

IT optimization

Get AI ready from the ground up



Contents

Introduction: Optimization
in the era of AI

[03 →](#)

Understanding
affordability

[05 →](#)

Leveraging existing technologies,
overcoming tech debt

[12 →](#)

Using consumption-based
approaches

[22 →](#)

Enabling multicloud
management and
technology alignment

[30 →](#)

Next steps:
Ready to learn more?

[40 →](#)

Introduction

Optimization in the era of AI

Technology seasons may change,
but the need for an optimized IT
infrastructure stays evergreen.
So, how can IT leaders overcome
optimization challenges in the AI era?

Same challenges, different font

What do technology leaders dream of? Autonomous IT? Zero-latency networks? Sentient algorithms? Possibly. But it's far more likely their preoccupations revolve around optimizing IT and ensuring it aligns with business goals—as they always have. [Technology seasons may change, but the need for an optimized IT infrastructure stays evergreen.](#)

85% of CIOs are expected to leverage organizational changes to harness AI, automation and analytics, driving agile, insight-driven digital businesses by 2028.¹

AI is the latest in a long line of disruptive technology innovations IT leaders have had to contend with. And in this season of generative AI (gen AI), CIOs and CTOs find themselves grappling with many of the same IT challenges they've always grappled with. Once again, they're dealing with diverse IT environments, high costs, skills gaps, slow technology adoption, data security, data management and other issues. But this time the complexity and risk inherent to gen AI is added to the mix.

And the complexity is dialed up today—because running gen AI production architectures means operating multiple models, databases and applications across different workflows and platforms. In fact, 49% of leaders highly involved in AI report their organizations struggle to estimate and demonstrate the value of AI.²

With workloads, resources and data spanning multiple cloud estates, on-premises environments and edge locations, IT optimization for AI can be daunting, especially if one isn't sure how or where to begin. To help CIOs and other IT leaders optimize their IT estates for AI the correct way, we looked at *four* focus areas and answered questions they didn't even know they should be asking.

What we found as we explored these areas—affordability, legacy technology and technical debt, consumption models, and multicloud management and technology alignment—is a common need for intentional, consistent architectural decisions.

Let's dive deep into these IT focus areas and understand the importance of an intentional approach to technology adoption and management.

01

Understanding affordability

How can you afford AI? AI affordability is a complex matter, influenced by multiple factors. So, the answer to that question is equally complex. The key is balance.

USD 371.6 billion

is the expected IT spend on gen AI-related investments from 2024 to 2027.³

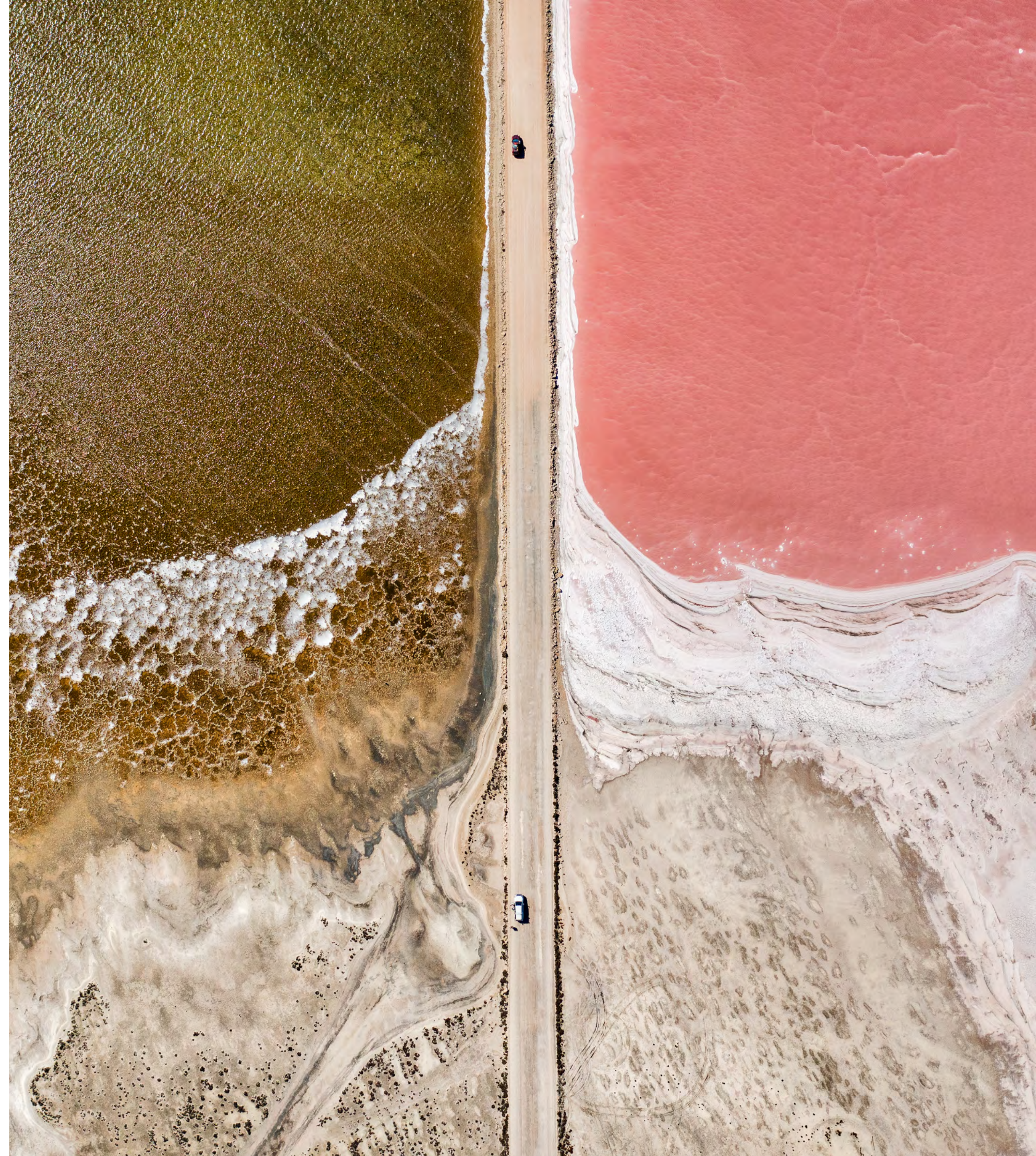
According to IDC, gen AI is triggering hyper-expansion of AI spending, with organizations expected to spend USD 371.6 billion on products and services to implement gen AI from 2024 to 2027. More than USD 150 billion of that estimate will be spent in 2027 alone.³

The world is all-in on AI and organizations are showing no signs of reining in their AI ambitions, but one question most likely plagues you: *How can I afford AI?*

Perhaps, new investments for AI aren't part of your budget nor do you know a way to use your current IT investments to support AI workloads. So, how can you afford AI?

AI affordability is a complex matter that's influenced by several factors, including initial costs, ongoing costs, skills, compute requirements, data availability and more—the answer may not be simple. AI is fundamentally transforming the way our world works. So, whether you can afford it or not, it's a must-have technology if you wish to drive the next wave of innovation and unlock new value in your organization.

