Digital Article / Generative AI

How to Make Enterprise Gen AI Work

Move from individual experiments to robust corporate applications.

by Melissa Valentine, Daniel J. Politzer, and Thomas H. Davenport

Published on HBR.org / September 18, 2025 / Reprint H08W2M

Eugene Mymrin/Getty Images

If your organization is still relying on ad hoc employee experimentation with gen AI, it's time to shift gears. While experiments like using Claude to draft emails or ChatGPT to brainstorm can yield learning and productivity benefits on an individual or unit level, they are typically unstructured and unmeasured and rarely yield large-scale results. (We believe that this is part of why companies aren't yet seeing bottom-line impact from AI investments.)

To find measurable business value from gen AI, leaders must move from open-ended individual experimentation to structured, enterprise-aligned deployments. Whether they focus on just one business unit or the whole organization, these tools translate gen AI's open-ended capabilities into applications that address specific use cases. They include everything from enterprise knowledge assistants and customer service agents to data ingestion and verification systems, automated regulatory compliance monitoring, and large-scale marketing content generation engines.

We are researchers who study AI and its business applications, and we have been specifically exploring how AI can support process management in large organizations. As we've observed real-world companies move beyond individual experimentation to build enterprise AI tools, we've seen that doing so requires changes in how businesses approach both data infrastructure and collaboration. In this article we'll offer best practices for each and offer real-world examples of companies that are successfully making the shift toward an enterprise focus—and looking ahead to agentic AI.

Shifting from Individual to Enterprise AI

Because of the way they are built, structured, enterprise-aligned AI applications offer multiple benefits. Integration with multiple systems means a wider array of data, which improves recommendations and outputs. It's easier to achieve scale, because the same tool can support thousands of users. Outputs are more consistent, as they can be standardized. Use is more closely controlled and auditable, reducing governance risk. Because they are evaluated closely for impact on cost, speed, quality, or innovation, we also believe they will be more likely to yield greater ROI, since with measurement comes the ability to rapidly address flaws and calibrate costs.

Some companies have already made the necessary pivot to building these kinds of systems. Earlier this year, for example, Johnson & Johnson concluded that most individual experimentation was not yielding measurable business value and decided not to devote additional resources to it. Instead, the firm prioritized just a few enterprise-level gen AI projects focused on applying the technology to strategic priorities for the firm: drug development, HR policy access, assisting sales reps in communicating with physicians, and identifying and mitigating supply chain risks. Coca-Cola has also decided to focus on large-scale, enterprise-orchestrated AI projects, such as creating new marketing content. The company is using gen AI to customize 10,000 different versions of 20 proprietary marketing assets for the 180 countries and 130 languages in which it does business around the world.

The same qualities that make enterprise systems more beneficial, however, make them difficult to build. Integrating high-quality data across complex systems, orchestrating data flows, and aligning outputs with business are huge challenges. While the findings from recent studies citing low figures for the production implementation of AI initiatives have been debated, it is clear that companies are struggling on these fronts.

Preparing your data flows for an enterprise use cases and focusing on how business and development teams work together during the project can help address these issues. Here's how.

Build Data Readiness

Organizational readiness for gen AI starts with making unstructured data and its flows more visible, structured, and strategically prioritized. Take Northwestern Mutual's knowledge assistant, which integrates internal proprietary data that has been systematically and reliably curated, updated, and made searchable first. But most organizations have focused their data quality and management efforts on structured data over the past decades, and haven't addressed the unstructured data that organizations need to adapt LLMs to their specific business issues.

This work proceeds best when led at the functional group level. Groups should map their data flows and work processes and identify current data being generated. Assess whether new data assets will be needed and if so, which ones.

This work can pose a challenge for business teams because they often see it as secondary to—and a distraction from—their "real" responsibilities. To combat this, frame the effort by underscoring the benefits of enterprise AI, explaining that information about a product or service can be just as valuable to the organization as the product itself. Leaders should designate clear ownership within the unit for data curation and quality going forward, with incentives, if possible.

Let's look at how Accenture has met this challenge. The company's marketing arm, under chief marketing and communications officer Jill Kramer, has created 14 custom AI agents that have sped up the organization's workflows considerably.

As the project kicked off, Kramer prompted her entire function to question their internal workflows and documentation norms, explaining that proceeding on any AI development without this work would just "accelerate chaos." Data stewardship had never been central to marketers' identity, but Kramer reframed it as a leadership imperative—as "the only way" to ensure the marketers had an optimal experience using agents. She asked, "What *is* our process? What parts are 10 times more annoying than they should be?" and designated ownership of the process of finding the answers to these questions broadly across the function.

