



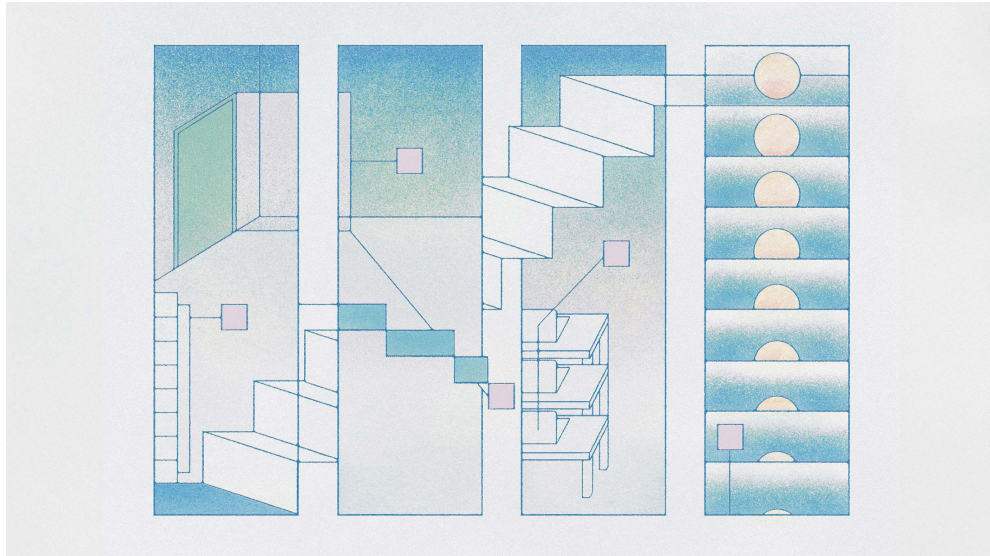
Spotlight / Generative AI

The Secret to Successful AI-Driven Process Redesign

Strong leaders put business transformation in the hands of all employees.

by H. James Wilson and Paul R. Daugherty

From the Magazine (January–February 2025) / Reprint S25012



Alanah Sarginson

In the late 1940s an engineer named Taiichi Ohno began developing the Toyota Production System, basing it on the Japanese principle of kaizen, or continuous improvement. At Toyota it led to constant small enhancements, with key suggestions coming from employees at all levels in manufacturing. Rather than revolutionizing its industry through bold, innovative, and risky endeavors, Toyota chose incremental but relentless improvement. Today it's the world's largest

automaker, and the Toyota Production System continues to be a model of how to manage processes across an enterprise. Some notable concepts that emerged with it have enjoyed a long afterlife: worker empowerment, a focus on perpetual cost reduction, total quality management, just-in-time manufacturing, root-cause analysis, data-driven processes, and automation with a human touch (*jidoka*).

As more operations become digitized, kaizen—augmented by generative AI and other advanced technologies—is once again reshaping process management. Now that features like natural-language interfaces have made gen AI accessible to nontechnical employees, it's driving both large and small process changes. With the help of AI, employees can synthesize data of all kinds, including unstructured data. They can turn once-inscrutable masses of numerical information into insight-driven workflow improvements, continuously increasing performance, reducing waste, and achieving higher levels of quality. Rather than displacing humans, as gen AI is widely presumed to do, kaizen 2.0 is moving them to the center of new machine-assisted processes and achieving a long-held aspiration of much management theory: putting business transformation in the hands of all employees.

But successfully reimagining business processes isn't as easy as asking ChatGPT to audit workflows. To get up to speed, leaders need to learn which processes are ripe for algorithm-powered redesign and understand how other companies have used gen AI to revamp them.

In this article, drawing on decades of experience providing advice to clients about technology and innovation, we'll describe how the best companies are deploying gen AI. We'll also introduce you to the future of kaizen: one in which fully autonomous agents can act independently to achieve goals, adapt strategies, analyze their environments, and complete complex tasks. However, as with all technological adoption,

humans will remain the linchpin behind gen AI's success and its ability to improve business processes.

Empowering Employees Throughout the Enterprise

Across industries from automaking to life sciences to consumer products, and across functions from R&D to manufacturing to supply chain management, gen AI is boosting employee empowerment in new ways. At Mercedes-Benz, for instance, this is happening on the shop floor, in the supply chain function, and in software design.