The key is balance. Balance between ambitions and costs, between new investments and the current ones, and between going at it on your own and relying on trusted experts.



“CIOs don’t have to tackle this challenge alone; **the trick is to start small and go from there.** Don’t try to boil the ocean.”

Joe Cropper ✓

Distinguished Engineer,
Power as a Service & Hybrid Cloud
IBM Master Inventor
IBM

Adopting AI on a budget

Like most IT leaders, you’re expected to tackle seemingly unlimited to-do lists with decidedly limited IT budgets. These days, AI adoption is chief among your priorities, but investing in AI means a large capital outlay that has no room in your stretched-thin budgets.

How do you ensure that these tight budgets don’t hinder your AI journey? Here’s how: start with small and focused use cases. Starting small has many benefits, but the most notable one is that it allows you to achieve better outcomes with less risk.

An important point to remember is that AI isn’t a technology that you can install one day and hope to see immediate benefits the next day. You need to be realistic and clear-sighted about your goals and have a three-to-five-year vision of what your AI journey will be.

Pro tip: Seek the help of industry leaders, such as IBM, that have a strong point of view and a proven track record to help identify the right use cases for your organization.

Drilling down with management tools

How accurate do you feel your costs are? To make AI more affordable and take the necessary modernization and investment decisions, you need to have a handle on what you're spending—and where. In many cases, you may have a high-level view of your IT spend, but what's the view like at the workload or application level?

You can't optimize what you don't understand. A comprehensive view into your IT operations, consumption and spend will help you identify and target not just overspending and overprovisioning but also inefficiencies. Here's where automation and IT management solutions—including ITOps, AIOps and FinOps tools—come in handy.

These solutions help you have full oversight of your technology assets throughout their operational lifecycle—from inception to retirement.

They help provide a centralized view of your IT operations, deliver real-time insights on costs and resource utilization from across your organization, and automate tasks to optimize performance and security. When you understand your IT environment better, you manage your costs better.

Pro tip: Seek emerging management tools you can use to increasingly capture the value of AI, not just the costs.

Betting on hybrid cloud

You want to get the most from your existing infrastructure—that's the only cost-effective path to AI you see ahead of you. An intentionally designed hybrid cloud enables you to do that. It gives you the choice and flexibility to modernize your infrastructure at your own pace—so you can pivot quickly if or when you need to.

While hybrid cloud isn't a new concept to most organizations, an IT foundation that's merely hybrid by default won't deliver the expected benefits. AI and other emerging technologies require an IT foundation that's intentionally designed to support their implementation and use.

With a hybrid by design approach, you can use the public cloud to train your AI models—and lessen the investment burden on your on-premises infrastructure. At the same time, you can use your on-premises infrastructure for inferencing to meet any data sovereignty or privacy regulations that you need to comply with.

This approach also enables you to have seamless access to data across your IT estate. Since data quality can have a direct impact on AI costs, easy access and management of data is crucial.

Pro tip: Understand your current cloud and IT environment. Don't think of cloud as a location—rethink and reframe your cloud strategy to encompass a hybridized, connected IT.

“In today’s era of generative AI,

your data and your models are going

to be two of your most important assets.

However, finding those assets in a complex IT environment sometimes can feel a lot like playing a game of hide ~~Ⓢ~~ and Ⓢ seek.”

Questions to consider

01

Will the use case you've selected introduce AI into back-office processes to help you increase efficiency within key functions, HR for example?

02

Will it greatly streamline your ability to handle queries without needing the number of reps or personnel you typically need to field those questions or take care of those tasks?

03

What would be the potential return on investment if those productivity realizations were gained?

04

To address the skills gap, should you hire the right people externally or look at reskilling your current workforce?

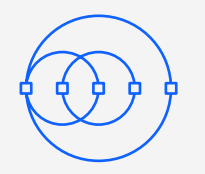
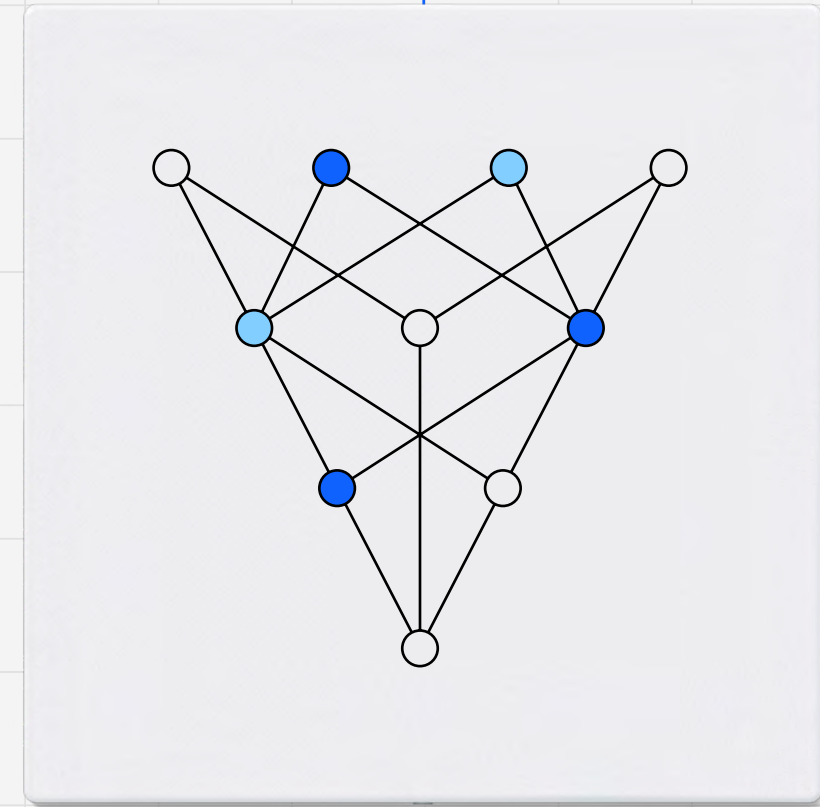
05

How will you make sure you've got the right quality of data?

06

When it comes to the IT infrastructure, how will you make sure to use IT resources properly?

The bottom line



Give your AI projects time to yield returns; don't expect immediate results.

Start small, with focused use cases that are more likely to be successful.

Use the right tools to stay on top of the costs.

Take advantage of hybrid IT to limit up-front investments and ease data management.

02

Leveraging existing technologies,
overcoming
tech debt

The road to AI is not easy. When traditional IT, tech debt and other IT challenges threaten to stall the journey, how do IT leaders keep moving forward?

Only 16% of surveyed executives say they're very confident their cloud and data capabilities are fully ready to support gen AI investments in 2024.⁴

While 27% of executives say they're unsure of readiness.⁴

Yes, AI has the potential to transform your organization, but any change most likely starts with you transforming parts of your infrastructure architectures.

Because when you run compute-heavy and data-intensive AI workloads on traditional IT infrastructure, you may be setting yourself up for failure. These workloads typically have wildly varying processor, memory and power requirements that most traditional IT and cloud infrastructures are simply not designed to meet. Then there's the matter of tech debt.

