



How to successfully navigate your AI journey

Has your business reached a tipping point in its AI journey? Enterprises looking to move from simply consuming AI to strategically wielding AI for competitive business advantage need to assess their architectural and team readiness. Here's how.



At a glance

1. The bulk of IT leaders say their organization is still in the planning or initial application phase of its AI journey, with clear intent to continue investing and progressing efforts.
2. However, to drive AI projects from operational-efficiency initiatives to models that unlock real business growth, organizations must sidestep potential oversights on required infrastructure and staffing resources.
3. For IT infrastructure, **open, hybrid,** and **AI-by-design** are buzzwords for the next level of success while **inclusive** and **collaborative** are descriptors to keep in mind for staffing resources.



¹ "Architect an AI Advantage," Sapio Research, January 2024

Most organizations have big ambitions for their AI rollouts, with 94% planning to increase their AI budget over the next 12 months to accelerate the pace of deployment, according to findings from the **Architect an AI Advantage**¹ survey.

Despite high investment levels, the bulk of IT leaders we surveyed also say their organization is only in the planning or initial application phase of their AI journey. Additionally, HPE survey results showed that most organizations are still predominantly using AI for projects that relate to operational efficiency over business growth, which is where the magic really happens. It is here that AI can be leveraged to open the most exciting opportunities—for example, using data modeling to deliver trend forecasting that allows businesses to preemptively pivot their offerings to meet customer demand ahead of competitors.

Reaching the business-growth end goal of any AI journey isn't easy. AI projects won't deliver value if an organization isn't able to access relevant data to train and run their models. Beyond this, potential obstacles to AI success are myriad: perhaps, there wasn't a proper understanding of what the AI model was aiming to do, the right architecture wasn't in place, or the correct internal resources were not consulted. So, how do organizations sidestep these oversights to make AI appropriate for next-level development and deployment across the enterprise?

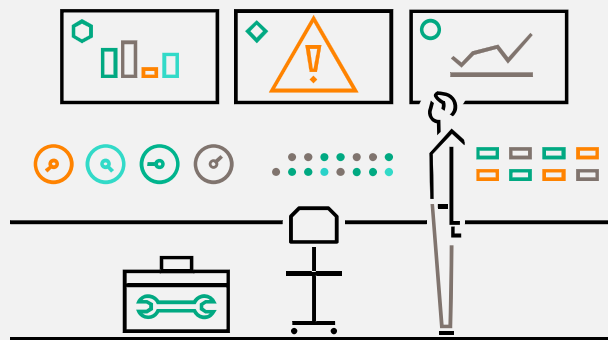


Considerations for a pilot-to-production move

Organizations can run smaller efficiency projects—such as chatbots to automate responses to help desk queries—with a single server. However, when they want to productize AI, or truly make it part of their organization's DNA, they need to take their setup to the next level. What exactly does this entail? It means implementing enabling architecture and balancing the expertise of staff resources.

If the goal is to have enterprises progress past initial AI applications toward business-growth projects, the first step is to understand what needs to be accomplished, using AI as a tool to deliver business outcomes. Once this understanding is clear, enterprises must look at what kind of team they need to plan, manage, and monitor AI. **Are the right people involved in the organization's AI projects?** Is there a balance between business decision-makers and IT influencers?

Next, **businesses must look at their existing infrastructure.** This might require a unification of their multigen IT environment. It may necessitate a move to a hybrid-by-design approach, including AI-specific hardware, software, and data management solutions. Each business will have a different path, but the bottom line is that the IT environment must be capable of handling compute-intensive AI workloads seamlessly.



Upping the ante on internal staffing resources

According to our latest research, there appears to be a **disconnect between business perception and actual needs when it comes to bringing AI teams together.** While 97% of IT leaders believe the right people are currently involved in their organization's strategy discussion, AI is still primarily viewed as an IT cost for the bulk of respondents, and projects are geared toward driving efficiency of IT systems and not unlocking business-led growth via insights and analytics.

When moving onto business-led projects, the resource needs will be more complex and, as such, **the expertise of AI-skilled data scientists and machine-learning developers or engineers cannot be overlooked.** AI teams need to be broader than top-level decision-makers, with the full team (including domain experts) made part of the discussions right from the strategy stage and not just during the rollout.



The lack of skilled personnel is always a challenge as AI use cases expand. Though we have a lot of data analysts and business analysts inside our organization, not many of them are trained in AI and ML.”