The company's MO360 Data Platform connects its passenger-car plants worldwide to the cloud, enhancing transparency and predictability across its production and supply chain operations and enabling the deployment of AI and analytics on a global scale. "With the MO360 Data Platform, we democratize technology and data in manufacturing," Jan Brecht, then the chief information officer of Mercedes-Benz Group, noted earlier this year. "Data is becoming everyone's business at Mercedes-Benz. Our colleagues on the shop floor have access to production and management-related real-time data. They can work with drill-down dashboards and make data-based decisions."

As more operations become digitized, kaizen—augmented by generative AI and other advanced technologies—is once again reshaping process management.

Using prompts in everyday language, rather than technical database queries, a production employee can ask about assembly-line bottlenecks or hard-to-notice opportunities for streamlining processes and receive data-rich insights from the AI. Such insights amplify, rather

than replace, workers' ability to generate improvements based on their own experience, powers of observation, and creativity.

The platform also helps teams identify bottlenecks in the supply chain. Meanwhile, the company's software developers are using GitHub Copilot, the AI-powered assistant that turns natural-language prompts into coding suggestions. This frees them up to spend more time addressing complex process issues and integrating software development across the enterprise.

To make data use more democratic, the company is helping people across the workforce acquire new qualifications in AI. The HR department has established [Turn2Learn](#), a program that gives frontline employees access to more than 40,000 courses on data and AI, including extensive training in skills from prompt engineering to natural-language processing. Thanks to generative AI, skilling initiatives, and digital ecosystems like the MO360 Data Platform, process change has gone from a niche technical skill to part of employees' everyday work experience at the company.

At the automaker Mahindra & Mahindra production teams can [send queries to gen-AI-driven virtual assistants](#) and receive step-by-step instructions for repairing industrial robots. That helps them quickly resolve technical issues and reduce machine downtime. Bhuwan Lodha, the head of AI at Mahindra Group, says the technology has significantly raised shop floor morale, delivering on the worker fulfillment that kaizen promises.

Redesigning Scientific Processes

In the pharmaceutical industry, gen-AI-powered synthetic data is helping workers create data-rich processes, reduce waste, speed up analysis, and strengthen quality control. Take the drug inspection

process. Pharma companies rely on automated visual systems to detect product defects. Unfortunately, the systems often generate false rejects, slowing the workflow and initiating expensive do-overs. This happens because the systems need to be trained with enormous amounts of images, but for many complex defects only a limited number of images exist.

To meet the challenge, Merck uses gen AI approaches (such as generative adversarial networks and variational autoencoders) to develop synthetic defect-image data. According to Nitin Kaul, the associate director of IT architecture, the gen-AI-enhanced system has helped Merck “understand root causes of rejects, optimize processes, and reduce overall false rejects across various product lines by more than 50%.”

Drug discovery is also being transformed by gen AI. Absci, a drug-development company, is now able to create and validate therapeutic *de novo* antibodies with a computer and zero-shot generative AI—in which a machine-learning model recognizes and classifies new concepts without having any labeled examples. In other words the AI designs antibodies that will bind to specific targets without using any training data on antibodies known to bind those targets. Creating them via AI instead of through trial and error could reduce the time it takes to get new biologics into the clinic from as much as six years down to 18 months, while increasing their probability of success. Waste, as kaizen has taught us, is not only a matter of materials but also a matter of time and effort.

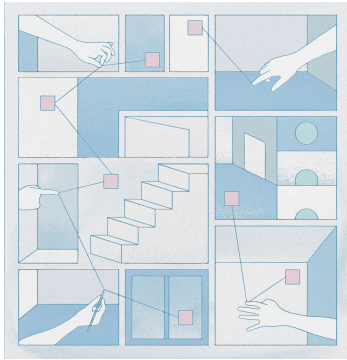
Augmenting Creative Processes

Several leading consumer-products companies are harnessing cutting-edge AI and digital technologies to catalyze the human creativity that drives growth in the sector. At Colgate-Palmolive, employees are

now using gen AI to speed up the process of devising new product formulations. Nestlé, Campbell's, and PepsiCo are reportedly using a gen AI platform that helps employees validate new product ideas and do market research. Coca-Cola is experimenting with a platform that combines the language capabilities of GPT-4 with DALL-E's ability to produce images based on text queries. The platform allows digital artists to incorporate distinctive branded elements from the company's vast archives, giving them a canvas on which they can create original artwork that will be used in billboard advertising.