As your organization has grown over the years, you may have left tasks undone or favored work-arounds—perhaps prioritizing speed over quality—and let the tech debt pile up. Now, it's time to pay up and it's going to be costly.

How do you unravel this Gordian knot* of IT challenges and advance your AI ambitions?

Managing modernization

Traditional IT systems and applications have earned a reputation of being one of the biggest inhibitors to transformation and innovation—*harsh, but fair*. From interoperability to scalability and security to data management and governance, challenges just keep bubbling up when you try to integrate a new technology into your traditional data center or IT environment.

But your “legacy” systems and applications are critical to the business—whether it's your storage devices housing core data, your virtual machine farm running core transaction processing, your database footprint, your mainframe footprint or more.

So, as you begin to optimize IT for AI, do you replace the old technology or figure out a way to use it with newer investments? Your decision can have implications on not just infrastructure readiness but also affordability.

IBM Global Product Executive Peter Brey says most organizations opt for a mix of both. For example, you can migrate hardware and software components to your new environment and take advantage of technologies such as containerization and hyperconverged systems. Alternatively, you can keep what you have today, and start surrounding it with new capabilities such as automation or start enabling your applications to be API invocable.

Another factor to consider is that sometimes workloads perform better with on-premises, purpose-built architectures. When you try to move them or replicate them in the cloud, they *just* don't deliver the expected value. But with hybrid technologies, you can extend your custom architectures to the cloud to manage workload placement and ensure fit for purpose regardless of location.

You've invested millions of dollars in your IT infrastructure. Modernization doesn't mean you have to throw it all away and start over. Make use of your current infrastructure and thoughtfully and strategically design it to stand up AI adoption and innovation. *Incremental modernization*, if you will.

Pro tip: Reconsider your rip and replace plans. Find ways to make the most of your existing infrastructure with a hybrid approach, slowly but intentionally adding new technologies to support your organization's short-term and longer-term IT needs.

“People have this perception that when they have to deal with legacy systems, innovation is going to be delayed and there's going to be slower time to market to bring out new capabilities. But as IBM Infrastructure CTO Hillery Hunter says, **it's not legacy, it's actually legendary,** and making sure people understand the value of that is important.”

Joe Cropper ✓
Distinguished Engineer,
Power as a Service & Hybrid Cloud
IBM Master Inventor
IBM

Tackling tech debt

55% of surveyed organizations report technical debt as an obstacle to achieving business goals.⁴

The wheel of technology keeps on turning, and your tech debt keeps on growing. After decades spent chasing the *next best thing* in technology, your organization is now shackled with a tech debt that may be preventing you from adopting the next best thing in technology today—AI. It's clear, you need to break this cycle of hasty technology decisions and expensive business outcomes.

On the other hand, is all tech debt bad? Chuck Smith, VP of Transformation Business Architecture at IBM, says, “You have to have tech debt to create tech results. But how do you minimize the debt? How do you make tech decisions that give you the flexibility to move to something new, grow in a different way or not limit your choices from a business perspective in a year or two?” You could say that, like financial debt, tech debt can be good or bad.

The key is to ensure you have the right processes and skills to manage the debt and stop it from turning into a liability for your organization.

As you get ready to introduce AI workloads into your IT environment, identify areas of significant complexity and take steps to resolve challenges through integration, automation, modernization and monitoring. And before you make any sizeable AI investments, consider how the benefit of solving your immediate business needs stacks up against the pain of paying future IT management expenses.

Pro tip: Though it can be tempting to just keep adding new, accelerated functionalities to hide your tech debt, take a breath and take stock of your existing investments to identify and remediate issues in high-impact areas.

Shifting to a hybrid infrastructure

70% of enterprises are expected to form strategic ties to cloud providers for gen AI platforms, developer tools and infrastructure by 2025.⁵

For many organizations, cloud is an easy choice for AI—getting the agility, scale and cost-effectiveness of cloud with on-premises infrastructure, that’s easier said than done. Yet, IDC states that by 2025, 50% of enterprises will break with the public cloud migration trend and primarily deploy high-performance computing (HPC) on premises, at a managed services provider or at a colocation provider.⁵

When you consider the two aspects to AI—model training and inferencing—you realize that AI infrastructure needs aren’t *one size fits all*. The more compute-intensive operations happen during training, and you may want to leverage the flexibility and scalability of cloud for such operations. At the same time, running training workloads on premises may be more cost-effective in the long run. On-premises infrastructure also gives you better control of your data, so you can handle any security and digital sovereignty issues you may face.

And while cloud and on-premises infrastructure both may be ideal for inferencing workloads, last-mile inferencing is steadily gaining traction and much of the inferencing in the near future will happen at the edge.

Once again, the benefit of following a hybrid cloud approach comes to the fore. With a hybrid IT infrastructure, you can deploy your AI workloads in a public cloud, on premises or at the edge, depending on their compute, governance and security requirements.

Pro tip: Define clear objectives and evaluate your workloads to plan for and select the right inferencing and training locations.



“The compute-intensive and data-intensive needs of many AI and next-generation workloads demand flexible, hybrid and multicloud infrastructure resources that can scale and adapt as the needs of the workloads evolve.”⁶

Mary Johnston Turner ✓
Research Vice President,
Future of Digital Infrastructure
IDC

Questions to consider

01

What are you trying to achieve? Are your modernization efforts intentional, thoughtful and designed to drive business outcomes?

02

Have you considered “blocking and tackling” your modernization efforts to improve optimization?

03

How do you intend to provide an integrated system for gen AI, machine learning or inferencing?

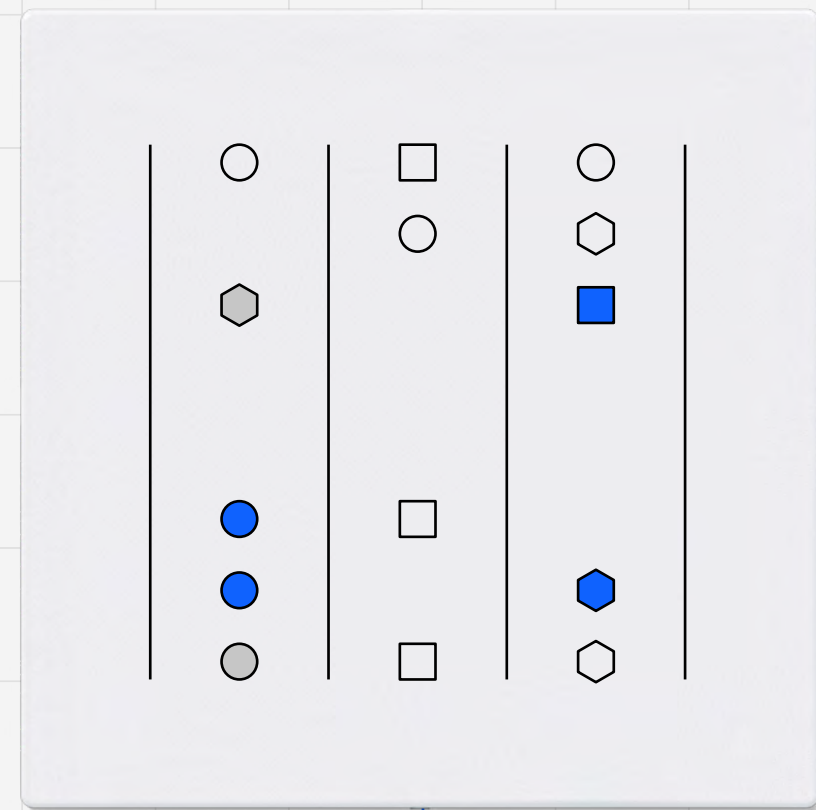
04

Have you considered where to start your modernization efforts? Did you know you can keep your existing platform as your crown jewels, then start wrapping new technologies around it?

