eBook

AI leaders reveal
what works in the modern AI era

Domino

Executive summary

Eighteen months after ChatGPT shattered preconceived notions of technological progress, generative AI (GenAI) continues to captivate executives across industries, from finance to agriculture. This surge of interest is not unwarranted; enterprises are poised to double their 2023 investments in GenAI from $19 billion to a staggering $151 billion by 2027.[1] This momentum is part of a broader trend, with overall AI investments nearing $1 trillion over the past decade.[2]

Even the most stoic executive can be forgiven for experiencing a Fear of Missing Out (FOMO) in this climate. The excitement and investment mandates surrounding GenAI are indicative of its perceived potential to revolutionize business operations and drive significant competitive advantages. But what is really happening inside the world's largest companies as they unleash AI? We asked leaders at nearly 300 enterprises across North America and Europe this question.

This study, commissioned by Domino Data Lab and conducted by **BARC**, explores the multifaceted landscape of AI implementation, offering a comprehensive guide to understanding and navigating the complexities involved. As corporate boards increasingly throw their support behind AI initiatives, companies are rapidly deploying both GenAI and predictive AI into production. However, this rapid deployment is accompanied by the need for robust governance frameworks and scalable infrastructure to support these advanced technologies.

[1] https://www.idc.com/getdoc.jsp?containerId=prUS51572023
[2] https://edgedelta.com/company/blog/ai-investment-statistics

This eBook provides insights into:

- Diverse strategies for implementing GenAI and scaling AI initiatives (board support, production deployment, and infrastructure adaptation)
- The critical need for updated AI governance (manage model risk, ensure compliance, and maintain transparency)
- How AI leaders and teams can completely and effectively leverage AI's potential
- A comprehensive roadmap for achieving AI innovation and competitive advantages

Key findings

1. **Corporate boards are all-in on AI:** 21% of enterprises have a blank check from the board for all types of AI.
2. **Companies are rapidly putting GenAI into production,** but traditional forms of AI (predictive AI, data analytics) remain the workhorses.
3. **No single path to GenAI:** multiple approaches to implementing a GenAI stack have taken hold.
4. **Everyone needs to upgrade AI governance:** 95% face a governance remodel or reboot to update their frameworks and processes for today's modern model landscape.
5. **Compute scarcity? Not a problem for AI-ready companies:** 92% of companies report no hindrance in scaling AI due to compute scarcity.

1. Corporate boards are all-in on AI

Corporate boards of directors are the ultimate arbiters when it comes to company decisions on strategy and investment. But, in today's age of AI, where do today's boards stand on AI?

The answer? Nearly all boards are paying attention to AI. Only 5% of respondents said their board hadn't engaged or set strategy on AI and 72% reported their boards were providing sufficient support for all AI, including GenAI. Indeed, 21% report that they have a blank check from the board to do whatever is necessary to support all types of AI.

72% of AI leaders say they have sufficient budget for all types of AI

One in five companies have a "blank check" for all types of AI

However, some boards may be overly excited by GenAI. A third (34%) say GenAI initiatives get more board support than they deserve, depriving support from other forms of AI and 8% said it was difficult to get funding for anything other than GenAI. This may show that exuberance over GenAI comes at the cost of continuing to innovate with predictive AI or traditional data analytics.

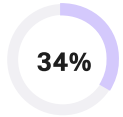
One striking finding here is the revelation that one in five corporate boards has written a "blank check" for AI, regardless of the type. This data shows that most of these companies are on the smaller side in terms of revenue; a finding which may indicate that these companies see AI as a path to competitive differentiation and growth.

These choices make clear that AI is at a crossroads:

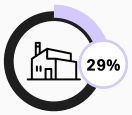
- Nearly one in three (29%) companies with $100M-$499M annual revenue (AR) offer AI investments with a carte blanche.
- One in four (24%) companies with AR ranging $500M-$999M have the same.

Will enterprises be content with having a few internal chatbots, email writers, and code assistants or will they wield AI to drive a significant step change in revenue and competitive advantage?

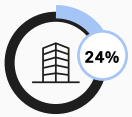
GenAI gets more support than it deserves, causing other forms of AI to lack support



"I have a blank check to do what's necessary for AI, regardless of type."



