



HOW TO MAKE AI INNOVATION EFFICIENT

CXOS THAT SELECT THE RIGHT CLOUD INFRASTRUCTURE FOR AI WILL BOOST CLOUD EFFICIENCY AND PRODUCTIVITY



SEPTEMBER 2023



POWERED BY
NUTANIX[™]

Organizations are dealing with two very opposing challenges. On the one hand, inflation is increasing operational costs and, for some nations and vertical markets, reducing sales. On the other hand, AI has the potential to revolutionize business processes and could cut those inflationary costs.

However, left unchecked, AI has the potential to increase cloud computing costs and, rather than solve inflationary pressures, exacerbate the impact on profit margins. Experts in cloud economics, business transformation, and technology optimization believe that a business-focused approach and the right AI-enabled infrastructure can deliver AI without breaking the bank.

In July 2023, the Producer Price Index, published by the [U.S. Bureau of Labor Statistics](#), revealed that data processing, hosting, and related services had increased in price by 2.8% year-on-

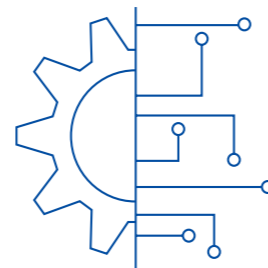
year and reported that these prices have been increasing since September 2022. The report states that generative AI demand and rising staff costs are the two main factors for this increase.

Despite these costs, industry analysts at Gartner forecast that CXO spending on enterprise cloud computing will grow by 21.7% in 2023, to a total of \$597.3 billion. Like the Bureau of Labor Statistics, Gartner says generative AI is behind these figures.

“Organizations today view cloud as a highly strategic platform for digital transformation, which is requiring cloud providers to offer more sophisticated capabilities as the competition for digital services heats up,” explains [Sid Nag](#), vice president analyst at Gartner. “Generative AI is supported by large language models, which require powerful and highly scalable computing capabilities to process data in real time.”



Organizations today view cloud as a highly strategic platform for digital transformation, which is requiring cloud providers to offer more sophisticated capabilities as the competition for digital services heats up”





AI demand

CXOs can expect their cloud costs to increase as demand for generative AI is on the increase. Business advisory firm McKinsey found in its latest survey that one-third of its respondents are using generative AI in at least one business function, and 40% plan to increase AI investments to keep up with advances in technology.

“Can you afford to be left behind?” asks Steen Dalgas, senior cloud economist at Nutanix. He goes on to say that despite the hype, business technology leaders fear becoming the laggard in their vertical market. The McKinsey research corroborates Dalgas’ observations.

Three quarters of survey respondents expect generative AI to significantly disrupt their vertical

market over the next three years, according to McKinsey. The survey also found that 28% of leadership boards have the use of generative AI on their agendas.

“There is lots of hype around AI, so it is really important to focus on what AI is good at doing today, Dalgas notes. “So drill down on the challenges that you have as an organization, and then can you deliver a business case for AI.”

“The business case will come down to revenue opportunities, lower costs, and decreased risk, as all business cases do,” Dalgas continues, explaining why CXOs need to bring their business acumen to AI adoption and enthusiasm.



There is lots of hype around AI, so it is really important to focus on what AI is good at doing today”

Dalgas says this is leading to two types of AI implementation. Organizations that need to develop and use sizable large language models (LLMs) are forming partnerships with hyper scalers that have identified the AI opportunity. Businesses partnering with Nutanix are modernizing their cloud infrastructure and using AI for new and powerful edge-computing business cases.

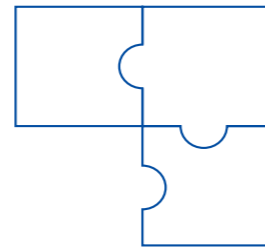
“We are working with a retailer to use AI at the edge for monitoring customer baskets in a self-service setting,” says Dalgas.

These are significant workloads that require the right infrastructure and cost structure to be strategically viable. A similar model is being applied in mining and manufacturing, where AI at the edge can monitor sensor data from machinery and alert the business of potential failures or maintenance requirements.

“The opportunities for AI at the edge are huge,” Dalgas points out. “And these are examples of organizations moving away from the hype and working on what AI can deliver today.”



To be successful, it requires a tight collaboration among the disciplines of governance, architecture, operations, product management, finance, and application development.”



Right AI, right cost

The demand for AI comes at a time when CXOs are closely monitoring cloud costs and, in some cases, moving away from public cloud providers and towards a hybrid cloud computing environment that includes their own IT estate. In 2022, technology market researchers at IDC found that 71% of the businesses it surveyed plan to move part of their workloads out of the public cloud.

“Cloud cost management is not just an operational concern,” Gartner states in an advisory note on cloud cost management. “To be successful, it requires a tight collaboration among the

disciplines of governance, architecture, operations, product management, finance, and application development.”

The [Gartner report](#) aligns with Dalgas’ observations and states that key performance indicators (KPI) for business have to be centralized to manage cloud costs in public, private and hybrid environments. It enables organizations to assess the business impact of cost growth and optimization. Driving costs down as a principle must not be done at the expense of being unable to fully support the business goals.”





