improvements in IT modernization efforts (see sidebar "LegacyX").

— *Focusing on scaling value.* The value of the multiagent model comes from industrializing it so it can scale and be applied to multiple areas of the business and continuously pay down tech debt.

---

[1] Nia Batten, "Fix it, even if it 'ain't broke': The price of legacy technology," TechRadar, October 11, 2023.
[2] Aamer Baig, Sven Blumberg, Arun Gundurao, and Basel Kayyali, "Breaking technical debt's vicious cycle to modernize your business," McKinsey, April 25, 2023.

What are legacy systems, and how do they hold organizations back?

Legacy IT systems permeate every industry, often serving as the backbone of critical operations. In financial services, for instance, core banking platforms and investment management solutions handle transactions amounting to trillions of dollars daily on a global scale. Insurance administration systems manage policies worth $1.9 trillion in annual premiums in the United States alone.[1] Benefits management platforms facilitate the distribution of over $830 million annually in social services and benefits to millions of US citizens.[2]

Technology debt in these systems creates significant business issues:

— *Slow pace of innovation:* Systems built using outdated and difficult-to-understand languages severely hinder the organization's ability to adapt and innovate.

— *Limited compatibility with modern channels.* Legacy technologies typically do not integrate well with modern channels that demand real-time data and advanced connectivity.

— *Slow responsiveness to regulations:* Continually changing regulatory requirements become harder and more expensive to adhere to.

— *Resiliency risk:* A lack of understanding of how systems work, coupled with limited automation capabilities and often manual-testing processes, introduces instability issues that are hard to predict and fix.

— *High run costs:* The cost of running legacy systems is often significantly higher than that for modern systems. For a large European bank, for example, 70 percent of its IT capacity was spent maintaining legacy systems.

— *Inefficient use of capital:* The cost of running legacy systems prevents investment in more value-building development.

— *Unattractive to top talent:* Top programmers and software engineers are less likely to join a business that uses legacy systems and outmoded languages that they do not understand and are not valued in the talent marketplace.

---

[1] Annual Report on the Insurance Industry, Federal Insurance Office, US Department of the Treasury, September 2024.
[2] "Project: Public welfare expenditures," State and Local Backgrounders, Urban Institute, April 26, 2024.

How gen AI agents can improve both code and business outcomes

At its core, technology modernization involves transforming existing applications to take advantage of modern technologies, frameworks, and architectures. This process can include switching to easier-to-use programming languages, transitioning to modern frameworks that provide more functionality, restructuring systems to create modularity, or even remediating and migrating applications to run on cheaper cloud environments. Harnessing gen AI capabilities to make these kinds of changes, improve ROI on cloud programs, and

LegacyX

LegacyX is a McKinsey capability, powered by QuantumBlack, that simplifies legacy IT system modernization and delivers business value by using generative AI. It employs a range of specialized agents to handle end-to-end workflows, focusing on deriving the intent of legacy systems to develop better processes and accelerate modernization with a repeatable process. The tiered multiagent factory framework automates complex software development flows, modernizing both processes and applications simultaneously.

pay down technical debt requires a focus on the three areas mentioned above.

Improve business outcomes
Companies have tended to use gen AI in a blunt-force way by, for example, feeding legacy code directly into a gen AI tool that translates it into modern language. This code-and-load approach, however, essentially migrates your tech debt into a modern context. This trap is similar to the one that many companies fell into during the early days of cloud computing, where the focus was on "lift and shift," that is, moving existing applications to the cloud where legacy issues remained unaddressed.

The goal of a legacy-tech modernization effort should not be to convert as many lines of code as possible. It should be to improve systems and processes so the business can generate more value. That means using gen AI to understand the code you already have, determine what's needed to generate business value, and then modernize the processes that are necessary for achieving that outcome.

When properly applied, gen AI is able to translate the often-impenetrable legacy elements—documentation, code, observability data, call logs, programming approaches, etcetera—into simple English process descriptions in a matter of minutes. Engineers at one financial-services company interviewed a number of experts to supplement the limited documentation available and fed the transcripts into the gen AI model to provide it with better guidance.This clarity not only helps engineers to understand what the systems are doing but also allows business experts to help determine what's really needed. In this way, business and engineering experts can work together to determine what they want, what should be updated, and what can be discarded.