– Healthcare manufacturing CISO, on the need for specialized AI skills and training



Without a technically skilled team helping to shape the rollout plan, organizations may discover that their AI initiatives aren't delivering results. They may find that the insights gleaned aren't fit for purpose or don't address critical problems. Flipping this, if organizations go into an AI project lacking clarity on the expected business-aligned outcome, the insights themselves might be correct, but the organization might be unwilling to make the necessary operational or technical changes to act on them—a scenario that rings true for 42% of the IT leaders we surveyed. As such, **inclusive** and **collaborative** are words that must be central for AI initiatives.



Zeroing-in on infrastructure considerations

In the early stage of an AI journey, enterprises tend to leverage **off-the-shelf** models for building AI-driven applications and are often willing to pay for this ease of use. When looking to operationalize AI and unlock more customized outcomes, businesses will want to build their own AI models, seeking software and infrastructure designed to accelerate training and tuning. **In line with the identified early-stage AI position that surveyed organizations find themselves in, our results reveal that less than 20% of IT leaders are currently building their own AI models.**

Those who opted to build their own did so to keep private data safe (58% cited this as their top reason), or to gain a competitive advantage (true for 53%). Comparatively, the bulk of IT leaders are using free preset models (50%) or buying them (46%) and then tuning them with their company's own data. The top reasons behind this choice included the fact that it is an easier or more practical avenue to getting started with AI, that the existing models suit the use cases organizations had in mind, or that they were worried about their ability to secure their data on custom-built models.

The takeaway is that off-the-shelf models are the best starting point, but those wishing to advance their level of AI power or who find themselves exploring very specific use cases may end up moving to an own-build model in the future.

And what of model hosting? For the smaller, starting AI projects, organizations usually turn to public clouds for a fast on-ramp to AI. This is because public platforms offer a host of free tools and apps for basic functions, and the business can rely on the AI management expertise of the cloud provider. This often—though, not always—means that AI testing is easier and cheaper. **When an enterprise is ready to level up its AI projects to develop business-outcome models, they will need to move to hybrid clouds or on-prem infrastructure**, which are designed to handle larger, unique workloads and address the operationalizing needs of more complex AI projects.

Our survey findings show that **80% of IT leaders are currently running their AI models outside of the public cloud**, with 31% using private cloud and 33% using a hybrid solution. This is great news on the infrastructure front, but it does mean that the **management onus will fall to the organizations themselves**, many of whom aren't as in tune with nuanced AI requirements as they believe themselves to be. And this is particularly evident when looking at network infrastructure and compute needs for flexible support of traffic across the end-to-end AI lifecycle—which spans various stages of data management, model training, tuning, and inferencing.





Leadership should understand the needs for AI and not be surprised by all the infrastructure, guardrails, access tools, and everything. It's an ongoing need. There are a lot of ops to tune and a lot of configuration has to be done.”

– Manufacturing medical devices data scientist,
on infrastructure impact

To effectively run more complex AI models, such as those required for deep learning, and advanced generative AI, **businesses need to extend their cloud-native environment to include an AI-by-design approach.** But while 93% of IT leaders told us their network infrastructure is set up to support AI's particular traffic needs, less than half said they fully understand what the actual network needs of the AI lifecycle are. Similarly, while 84% state that their organization has flexibility in compute power, less than half said they are aware of the distinct compute needs across each stage of the AI lifecycle.

What this implies is that **enough focus might not be given to putting AI workloads where they are needed or are being used**—for training, tuning, or inferencing. Each phase in the AI lifecycle will require different infrastructure and locations, supporting the wisdom behind a hybrid end-to-end approach. What's required is an adaptive architecture of software and infrastructure to deliver significant scaling-up capabilities. This will enable businesses to harness vast amounts of both public and private data to fuel data analytics applications and AI model development from training to tuning. At the inference stage, for example, the architecture should be hybrid by design, allowing AI workloads to be placed where they matter, from edge to cloud.

A final note is for businesses to consider incorporating an open system from the start. Why? Because an open architecture gives the enterprise access to its tools of choice, **eliminating vendor lock-in.** A truly cloud-agnostic hybrid environment would also allow IT professionals to **use a single management interface to monitor any AI workloads, no matter where they are.**