Product and component design has long been a mix of art and science—combining the experience and sensibilities of a designer with the rigor of prototyping and testing. Across industries, gen AI is accelerating and transforming numerous elements of the process: creating 3D models of new ideas, suggesting modifications to designs, recommending the materials to be used, optimizing costs, rapidly creating digital prototypes, and identifying which ideas are most promising.

To what extent can such tools empower employees to improve creative processes? Consider the case of a single research engineer at NASA. Using commercially available AI software, Ryan McClelland reinvented the design process for specialized one-off parts at NASA's Goddard Space Flight Center, in Greenbelt, Maryland. Few organizations make more one-of-a-kind components to more-exacting standards than the space agency. These components can be critical in everything from astrophysics balloon observatories to atmosphere scanners, planetary instruments, space weather monitors, space telescopes, and even the Mars Sample Return mission.



Alanah Sarginson

In McClelland's new process a computer-assisted-design specialist starts with a mission's requirements and draws in the surfaces where the part will connect to an instrument or a spacecraft—as well any bolts and fittings for other hardware and electronics. The designer might also have to block out a path for a laser beam or an optical sensor. “The algorithms do need a human eye. Human intuition knows what

looks right,” McClelland notes. “Left to itself the algorithm can sometimes make structures too thin.” Under the supervision of expert human engineers, the AI then produces complex structure designs in as little as an hour or two. In traditional mechanical design, coming up with a design and analyzing it might take a week, followed by more iterations until an expert assesses the design for manufacturability. So it can take months of work to arrive at a solution. The structures designed with AI may look a little weird, but they are two-thirds lighter and 10 times less subject to stress than components created by the traditional design process.

Animating Physical Operations

Gen AI is also transforming the ways humans interact with complex physical systems, from robots to the human body to organizations like hospitals.

Stuttgart-based Sereact, a provider of AI-based software that automates warehouse operations, has pioneered the first commercially available solution that uses the transformer technology underlying ChatGPT to enable robots to understand natural language. The robots, trained on billions of simulated images, perform “pick and pack” tasks, which typically account for 55% of warehouse costs. Called PickGPT, the

technology allows human operators to simply type text commands into a chat interface; users with no technical expertise can direct and debug the system. Ralf Gulde, the CEO, calls it “the world’s most accessible way of interacting with robots.”

What’s next? The convergence of gen AI and digital twins, already underway, provides a glimpse of a future in which continuous process improvement becomes even more democratic. Digital twins are used to model complex systems—such as jet engines, wind turbines, factories, and human hearts—and simulate their functioning with an accuracy that allows users to remotely create solutions to any problems that arise (and often before problems arise). Digital twins can be used to make production processes more efficient, improve quality, increase operational performance, and create more-robust and -resilient supply chains.

Consider how twins are used in healthcare. Some 90% of the world’s top drug and healthcare laboratories already employ them in areas like preclinical drug development. [Atlas Meditech has built a platform](#) that lets surgeons practice on a virtual brain that matches the patient’s brain in size, shape, and lesion position. [Digital twins of hospitals](#) are used to make day-to-day decisions about staffing, operations, and bed management. A hospital can also use its twin to stress-test the organization against future scenarios like an earthquake with mass casualties. [Researchers foresee the day](#) when digital twins will be used to deliver precision medicine, diagnose diseases, and predict health and disease outcomes.

With gen AI now poised to expand the capabilities of digital twins, including by adding natural-language interfaces, we imagine that many more healthcare workers will have the tools to adapt processes

and almost instantly respond to new needs, a giant step forward for continuous improvement.

Autonomous Agents

The new AI agents take kaizen to a new level, not only offering advice but making decisions, taking action, and improving processes on their own. They range from simple chatbots to self-driving cars to robotic systems that can run complex workflows autonomously.