Enable autonomous gen AI agents
We have previously written that in software development, using gen AI agents to assist developers with coding can help some developers increase their productivity.[3] Our experience has shown that the next horizon of this acceleration will enable hundreds of gen AI agents to operate independently with human oversight, especially as the model scales (exhibit).
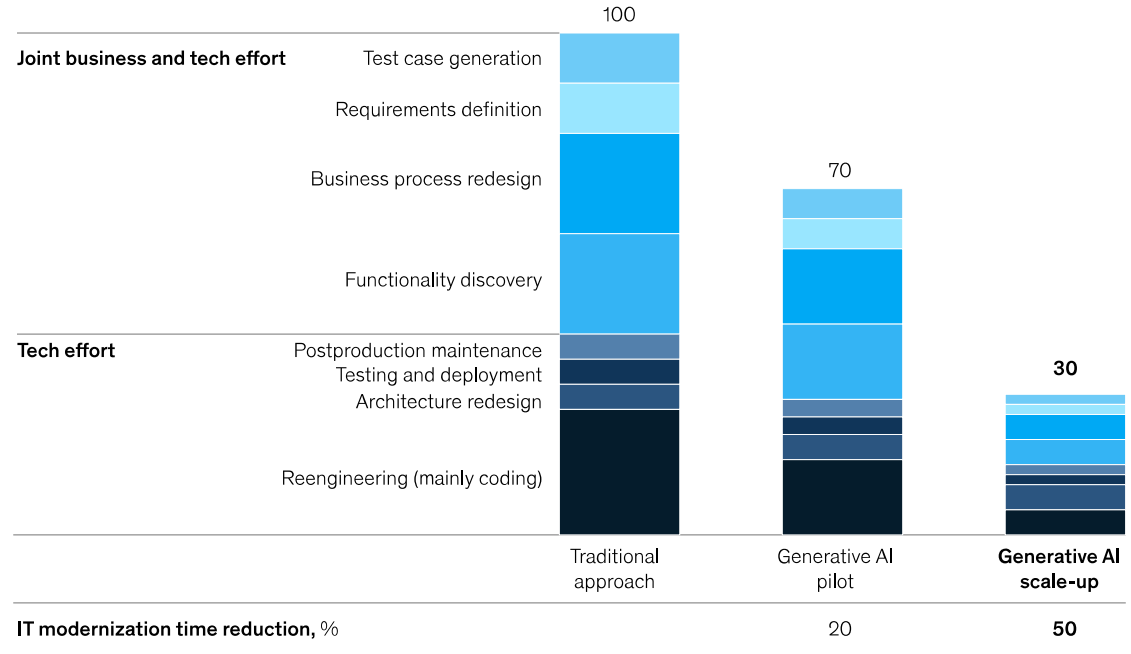
---

[3] "Unleashing developer productivity with generative AI," McKinsey, June 27, 2023.

As companies scale their generative AI agents, speed and productivity for IT modernization increase.

Journey of application modernization effort required, index (baseline = 100 full-time employees)

Source: McKinsey LegacyX experience implementing IT modernization efforts with generative AI

McKinsey & Company

The core of this autonomous-agent approach is the deployment of many specialized AI agents, each with distinct roles and expertise, collaborating on complex tasks. The tasks they can perform include data analysis, orchestrating sophisticated integrations, designing and running test cases, and refining outcomes based on real-time feedback from humans.

The real value, however, comes from orchestrating agents to complete not just tasks but entire software development processes. Data mapping and storage agents, for example, perform data analysis, compliance analysis, QA, as well as create relevant documentation. These agents work with security design agents that focus on threat analysis, information security policy, security design, and QA agents to develop safe, secure, and effective code.

To ensure gen AI agents deliver the right outcomes, it's important to implement a range of controls. Constructive feedback loops, for example, allow agents to review and refine one another's work. Gen AI agents can also be programmed to teach themselves to solve problems or escalate them to a human manager if they can't figure it out. Some gen AI agents can even ask the manager direct questions. Assigning IDs to each gen AI agent allows managers to quickly identify the source of an issue and address it. Similarly, organizations can develop specialized agents to automatically test and remediate the output of other agents based on identified ethical and bias concerns.

The role people play will continue to be vital to directing and managing gen AI agents. Experts like product owners, engineers, and architects will need to understand the intent of legacy systems, figure out what processes are important for the business, and develop and set goals and target states.

The power of the orchestrated gen AI agent approach became real at one banking company that had been trying unsuccessfully to modernize its mainframe for years. When it deployed a large collection of gen AI agents, the bank was able to migrate and improve a number of mainframe components as part of a migration to a Java, Angular UI, and PostgreSQL target state. Looking to modernize 20,000 lines of code, the company had estimated it needed 700 to 800 hours to complete the migration. The orchestrated gen AI approach cut that estimate by 40 percent. The relationship-mapping step, for example, went from requiring 30 to 40 hours to complete to just about five hours.