$100-499 MILLION



$500-999 MILLION

$1-5 BILLION

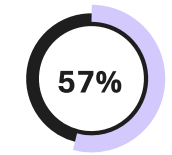
>$5 BILLION

2. Companies are rapidly putting GenAI into production, but more traditional AI is still the workhorse

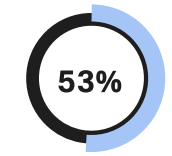
While GenAI gets all the attention, slightly more companies are getting predictive AI (53%) — i.e. machine learning (ML) — and advanced analytics projects (57%) — into production. However, GenAI has been adopted at breakneck speed and almost half (49%) of respondents claim they already have GenAI use cases in production today — a finding that may be explained by prevalent use of chatbot products like ChatGPT and Midjourney to assist in a broad range of tasks. The distinction between these types of projects is blurring though with 41% of leaders saying that they have projects that use both predictive and generative AI in production.

On the flip side, a worryingly large share of companies are early in their AI journeys. Over half (51%) are still in the planning, researching or proof of concept stage when it comes to GenAI, and 47% have not yet put predictive AI projects into production despite the fact that these technologies have been available for decades. This indicates that companies struggle to move from experimentation to productionalization across all types of AI.

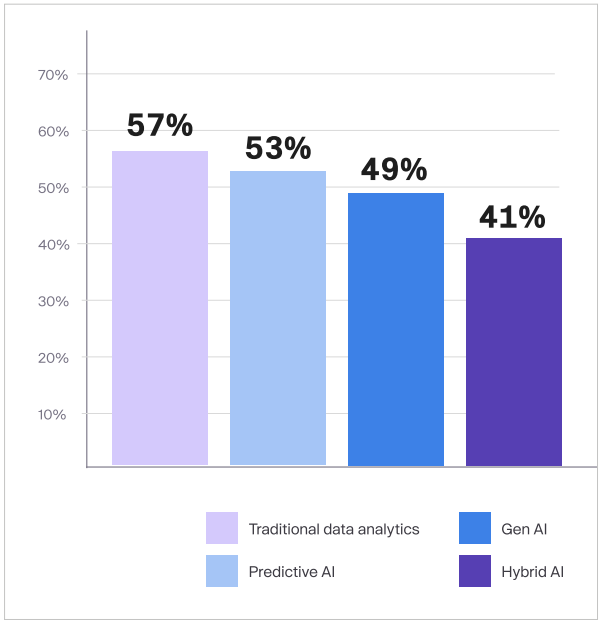
Why is GenAI, which only really debuted 18 months ago for most businesses, rising so fast? Though the technology surely has yet to come into its own, one clue could be found in the early returns on how useful AI leaders find it. Respondents ranked GenAI as the workload type with the highest ROI, followed by hybrid workloads and traditional predictive AI/ML. Traditional query-based data analytics ranked last for ROI.