05

How will you ensure your current optimization measures align with your future strategies?

The bottom line



Find ways to strategically manage tech debt. But also understand that sometimes tech debt is necessary to create tech results.

Make use of your existing infrastructure. You don't have to start from scratch and let decades' worth of IT investments go to waste.

Consider the nature of your AI workloads before you make deployment decisions.

Always focus on business and strategic outcomes.

Stepping up on skills

Not surprisingly, skills—or the lack thereof—is one of the biggest challenges organizations have to address in this season of AI. Firstly, to ensure more fruitful optimization efforts. But more importantly, to realize any degree of success in their AI adoption.

However, addressing the AI skills gap isn't just about reskilling your staff on AI-specific technologies. It also involves managing any strategy or infrastructure skills gap you may have.

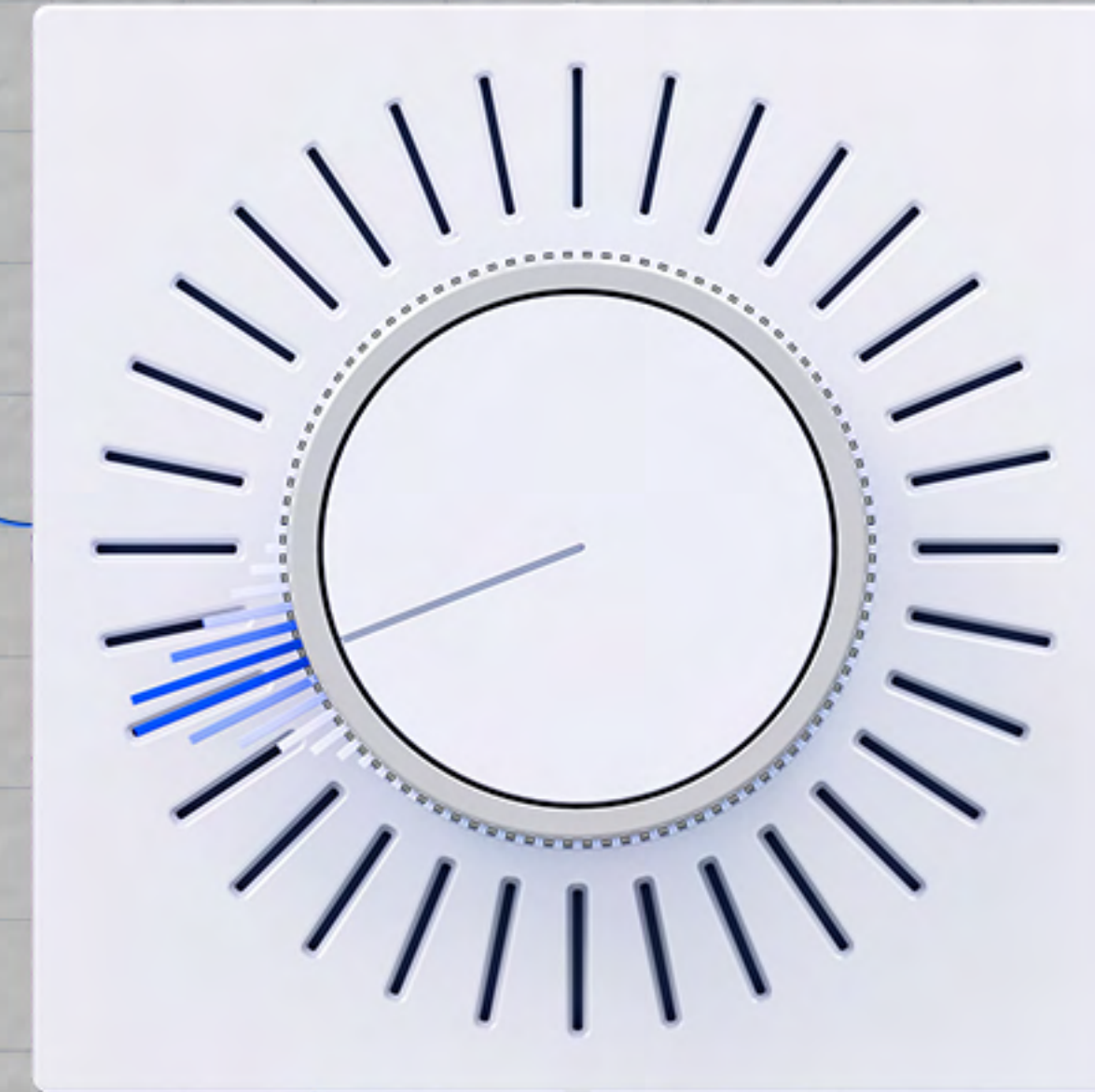


72%

of responding IT leaders say AI skills gap must be addressed urgently.⁷

“In many cases, organizations lack the resources to manage the traditional systems in their IT environment. Whether their staff have retired or have exited the company, they are struggling to get the skills and capability to either modernize the systems or just maintain them.”

Chuck Smith ✓
VP, Transformation Business
Architecture, Infrastructure
IBM



AI needs everything from AI strategy expertise to Kubernetes and Red Hat® OpenShift® skills. When it comes to successfully implementing AI, the data scientists who work with the data as well as the engineers who build the tools and solutions the data scientists consume—both are equally important. So, whether your plan is to hire externally or retrain your existing staff, remember, bridging the AI skills gap requires a threefold approach.

03

Using consumption- based approaches

AI affordability and acceleration goals may often seem at odds with each other, but with consumption-based operating models, IT leaders can make the twain meet.

46% of surveyed customers are expected to procure dedicated gen AI infrastructure solutions from OEMs in flexible consumption, XaaS contracts.⁸

With AI, overspending—on infrastructure, skills or energy—is always a clear and present possibility.

Fortunately, with anything-as-a-service (XaaS) models, you have options. You have the option to get started with your AI workloads at a lower CapEx. You have the option to control and manage IT costs with consumption-based billing. You even have the option to consume your on-premises resources in an as-a-service model.

In fact, consumption-based operating models can be the linchpin that holds your AI affordability and acceleration goals together. They give you the necessary impetus to begin your journey—and a safety net in the event the journey turns out to be riskier than you expected.

But whether consumption-based models are right for you depends on your AI priorities.