In another case, a top 15 global insurer used this approach to modernize legacy applications and services. The first step was to reverse engineer the code to better understand technical specifications and then use gen AI agents to generate code as well as automate discovery and conversion journeys. The result was an improvement in code modernization efficiency and testing by more than 50 percent, as well as a greater than 50 percent acceleration of coding tasks.

Focus on scaling value

The excitement surrounding gen AI has led companies to focus a significant amount of time on evaluating and selecting tools. That is important, but it pales in comparison to tackling the bigger issue and opportunity: how to scale gen AI. As one CIO recently said, "I don't want one tool to solve one problem; I need a capability to solve hundreds of problems."

Technology leadership should focus on developing a central, autonomous gen AI capability that can build sophisticated multiagent, end-to-end workflows. There are two primary components of this capability:

— *Factory.* A factory is a group of people who develop and manage multiple gen AI agents to execute a specific end-to-end process. The goal of the factory is to standardize and simplify the various processes that make up the development, deployment, and management of gen AI agents. A factory develops a standardized set of tools and approaches for agent development and management, such as monitoring, traceability, document management, and large language model access. An organization should consider developing five to ten agent factories in the early stages of maturity.

— *Platform.* A gen AI platform is a standardized set of reusable services and capabilities that factories can access. A platform should include a user interface, APIs that connect gen AI

services to enterprise services (such as Jira or ServiceNow), a range of supporting services (such as a data import service or agent orchestration service), and a library of gen AI agents that can be loaded to execute specific tasks. Companies should ideally develop one or two gen AI factories to determine exactly what sorts of services and capabilities they use in common, then standardize those elements and offer them through a platform. A dedicated team of relevant experts should both oversee the development and management of this platform and closely track usage against specific KPIs, such as service or feature use.

Next steps

Companies looking to move to this multiagent orchestration model should consider taking four steps:

— *Question any technology proposal that has a long timeline and requires many people.* Many large-scale IT projects have traditionally required many people to work for years to deliver value. Any proposals that follow this model should be treated with skepticism. That means reviewing all proposals and initiatives that are under way to determine how gen AI can reduce costs and shorten timelines. Be particularly thorough in reviewing programs and proposals that purport to use gen AI capabilities. The capabilities may be limited or ancillary and thus unable to deliver much value.

— *Focus gen AI on your biggest problems.* Small-scale initiatives lead to small-scale outcomes. Gen AI has the potential to radically redefine the cost-benefit of modernizing systems and reducing tech debt. Identify the largest and most complex technology problems—the ones that cost hundreds of millions of dollars, have multiyear timelines, and are responsible for large tranches of technical debt—and focus on developing gen AI solutions for them. Part of this effort should include revisiting previous tech modernization plans that were deemed too expensive or time-consuming.

— *Tie your business plan explicitly to value and track it vigorously.* While many companies have business plans, they are often superficial or limited in scope (for example, focusing on just the technology rather than the operating model). A strong plan provides a detailed view of the value at stake, the increments of value to be captured along the way (factoring in ongoing costs, like current and future infrastructure run costs, and one-time costs, like code modernization), the activities required to capture them, and a timeline that captures the break-even point. More important than developing the plan is revisiting it and ensuring that the modernization efforts are actually capturing the intended value. Without this kind of discipline, it's common for strong plans to slowly lose focus and default to delivering code rather than value.

— *Get ahead of the talent, technology, and operating-model implications.* As this multiagent approach scales, companies will need to understand and plan for the business implications. These include how to rethink your talent strategy and reskilling programs, how your operating model has to adapt, and how operating expenditures and capital expenditures will change, among other priorities. These are CEO- and board-level issues requiring thoughtful planning.

───────────

Companies have barely scratched the surface when it comes to effectively applying gen AI to modernize legacy technology and reduce technical debt. Focusing on how to orchestrate gen AI agents on meaningful business technology opportunities is the only way companies will be able to cut back tech debt and enable their tech estate to drive innovation and value.

Aaron Bawcom is a partner in McKinsey's Atlanta office; **Matt Fitzpatrick** is a senior partner in the New York office, where **Chi Wai Cheung** is a principal architect and **Dan Collins** is a senior principal; **Dante Gabrielli** is a principal product manager in the Philadelphia office.

This article was edited by Barr Seitz, an editorial director in the New York office.