Managing risk

Alongside the excitement towards AI, there is fear, from employees and customers. AI brings a new level of risk to organizations. To date, there has been a well-founded focus on the risks of information fabrication, often dubbed AI hallucinations. The ability to create deep fakes, increased cybersecurity risk, data security, and copyright infringements are among the potential threats from AI.

As Gartner's [Avivah Litan](#) says in her blog, "AI developers must urgently work with policymakers, including new regulatory authorities that may emerge, to establish policies and practices for generative AI oversight and risk management."

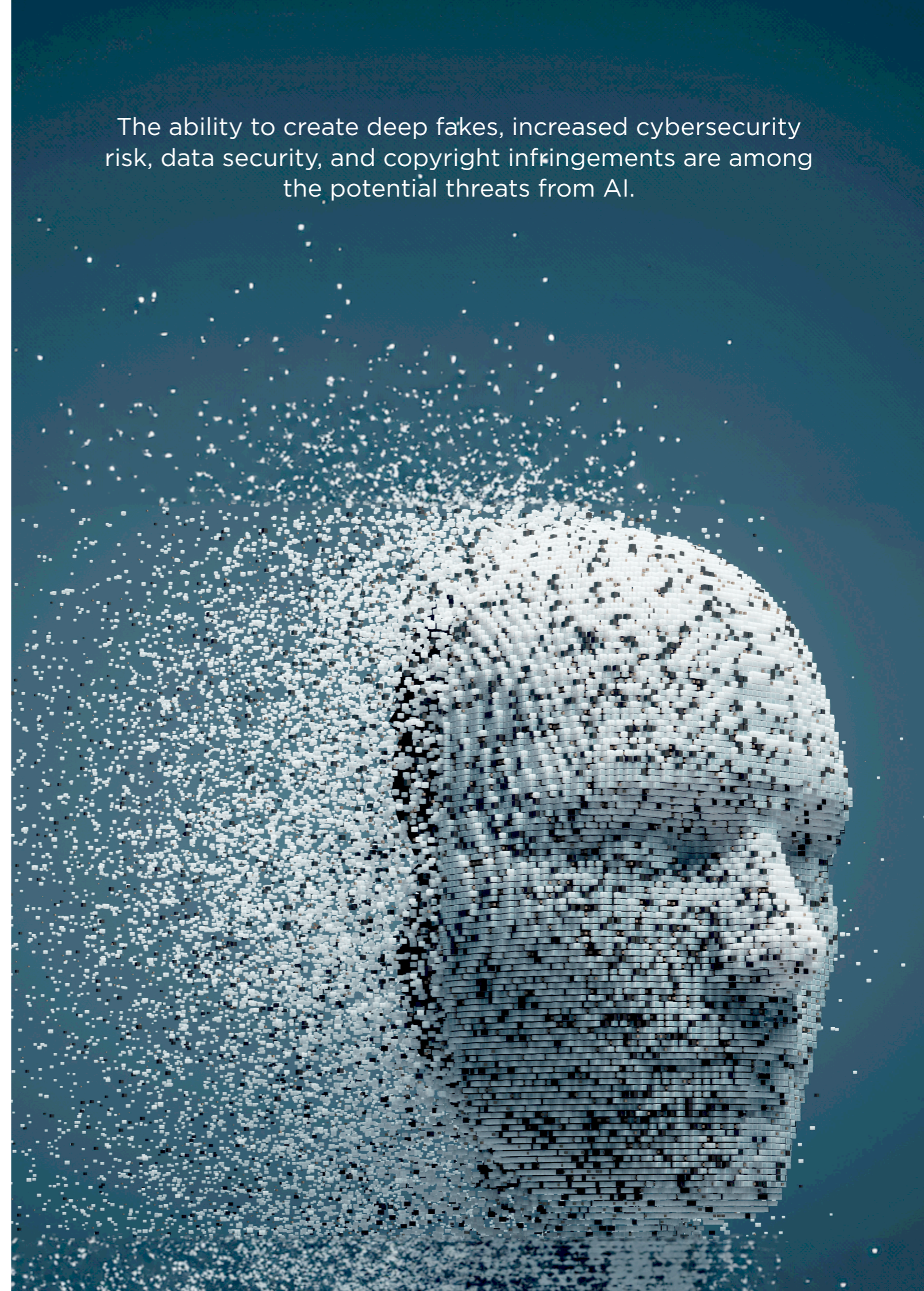
The cost control risks of AI have yet to be addressed, and these pose a significant risk to

organizations. But a hybrid cloud computing estate tailored to AI can enable CXOs to effectively manage and promote the use of AI. With governance of the infrastructure in place, organizations will be able to manage AI costs and contribute to protecting the business from the inherent risks of AI.

"You cannot afford to get this wrong as you'll be front page news," Dalgas notes.

Neither economic pressures nor the demand to use AI can be prioritized above the other. CXOs will and have to deal with both. To become AI-enabled and efficient, organizations need to maintain control of their data and applications and ensure that their

The ability to create deep fakes, increased cybersecurity risk, data security, and copyright infringements are among the potential threats from AI.





KEEP UP TO SPEED
WITH THE LATEST CONTENT

[NUTANIX.COM/CXO](https://www.nutanix.com/cxo)

POWERED BY

NUTANIX[™]