“There are cloud-based consumption models that customers are pursuing as an easy way to start. But many customers have data privacy or regulatory compliance requirements that mandate their data stays on premises. Those customers are working to figure out ways to do inferencing on prem with the infrastructure they have or with new infrastructure. So, I think the consumption model can be cloud or on prem, but generally, it’s guided by preference, any type of regulatory or compliance requirements, or cost,” says Rebecca Gott, CTO, Power Platform, IBM.

Advancing with XaaS models

You already know, cost is a big inhibitor to AI adoption. But with an XaaS model, you can make a case for AI while still staying on budget. An XaaS model largely eliminates the need for any up-front investments by providing flexible subscription options that help reduce risk. It gives you the flexibility to match your usage with your investment—on a predictable monthly stream—while providing visibility into what you consume. It also gives you the ability to consolidate the costs and convert them into OpEx.

Additionally, in a XaaS model, whether you access cloud or on-premises resources, the burden of managing the infrastructure is on the service providers. The only burden on you is managing workload deployment based on your chosen use cases and the specific requirements they may have.

“You might decide this consumption-based model makes sense in the long run. Or, depending on your growth and other factors, you might decide that it makes more sense to do a capital investment and gain more security. That’s where people have that choice. Public cloud and other consumption-based models open doors to people who would otherwise not have been able to get into the AI space,” says Joe Cropper, IBM Distinguished Engineer and Master Inventor.

This way, XaaS may also hold the answers to some of your modernization problems. It provides you flexible access to resources on an affordable monthly basis, which, in turn, allows you to grow and transform your footprint over time. Did someone say incremental modernization?

Pro tip: Take advantage of consumption models not just to make AI deployment more affordable and manageable but also to reduce the burden on your staff.



How XaaS helps advance your AI ambitions

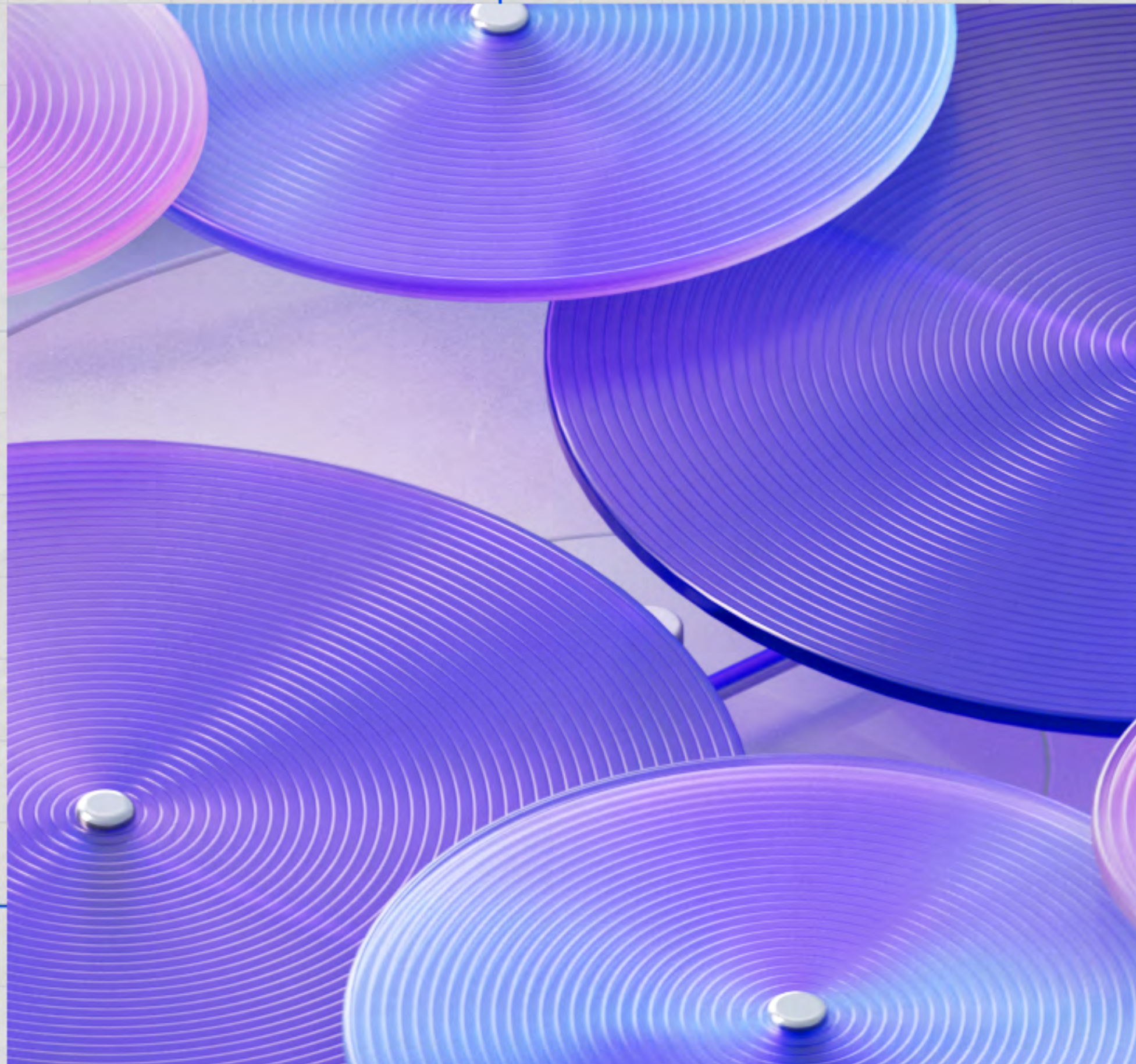
AI is accelerating the shift from traditional consumption models to XaaS models. Here's why. Organizations need an IT consumption model that can:

- Help manage unplanned growth requirements
- Provide ready access to capacity, services and IT skills
- Offer flexible and predictable cost structures
- Help reduce overprovisioning and overspending



78%

of surveyed respondents state that XaaS offerings are a key part of their future strategy.⁹





“XaaS models can help organizations embrace this new growth era by providing customers access to capacity, services, IT talent, and cost metrics to **reduce overprovisioning and overspending** as they integrate gen AI into their long-term planning.”

Future of Digital Infrastructure:
GenAI Growth and the Case for XaaS Adoption,
IDC Research, September 2023



Balancing long-term costs and benefits

There isn't one correct consumption model. You exercise the choice and flexibility offered to you and make decisions that are right for your organization. But you know at the heart of it all is the question: *CapEx or OpEx, which is better?*

CapEx is fundamentally more complex. But more importantly, in the eyes of your CFO, it's more expensive. It's a traditional funding model that's fast losing ground to cloud-based OpEx and on-premises OpEx. But these OpEx models, while great for agility and flexibility, can also drive up costs in the long term.

Is it a lose-lose situation, then?
Not necessarily.

Joe Cropper says, "There is a direct correlation between organizational benefit and organizational cost. In other words, if you're consuming resources, it's because you're getting some organizational benefit out of it."

Like with everything in life, there's a trade-off with CapEx and OpEx too. But an intentional hybrid cloud approach helps you take advantage of both the models, so you can balance investments and expenses, delay capital expenditures and limit operational costs.