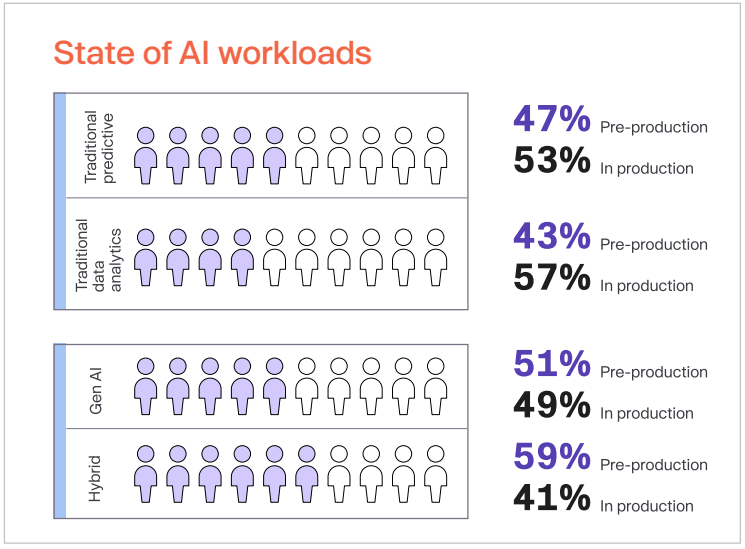


Traditional data analytics



Predictive AI





3. No single path to GenAI: Multiple approaches to implementing a GenAI stack

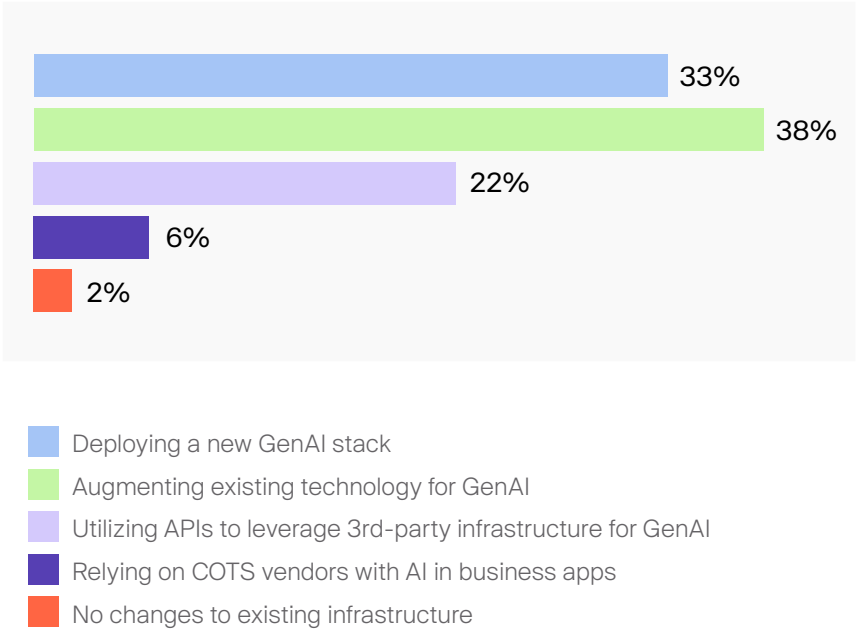
GenAI adds complexity and new challenges to building and operating AI, meaning that these new models require more data, more compute, and new techniques. For data science leaders, this means their AI stacks now need more GPUs and silicon, more access and volumes of complex forms of unstructured data, and the adoption of new processing and model training techniques.

So, what are they doing?

In terms of their approach, more than 90% of enterprises plan to make some infrastructure adjustments to account for their GenAI journey, most commonly using updated versions of their pre-GenAI infrastructure. 38% are powering their GenAI initiatives by augmenting their existing AI stack. However, nearly as many (33%) are investing further, deploying a new stack entirely for GenAI. Slightly fewer than that (22%) are simply turning to cloud providers or other third-party APIs to access third-party infrastructure for GenAI for capabilities such as inference.

Not surprisingly, respondents very seldomly reported that their companies were taking the path of least resistance. Only 2% are meeting the expanded role of AI in their business with no changes to infrastructure at all. 6% relied wholly on their traditional software vendors to bake AI into business applications, a finding that could point to a reality where companies need their own differentiated AI solutions, not just AI features baked into traditional software products.

Impact on AI tech stack strategy/investment



4. Everyone needs to upgrade AI governance

Enterprise plans to scale AI today require support for a diverse and evolving range of workloads, leveraging a variety of compute, frameworks, and tooling. But these companies do not operate in a vacuum. Regulators, lawmakers, and ultimately, their own customers increasingly expect AI innovation to go hand in hand with governance. Understandably, model risk and model governance have become significant areas of concern for companies.

It is therefore no surprise that enterprises feel underprepared for today's governance needs. Our study found 95% of enterprises face a governance remodel or reboot to update their frameworks and processes for today's modern model landscape.

What was once a conversation focused on explainability and bias is now a deeper discussion of how people and processes enable responsible AI. This concept includes specific steps to follow during model development, new roles that review information along the way, and the ability to audit all this work later. Data science teams demand that all this must happen while preserving the ability to self-serve the infrastructure they need, innovate with cutting-edge tools, but preserve a system of record down to the level of infrastructure, frameworks and drivers.

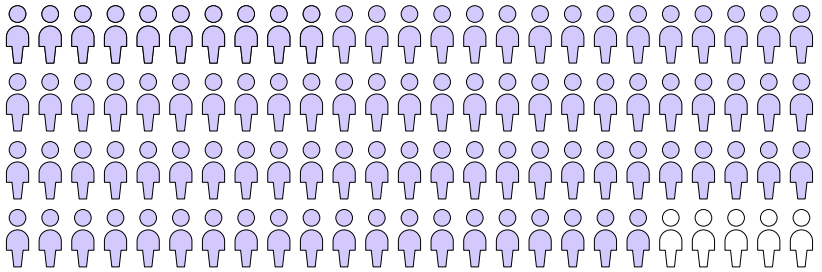
As a recent and complex development in scaling AI, GenAI has caused enterprises to grapple with how to streamline this process and address the ever-present hallucination risk. In a world with both traditional and GenAI to manage and soon to mix into hybrid workloads, how will companies enforce tracking model development without being too onerous? How can they enable process reproducibility and traceability? Can they leverage their prior processes defined for pre-GenAI governance?

Reassuringly, most companies say they have a baseline of necessary AI governance infrastructure and processes in place. They believe that they can incorporate more data sources and more data into the AI equation. Most (65%) of our respondents said their companies need only add new processes to existing governance frameworks to compensate for this change.

For others, GenAI requires a deeper revision of their governance plans. Nearly one in five (18%) say they will entirely rewrite their governance frameworks to address the advent of GenAI. 12% plan to replace their current governance framework entirely as a result of GenAI.