Consider DoNotPay, a company that aims to help consumers save money by performing a range of tasks from contesting parking tickets to canceling time-share memberships. Until recently, DoNotPay simply identified opportunities for customers to save money and encouraged them to act. But then the company integrated GPT-4 and AutoGPT into its software. The first user of these new features was DoNotPay's CEO. He gave the agent access to his financial accounts and prompted it with a concise yet complex command: Find me money. The agent discovered \$81 in unnecessary subscriptions and an unusual \$37 in-flight Wi-Fi fee. Then it offered to automatically cancel the subscriptions, drafted a letter to contest the Wi-Fi charges, and checked in with the CEO for review. As icing on the cake, it even drafted and sent emails that negotiated a 20% reduction in the CEO's cable and internet bill.

Gen AI is transforming the ways humans interact with complex physical systems, from robots to the human body to organizations like hospitals.

Traditional software is driven by precise, rule-based instructions and programmed to produce predictable outcomes. That significantly limits its ability to act autonomously. It lacks the capacity for humanlike

reasoning; decisions are hard-coded and don't incorporate the nuanced judgment and flexibility characteristic of human thought. In contrast, AI agents built on top of pretrained large language models are more dynamic and adaptable because of their ability to understand language and prompts. Agents built on multimodal foundation models have even more capability because they can generalize and understand, operate across, and combine many types of information simultaneously—text, code, audio, image, and video.

Autonomous agents exhibit three other similarities to human workers in kaizen-oriented settings:

Goal-oriented behavior. People set the goals, but AI agents act independently to achieve them, adapting their strategies when necessary. To do so, an agent can work across other software platforms at a company and interact with other organizations' software and language models to execute tasks.

Logical reasoning and planning. AI agents perceive and analyze their environments. They can break complex tasks down into their component parts and use reason to figure out the best way to achieve their goals.

Long-term memory and reflection. AI agents draw on past interactions to better understand intention and context. They learn from their experience to get better at their jobs.

Across industries, many companies are now deploying AI assistants or agents with varying degrees of autonomy. Walmart uses them to help manage inventory. At Marriott International, they optimize booking processes. At Nestlé, they improve supply chain processes. ADT is building an agent that will help its millions of customers select,

order, and set up their home security systems. Toyota is developing robot agents that could act as caregivers for elderly people or operate autonomously and smoothly in production processes on the factory floor. JPMorgan Chase is developing autonomous agents that will perform complex multistep tasks in the near future.

Meanwhile, technology companies, from tech giants to smaller firms and startups, are offering platforms and tools for creating autonomous agents and systems. Microsoft's AutoGen is an open-source programming framework for building such agents and coordinating them to perform tasks. It enables agents that can converse with other agents and can help create both autonomous and human-in-the-loop workflows. Meta's React is a free, open-source JavaScript library for building user interfaces, including with autonomous agents. Amazon Web Services recently introduced Amazon Q, which helps users set up semiautonomous AI agents that perform various tasks, including writing and debugging code. Google's Vertex AI Agent Builder and Semantic Kernel open-source development kit also help you easily build AI agents and integrate the latest AI models into your code base. OpenAI's Assistants API allows users to create agents within their own applications. The San Francisco-based startup LangChain offers an open-source framework for building applications based on large language models. LlamaIndex helps users build context-augmented gen AI applications, including autonomous agents and workflows. GenWorlds provides a platform for creating environments where AI agents interact with one another to execute complex tasks. MemGPT enables agent chatbots that can learn about you and modify their own personalities over time.

Tech companies are also including autonomous agents in their product offerings. Salesforce, for instance, sells Agentforce, a completely autonomous AI agent that can understand the full context of customer

messages and independently resolve a broad range of service issues without using preprogrammed scenarios the way traditional chatbots do. (Most chatbots can handle only specific queries that have been explicitly programmed into their system.)

Ecosystems of Autonomous Agents

Completing some tasks requires more than a single agent. In those cases companies may custom-build a system of agents wherein each one is expert in a specific task. Take the mortgage-underwriting process. When a human underwriter provides the instruction “Review this loan application based on our company’s lending policies,” one agent might extract relevant information from the application. Another agent might act as a librarian of bank policies, making them available to agents that compare the application against them. Yet another agent might generate a final report, recommending a course of action to the underwriter considering the loan. A “connector” agent might oversee and orchestrate the activity of all these agents.