Pro tip: With your fully consumption or as-a-service models, learn to balance the long-term costs with all the additional value that you'll gain. When your budget opens up and you decide to invest in capital, be mindful of the fact that what you gain in control and security, you may have to pay for in agility and scalability.

Questions to consider

01

Why do you think consumption-based models are right for you?

02

Does your operating model match the use cases you've identified?

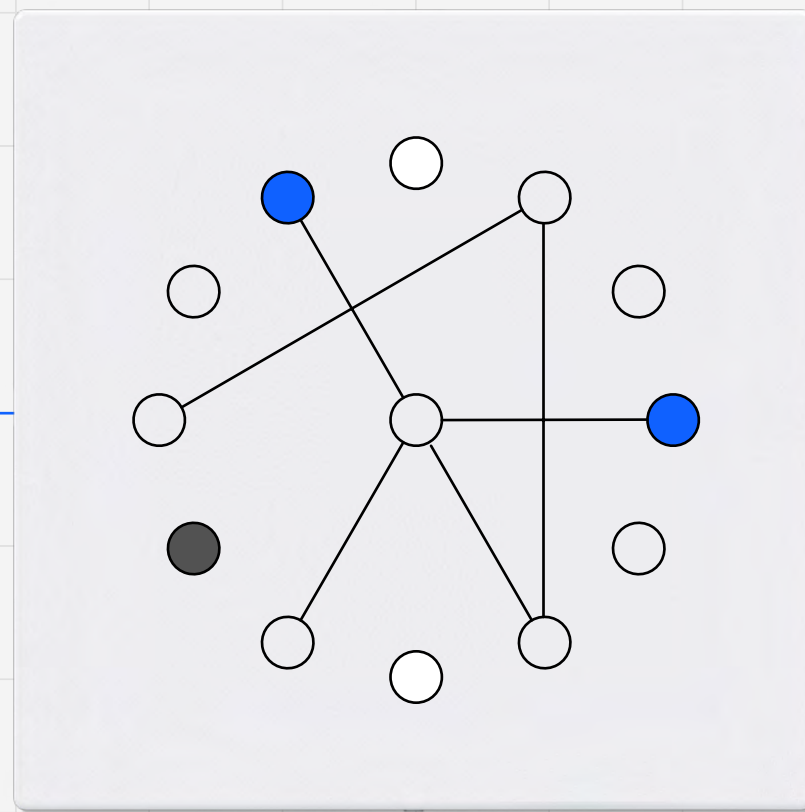
03

Have you done a detailed ROI comparison of CapEx and OpEx models?

04

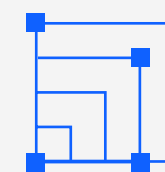
What will your primary focus be, training or inferencing?

The bottom line



Use XaaS solutions in your cloud or on-premises infrastructure in a flexible consumption-based model.

Choose an operating model that meets the specific requirements of your organization and your use cases.



04

Enabling multicloud management and technology alignment

How can technology leaders get their IT estate in order and put their AI strategies into action?
Hint: It involves aligning IT and business and making intentional technology decisions.

93% of surveyed organizations deploy more than one public cloud provider, and large organizations have a median of three public hyperscalers in active use today.¹⁰

When IT and business are not aligned, poor decisions are made. Such as uncoordinated investing, which leads to not only overspending, sprawl and poor ROI, but also the *dreaded* data silos—which in turn create data governance, management and integration issues.

So, when IT and business are not aligned, you may find that your data and applications are all over the place: across multiple clouds, on premises or at the edge. And you may find you have to temper your AI expectations and confront the complexity of integrating and managing these disparate data sources.

Given the obstacles in your way, how do you get your IT estate in order and put your AI strategies into action?

Curbing the cloud sprawl

You know what they say about too much of a good thing? Nothing illustrates the phenomenon of cloud and tech sprawl quite as well as that adage. Picture this: you've always been eager to take advantage of the obvious—and ubiquitous—benefits of cloud and other next-generation technologies. But over the years, you've invested in one cloud too many, one IT tool too many. And now, you have an acute case of the tech bloat and must reckon with the decisions you've made. *How did we get here?*

Peter Brey says, “For most organizations, sprawl is an issue that has built over time and become a situation the IT can’t control. Then you layer on technological change that keeps coming at them and the tech debt issue. It’s easy to understand how these organizations have gotten to where they are.” Simply put, sprawl is the result of nearsighted—or nonexistent—cloud and IT strategies that prioritize short-term gains. These strategies age poorly and are often unequal to support the needs of new and emerging technologies.

Now picture adding AI to this mix. If you can’t stop the sprawl and address the many attendant challenges—security and compliance problems, the proliferation of data silos, integration issues, high operational costs, and even skills gaps—it can be tough. Before you can effectively—and responsibly—deploy AI, you need to repair the fractured data flow between your disparate systems and ensure your AI algorithms have easy access to the data they need for training.

At the same time, you need to tighten your security, compliance and governance policies. And when you finally succeed, you need to make sure your AI plans don’t drive more unchecked technology consumption and further add to the sprawl.

“There are many applications that inferencing can be used for and we’ll see more sprawl there. On the training side, because of the expense and limited availability of GPU resources, there’ll be a natural damper,” says Brey. “But in the next three to five years, we’re going to see a shift where even the training side of the equation is going to become more ubiquitous than it is today—it’s going to be a bit of the Wild West. Then the question will be, how do you get a handle on the sprawl? How do you manage it?”

How *do* you keep any future sprawl under control? By developing more enduring strategies and taking a more thoughtful approach to technology adoption.

It could mean keeping a cap on the number of cloud providers, choosing applications or workloads with buy-in from both IT and business, leveraging automation and implementing monitoring, or setting up strong governance structures.

Pro tip: Use management tools to keep a handle on your operations, environment and architectures today and understand them. If you’re successful, you’ll be able to achieve the necessary optimization and efficiencies and be AI ready faster.

Building a hybrid cloud by design

Only 16% of surveyed executives say they're very confident that their cloud and data capabilities are fully ready to support gen AI investments in 2024.⁴

There's one way you can reconcile the often conflicting needs of IT and business, integrate cloud and business goals, minimize the impact of any misalignment, and pave the way for AI. It's by developing a durable cloud strategy and building an intentionally designed hybrid IT estate. Here's why.

An IT estate that's hybrid by design—and not by default—is the result of *thoughtful* and coherent technology decisions that are made when cloud and technology strategies align with the business strategies. It's connected and cohesive despite being a heterogeneous mix of multiclouds, on-premises systems and edge devices. And it can help consolidate your tech stack and standardize data management practices, breaking down silos to build data pipelines for your AI engine.

“The ability to have seamless access to data across the cloud or on premises—or even at the edge—and to manage all that is important for many organizations,” says Chuck Smith. “Because it all goes back to the data. Where is that data residing? Where is it being acted upon? Where is it being generated? Where is it being modified and placed? So, having a hybrid landscape is becoming more and more important for AI.”