95%

of enterprises must replace, re-write or update their AI governance frameworks and processes for today's Gen-AI enabled model landscape



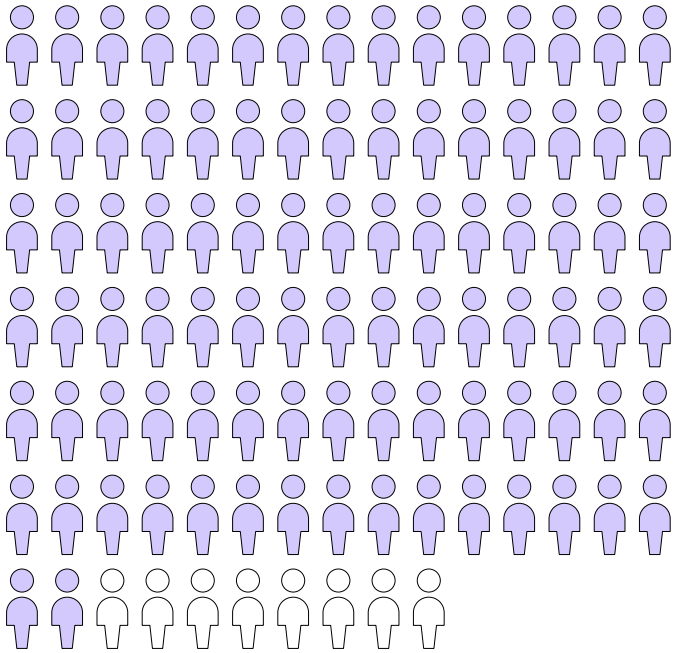
5. Compute scarcity? Not a problem for AI-ready companies

Securing access to today's most advanced GPUs is like trying to book a table at the trendiest restaurant in town — high demand, limited supply, and fierce competition. In one case, a venture firm was reported to have been luring potential portfolio companies into investment deals with a stockpile of cutting-edge GPUs. Nonetheless, our survey shows that, at least today, not many companies are hindered by compute scarcity and most are confident in leveraging their AI stacks across various vendor silicon and infrastructure.

- Despite the scarcity of some top-of-the-line GPUs, 9 out of 10 companies (92%) say compute scarcity, considering the full gamut of GPUs and CPUs, is not impacting their needs to scale AI.
- In a nod to their capable teams, when it comes to peripheral technologies — frameworks, libraries, etc. — 87.1% of enterprises indicated their staff is up to managing multiple vendors.

A number of conclusions could be drawn from this data. On one hand, these findings highlight that capable teams within today's enterprises appear to experience little pain leveraging various technologies composing their AI stacks. However, perhaps they are also a symptom of widespread immaturity we saw with workloads in production vs the research or POC stage. The question then becomes how these teams can achieve corporate AI transformation securely, cost-effectively, and efficiently. With 90% revamping their infrastructure for complete AI, this is the perfect opportunity for a platform that can handle multiple GPU vendors, mix and match cloud platforms, and the full spectrum of complex workload types.

Enterprises are confident in leveraging their AI stacks across a range of vendor silicon and infrastructure



92%

say compute scarcity, considering the full gamut of GPUs and CPUs is not impacting their needs to scale AI

Conclusion

The pressure to “do” AI is even greater now than it was before and data science and IT leaders must better understand how to navigate the complex AI landscape in order to maximize business value. After reading this eBook, you now understand the implementation challenges ahead, what it takes to effectively implement GenAI and scale AI initiatives, the critical need for updated AI governance, and what a comprehensive innovative AI roadmap looks like. Make the next step in your AI journey and leverage AI to its full potential.

Methodology

The Domino Data Lab survey was conducted by BARC among 279 respondents in North America and EMEA with Senior Director to C-level titles who are involved with or knowledgeable about AI strategy or projects at companies with $100M+ annual revenue, during May 2024, using an email invitation and an online survey.

Discover more at

domino.ai/platform/ai-governance →

eBook

AI LEADERS REVEAL WHAT WORKS IN THE MODERN AI ERA

About Domino Data Lab

Domino Data Lab empowers the largest AI-driven enterprises to build and operate AI at scale. Domino's Enterprise AI Platform unifies the flexibility AI teams want with the visibility and control the enterprise requires. Domino enables a repeatable and agile ML lifecycle for faster, responsible AI impact with lower costs. With Domino, global enterprises can develop better medicines, grow more productive crops, develop more competitive products, and more. Founded in 2013, Domino is backed by Sequoia Capital, Coatue Management, NVIDIA, Snowflake, and other leading investors.

Learn more at **www.domino.ai** →

Domino