This helped her team successfully undertake a major data infrastructure initiative, mapping workflows such as how marketing plans were built, how documents and assets were created and accessed, and how decisions were made across the function. This effort led to better data visibility and, ultimately, the development of agents that assist with research, editorial planning, and resource allocation, areas identified as bottlenecks through their workflow mapping work. These agents have allowed Accenture's marketing function to bring a campaign to market 25–35% faster.

Improve Collaboration

AI often requires new ways of collaborating—and new collaborations—for both business and tech teams. Business teams cannot simply hand off requirements as is typical in IT projects; they must stay deeply involved, curating data and iterating on AI outputs and use cases that the development team delivers. At the same time, development teams must accept that governance and evaluation cannot be solved purely through technical solutions, and that they must solicit and work with the business team's input. The development team's collaboration challenge is amplified by the need to orchestrate AI use for a *whole ecosystem* of teams across functions (BI, commercial teams, operations, HR, finance, etc.), many of whom may have very different structures, cultures, data flows, and needs. All groups must become active and continuous partners.

This makes role clarity—who owns what, how decisions get made, and how feedback loops are structured—both particularly critical and unusually complex. Typically that means business teams becoming more involved in the process than in traditional tech projects. They must become active partners, involved in the iterative process of design and evaluation of the new technology from inception, curating their domain-specific data, administering their own bots, and responding to user-driven feedback. And both groups need to collaboratively define access policies, evaluation approaches and standards, and user training.

Secondly, a measurement and evaluation framework is critical to help collaborating teams stay aligned. The team should define together at the start of the project how progress and success is to be defined, measured, and tied back to key business objectives.

Take, for example, JetBlue's development of BlueBot, an LLM-powered tool that enables business functions across the company, such as communications and HR, to access data and knowledge based on their specific roles and needs. (This story is drawn from a talk given by Sai Ravuru, Senior Manager of Data Science and Analytics at the airline, and conversations one of us [Melissa] had with him afterward.) Early in the project, the development team focused on building and the technology infrastructure that would enable users to curate their data and evaluate BlueBot's outputs for their function, and the business functions focused on selecting and curating the structured and unstructured data and giving ongoing feedback, with decision rights elevated to the C-suite. The involvement of the business team in these specific ways enabled diverse and decentralized business functions to take responsibility for the quality of BlueBot's outputs, and, ultimately for their own use of the tool going forward. They also defined increased productivity through time saved as their core business goal. The same role structure and goal continued after launch as the teams evaluated user inputs and feedback and BlueBot's outputs and addressed them to improve the bot.

As a result of this collaborative work, JetBlue has seen a 5–10% time savings because business teams no longer need to waste minutes searching for or requesting dashboards, and development teams no longer need to develop or maintain them.

Prepare for Agentic AI

It's time for companies to invest in enterprise gen AI now—but the bar for data and collaboration will become even higher for enterprise *agentic* AI use cases, which are the next technology on the horizon. Digital agents perform tasks with even more autonomy and so require even more upstream effort.

From what we've observed with companies like Salesforce that are beginning to roll out AI agents, the same requirements apply as for other enterprise AI tools—documenting data flows and needs and collaboration between development and business teams will be key. That's all the more the case because successful rollouts seem to be iterative, requiring shifts and continued input from all parties for real-time operational fixes and updates as knowledge, tools, and policies evolve. For example, Salesforce initially piloted the Agentforce system with only 200 selected users. In the next phase, 10% of help requests were routed to the agent. The team monitored weekly performance stats, identified weak spots, and made adjustments before scaling up. At this point, 85% of Salesforce's customer issues are resolved without human intervention, and even employees make use of the tool.

...

While it's time for companies to shift their AI resources to enterprise solutions, that doesn't mean that all individual experimentation should go away. We've seen that gen AI experiments continue to be fruitful in situations in which users are already tech-savvy and somewhat autonomous, such as game development, for example.

But we advise that the bulk of the work be moved to larger deployments —and that organizations build the mindsets and skills needed to prepare for and execute on them.

This article was originally published online on September 18, 2025.

Melissa Valentine is an associate professor at Stanford University in the Management Science and Engineering Department, and a senior fellow at the Stanford Institute for Human-Centered AI.

Daniel J. Politzer is a research associate at Stanford University in the Management Science and Engineering Department, and the founder of Alerce Advisors, a consultancy that offers AI advisory services and fractional management.

Thomas H. Davenport is the President's Distinguished Professor of Information Technology and faculty director of the Metropoulos Institute for Technology and Entrepreneurship at Babson College, a visiting scholar at the MIT Initiative on the Digital Economy, and a senior adviser to Deloitte's Chief Data and Analytics Officer Program.