Combining multiple agents and allowing them to communicate and collaborate with one another is critical to the development of AI systems that can autonomously manage end-to-end processes. Such solutions could transform entire functions like supply chain management, production, and marketing. At first glance, this might seem to be about achieving automation on a huge scale (and more fuel for sci-fi fantasists fearing world domination by AI). But the real story here is the opportunity for continuous improvement in the tradition of kaizen.

Consider an experiment conducted by researchers from Google and Stanford in 2023. They created 25 digital human avatars, each endowed with a distinctive personality and a back story, and set them loose in a simulated online world. As the avatars interacted and went about their daily “lives,” they created a credible facsimile of human behavior. They

made decisions based on their memory and past experiences, without real human intervention. The ability to rely on memory and to reflect on experiences and interactions like this is what enables agents to learn from one another and create ecosystems that continuously improve processes.

In another recent experiment, researchers at Stanford demonstrated that human-agent collaboration is a far more promising approach to automating complex workflows than robotic process automation (RPA), the most widely used technology. RPA is essentially a bot, hard-coded to perform a set of actions according to predefined rules. Lacking generalized reasoning and planning abilities, bots are easily tripped up by the variations and exceptions that inevitably crop up in complex processes. They're also expensive to set up and hard to maintain in the face of changing conditions.

The first workflow the researchers experimented with was revenue cycle management at a hospital. Most hospitals have departments that handle timely payment collection, patient insurance verification, prior authorization, and claims processing. A second experiment involved invoice processing at a large B2B enterprise, which was similarly complex, given the many contracts with widely varying conditions that the company dealt with. Most of the work in the two processes was still manual, despite attempts to automate it.

To overcome the limitations of RPA, the researchers employed a multimodal foundation model that learned from humans by watching video demonstrations and reading documents, greatly reducing set-up costs. The model, which encompassed a variety of agents executing different tasks, identified every step of each workflow with 93% accuracy. It leveraged its reasoning and visual-comprehension abilities to formulate plans of action, monitored itself and corrected errors, and

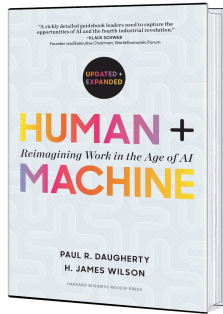
successfully identified the completion of a workflow with 90% precision and 84% recall. Those results suggest that the model could automate entirely new categories of workflows, such as those that contain hard-to-describe steps, require complex decision-making, or are knowledge-intensive.

As it executes a workflow, the researchers' model observes the effects of its actions and can compile a database of skills that can be transferred to other workflows. Though the goal was to achieve minimal human intervention, the researchers found that human integration into processes was critical: Humans were needed to ensure alignment with overall objectives, optimize models for interactions with people, and provide training and feedback to the agents.

...

As the Stanford research illustrates, though AI agents act on our behalf and in concert with one another, that doesn't mean that humans are out of the loop. Success with artificial intelligence will depend as much on people as it does on technology. While employees optimize agent models for human interaction, the agents will make decisions and operate with a greater degree of autonomy. Agents will constantly improve as they gain experience, and the humans overseeing the process will continually refine their design and performance. When both employees and AI agents are empowered, continuous improvements will be generated on both sides of the human-machine equation.

Even in the face of increasing machine autonomy, processes remain human-centered. In the coming age of autonomous agents, that will be one of the keys to kaizen.



H. James Wilson and Paul R. Daugherty are the authors of *Human + Machine: Reimagining Work in the Age of AI*, New and Expanded Edition (Harvard Business Review Press, 2024), from which this article is adapted.

[Buy it here](#)

A version of this article appeared in the [January–February 2025](#) issue of *Harvard Business Review*.



H. James Wilson is the global managing director of technology research and thought leadership at Accenture Research. He is the coauthor, with Paul R. Daugherty, of *Human + Machine: Reimagining Work in the Age of AI*, New and Expanded Edition (HBR Press, 2024).

✉ [@hjameswilson](#)



Paul R. Daugherty is the former chief technology officer of Accenture and currently a senior technology adviser at the firm. He is the coauthor, with H. James Wilson, of *Human + Machine: Reimagining Work in the Age of AI*, New and Expanded Edition (Harvard Business Review Press, 2024).

✉ [@pauldaugh](#)