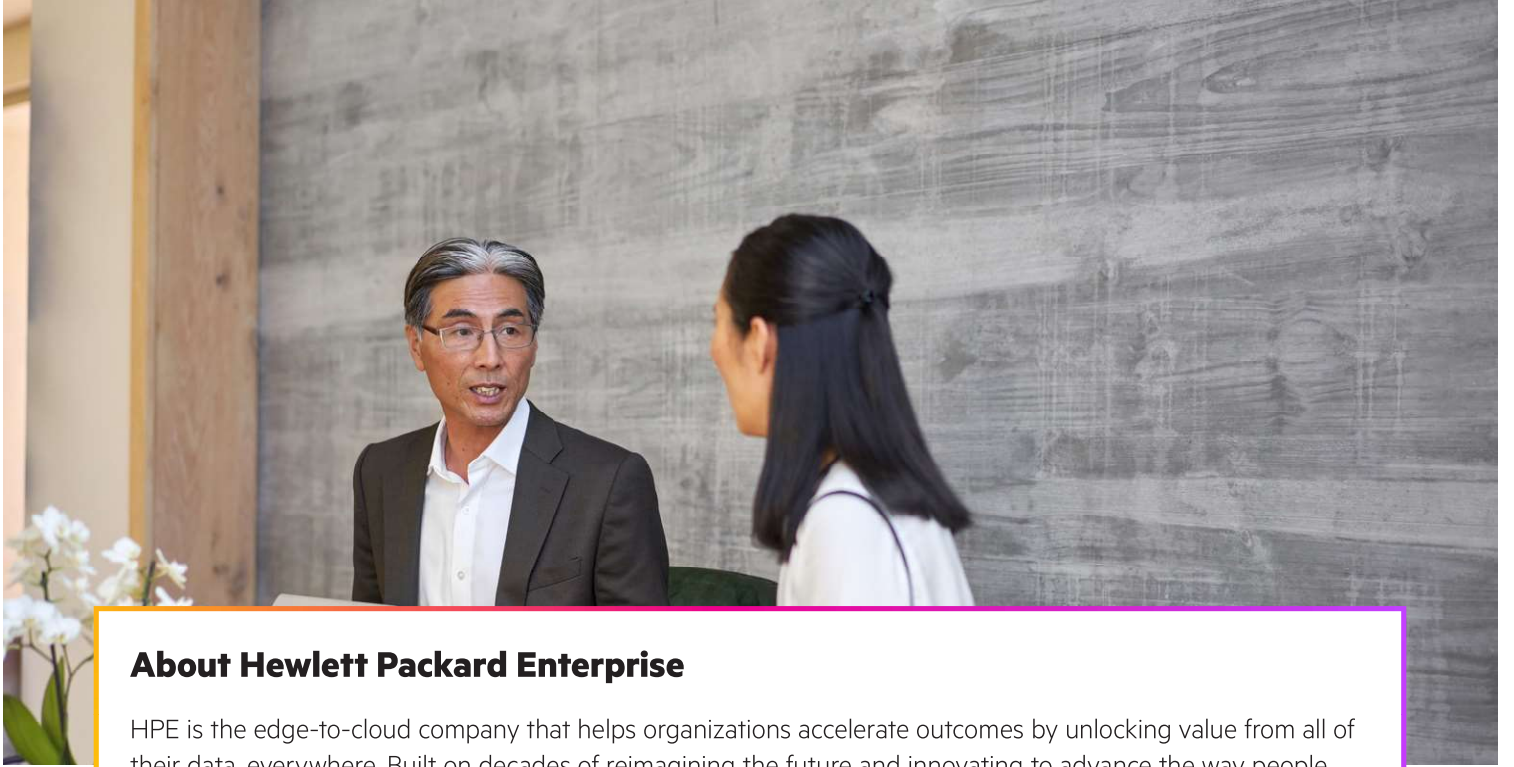
Is your business ready for the next step?

How do you know if your business is ready to progress its AI projects to the next level? **What is the tipping point for an organization to move from just consuming AI** (via off-shelf models they deploy) to producing, tuning, developing, and deploying their own in a very methodical, bespoke way across the business?

The tipping point for operationalizing AI is reached when a business has truly realized the importance of AI—and has ambitions to make it core to their operational model. At this stage, use cases have progressed beyond computer vision, calling for bespoke models that deliver a real competitive advantage in market—not simply internal business efficiency.

Cloud-native architectures have carried enterprises through digital transformation in the last decade. But those on the brink of that next-level AI tipping point are **embracing the move to an open, hybrid- and AI-by-design architecture** powered by supercomputing technology. And they're bringing together a considered team of cross-department influencers and decision-makers to make the most of this new era.





About Hewlett Packard Enterprise

HPE is the edge-to-cloud company that helps organizations accelerate outcomes by unlocking value from all of their data, everywhere. Built on decades of reimagining the future and innovating to advance the way people live and work, HPE delivers unique, open, and intelligent technology solutions, with a consistent experience across all clouds and edges, to help customers develop new business models, engage in new ways, and increase operational performance.

Learn more at

[HPE.com/ai](https://hpe.com/ai)

Visit [HPE GreenLake](#)



Chat now (sales)



**Hewlett Packard
Enterprise**

© Copyright 2024 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties for Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein.

a50010469ENW