“By breaking down workload and data silos using open standards and APIs, cloud-native container orchestration, unified control planes, and automated data logistics capabilities, enterprises can increase the value of their existing data assets and take better advantage of infrastructure innovation regardless of where it is physically deployed.”¹¹

Mary Johnston Turner ✓
Research, Vice President,
Future of Digital Infrastructure
IDC

Technology leaders understand intuitively that, even at an architectural level, data is super important—whether it is a dataset for modeling and training or a dataset for inferencing. “Because there’s also this fear of the data being outside their four walls and outside of their control, or the data being noncompliant. All those things are top of mind for most organizations,” says Chuck Smith.

To take the first step toward building an IT estate that’s hybrid by design, prepare to carry out a clear-eyed assessment of your current IT estate. You may discover that your systems are too disparate and disconnected to support the fast and free flow of data necessary for collaboration and innovation.

Or you may uncover that shadow IT has encroached your technology landscape, resulting in cost and effort duplication as well as increased risk exposure. Once you have a complete understanding of the state of your IT, unify your technology assets by applying a tested, codified architectural framework to gain the integration, speed and security your organization needs.

With the *hybrid by design* approach, you can get your IT optimized and ready for AI today—and lay the foundation for enduring advantage to achieve business outcomes in the future.

Pro tip: Turn to experts who can assess your environment, help establish clear goals and expectations, and implement best practices around hybrid cloud and AI.

Questions to consider

01

How aligned are your IT
and business strategies?

02

Do you have the right tools and technologies
to manage cloud and tech sprawl?

03

How are you managing your multicloud
and on-premises environments?

04

Is AI going to increase overprovisioning
and sprawl in your IT environment?

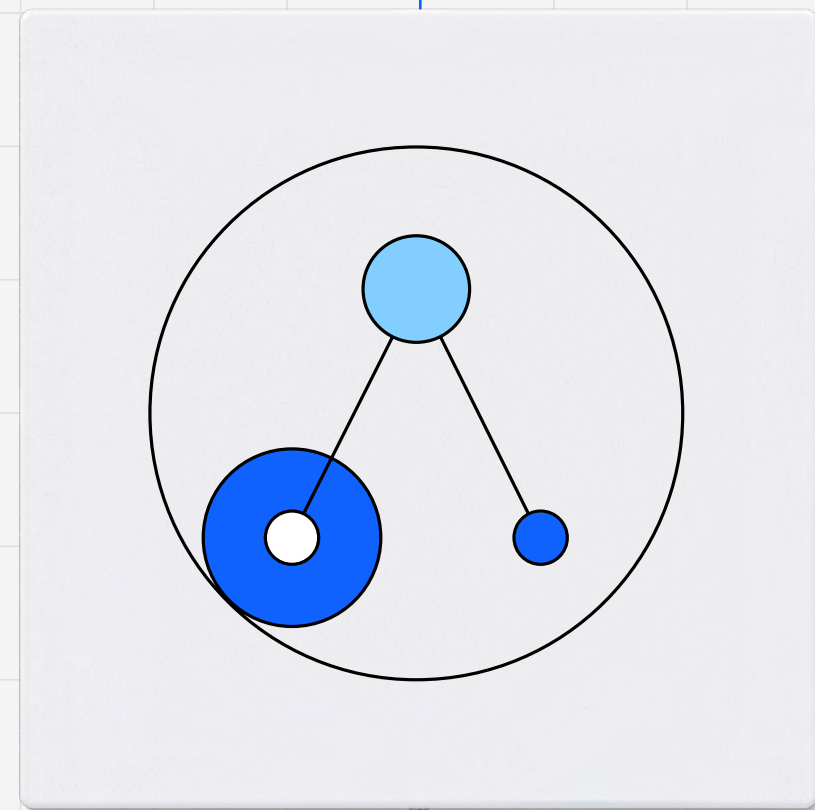
05

Are your cloud budgets spiraling
out of control?

06

Do you have a hybrid IT estate? Was it built organically
over time or planned according to key business strategies?

The bottom line



To successfully implement AI, make sure IT and business are in sync and take an intentional hybrid by design approach. Aligning IT and cloud strategy with business strategy can help ensure that the AI initiatives are driven by business goals and supported by business leaders.

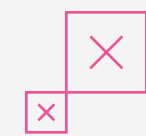
Track usage across the IT estate with monitoring and optimization tools to prevent uncontrolled spending and identify architectural pain points that may need remediation.

Understand that hybrid by design delivers durable advantage—most organizations will realize benefits over time, rather than all at once.

The hybrid by default problem



Poor alignment between IT and business, which has resulted in a motley mix of IT architectures and operations



Underutilized investments and growing tech debt



Cloud and tool sprawl, which can lead to security and compliance problems, proliferation of data silos, integration issues and high costs

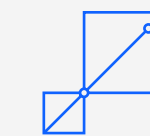


Fractured data flow between disparate systems

The hybrid by design solution



Applies a tested framework to help unify your technology estate and align your IT and business priorities



Provides choice and flexibility to modernize your infrastructure at your own pace



Enables you to consolidate your disparate IT systems into a single, integrated estate that's easier to secure and manage



Helps break down silos to build data pipelines for your AI engine

“An intentionally designed hybrid cloud architecture allows you to facilitate rapid scalability by tuning and deploying AI wherever your data resides. This is a central step towards more efficient, cost-effective value creation.”



Explore IBM's AI-ready infrastructure solutions and hybrid cloud expertise, and design a hybrid IT estate for AI.

Next steps

Ready to learn more?

Over the course of four chapters, we highlighted the different challenges you may face while optimizing your IT estate for AI, providing insights and advice from a panel of IBM technology experts as well as industry analysts.

[Now, put the advice into action and optimize your IT for AI the correct way.](#)

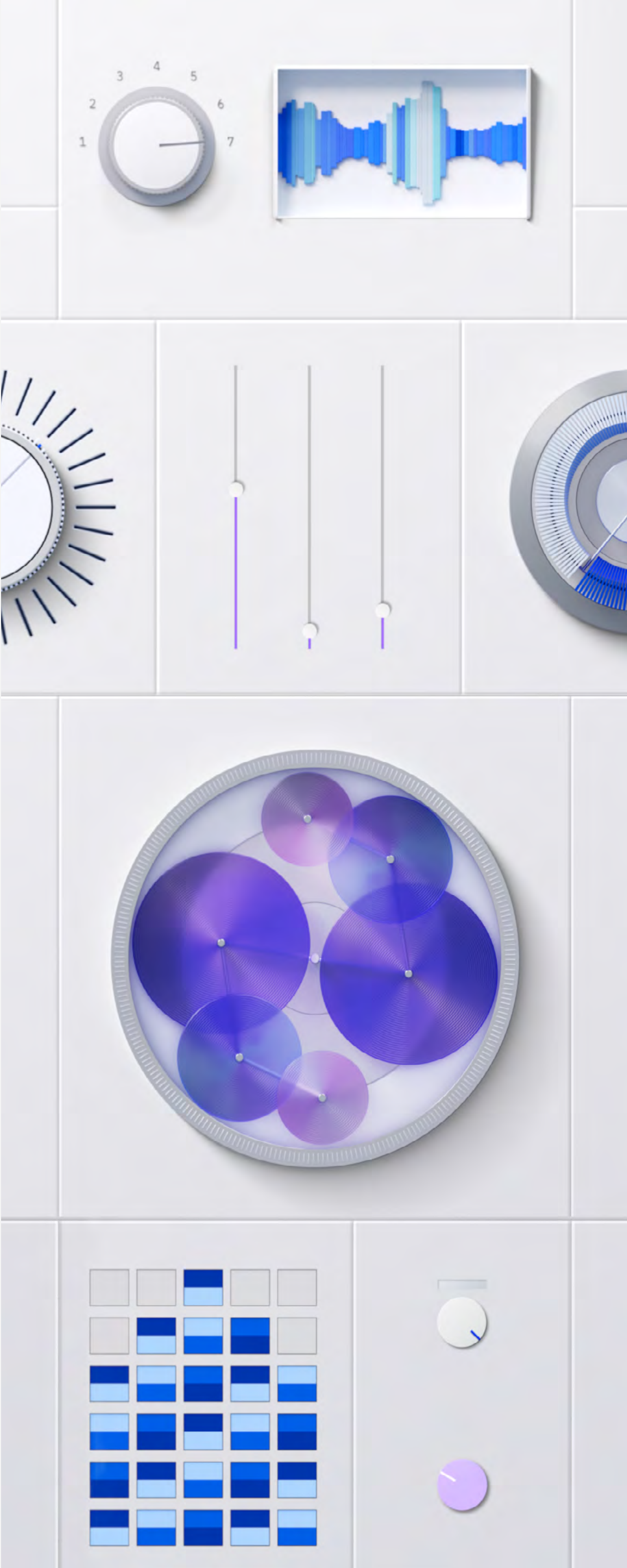
Take advantage of IBM's AI-ready infrastructure solutions and unparalleled hybrid cloud expertise to design a hybrid IT estate that can help you realize your AI outcomes.

AI-ready infrastructure

Build your AI-ready hybrid cloud by design with [IT infrastructure solutions](#) from IBM that can help you modernize and scale on premises and in the cloud at your own pace. They're security-rich, scalable, open, indemnified and designed to support the highly dynamic and performance-intensive nature of AI workloads.

AI-ready hybrid cloud expertise

As much as AI adoption and infrastructure modernization have found a prominent place in the technology zeitgeist, many organizations still need a guiding hand to help them navigate the uncertainties surrounding evolving technologies. Let our consulting and industry experts help you figure out not just your AI goals but also the roadmap for your IT optimization journey. [IBM Consulting®](#) has worked with some of the largest companies in the world and helped them deploy hybrid cloud and AI technologies. Our experts can help you [build open, secured, hybrid cloud architectures](#) that align with your business strategies and implement AI across your environments to optimize critical workflows and business processes.



Ready to learn more?

IBM's portfolio of hybrid cloud solutions

[IBM Cloud®](#)

IBM's enterprise cloud platform that is AI-ready, security-rich and hybrid by design

[IBM Z®](#)

A family of modern mainframe infrastructure that helps delivers the security, resiliency, performance and scalability your business needs

[IBM® Power®](#)

A family of servers based on IBM Power processors that are designed to help users respond faster to business demands, protect data from core to cloud, and streamline insights and automation

[IBM Fusion](#)

A hyperconverged infrastructure (HCI) system that provides a fully integrated, turnkey platform for running and maintaining on-premises Red Hat OpenShift applications

[IBM Apptio™ Cloudability](#)

A cloud financial management platform that ingests, normalizes and structures billing data across all major cloud providers to deliver single-pane-of-glass cost visibility

[IBM Concert®](#)

An AI-powered tool that helps simplify and optimize app management and technology operations with gen AI-driven insights

[IBM Instana® Observability](#)

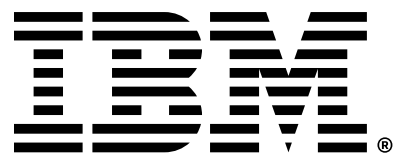
An observability platform that's designed to unlock cloud-native application performance with AI-driven automation

[IBM Turbonomic®](#)

A software platform designed to optimize the performance and cost of IT infrastructure, including public, private and hybrid cloud environments

[IBM Consulting](#)

The only major global consultancy within a technology company, led by experts that follow principles grounded in open innovation, collaboration and trust—working with clients and partners—to design, build and operate high-performing businesses



1. [IDC FutureScape: Worldwide CIO Agenda 2024 Predictions](#), Doc # US51294523, IDC Research, October 2023.
2. [Gartner Survey Finds Generative AI Is Now the Most Frequently Deployed AI Solution in Organizations](#), Gartner, 7 May 2024.
3. [Worldwide Core IT Spending for GenAI Forecast, 2023–2027: GenAI Is Triggering Hyper-Expansion of AI Spending](#), Doc # US51539723, IDC Research, December 2023.
4. [From chaos to cash: How hybrid by design creates business value](#), IBM Institute for Business Value, 7 May 2024.
5. [IDC FutureScape: Worldwide Cloud 2024 Predictions](#), Doc # US51294723, IDC Research, October 2023.
6. [IDC PeerScape: Practices to Align Digital Infrastructure Spending for AI and Next-Generation Workloads with Business Priorities](#), Doc # US52103524, IDC Research, May 2024.
7. [AI Skills Gap Must be Addressed Urgently, say 72% of IT leaders in Red Hat Survey](#), Red Hat press release, 5 December 2023.
8. [Preference for Flexible Consumption of Dedicated GenAI Infrastructure](#), Doc # US52044624, May 2024.
9. [Digital Transformation in the Era of AI Drives Interest in XaaS](#), IDC, December 2023
10. [IDC MarketScape: Worldwide Multicloud and Hybrid Cloud Management with Automation 2024 Vendor Assessment](#), Doc # US52084624, IDC Research, May 2024.
11. [Hybrid and Multicloud Interoperability Building Blocks for Future Enterprise Innovation](#), Doc # US52076924, IDC Research, May 2024.

*The Gordian knot is a metaphor for complex or difficult problems. The term has its origins in ancient Greek mythology; today it's commonly used to describe challenges that require bold or unconventional approaches to be solved.

© Copyright IBM Corporation 2024

IBM, the IBM logo, IBM Apptio, IBM Cloud, IBM Concert, IBM Consulting, IBM Instana, IBM Z, Power, and Turbonomic are trademarks or registered trademarks of International Business Machines Corporation, in the United States and/or other countries. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on ibm.com/legal/copytrade.

Red Hat and OpenShift are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.

IBM products are warranted according to the terms and conditions of the agreements under which they are provided. No IT system or product should be considered completely secure, and no single product, service or security measure can be completely effective in preventing improper use or access. IBM does not warrant that any systems, products or services are immune from, or will make your enterprise immune from, the malicious or illegal conduct of any party.

The client is responsible for ensuring compliance with all applicable laws and regulations. IBM does not provide legal advice nor represent or warrant that its services or products will ensure that the client is in compliance with any law or regulation.