



So you want to adopt Generative AI

Key questions and best practices
for how to get started

Clear path. Open road.



If you're like **73%**¹ of executives, you're actively looking to leverage Generative AI (GenAI) to advance your business goals.

Smart move.

With GenAI, you can supplement and enhance your current efforts with:

-  Better strategic decision making
-  Higher productivity
-  Lower costs
-  Accelerated innovation
-  Faster time to market

...and the list goes on.

¹ PwC Pulse Survey, June 11, 2024, indicating that 73% of executives plan to use GenAI to make changes to their company's business model.



But, like many of your peers, while you may know that you need GenAI to remain competitive, you might not know where to start.

That's where this ebook comes in.

Here, you'll find information to help you move forward.



Tips and best
practices



Key questions
to consider

Let's jump in.



Step 1: Take stock

As with any big undertaking, it's imperative to start with an honest stocktaking exercise. You need to understand what you have, what you need, and what the appetite for change really is.

Here are a four key steps to help you strategize

Evaluate data quality and availability.

Access to high-quality, relevant data is essential for AI. Data quality is critical for training accurate AI models, so it's important to assess the completeness, accuracy, and relevance of your data.

1

2

Assess technological infrastructure.

Determine if your current infrastructure can support AI workloads. This includes ensuring security and evaluating the availability of high-performance computing resources, storage solutions, automation, and network capabilities. If necessary, explore what kind of infrastructure you'll need to fill any gaps.

3

Identify where skills are needed.

Evaluate the availability of AI expertise within your organization. Assess the current skill set and identify where training or specialized skills may be needed.

4

Review strategic alignment.

Make sure that your AI initiatives align with your business goals and strategies. AI projects should support your organization's overall strategic objectives and deliver measurable business value.



Step 2: Set up for success

Your next step is to set yourself and your team up for success.

That begins with a basic task: identifying use cases—the opportunities within your organization that AI can address. A well-defined use case will guide your AI adoption efforts and focus resources.

Revenue Generation

- Chatbots
- Campaign and content marketing
- Developer assistants
- Guided selling
- Drive product innovation

Risk Management

- Sentiment analysis
- Predict employee attrition
- Contract risk assessment
- Fraud detection
- AI-assisted security operations

Cost Optimization

- Automated AI support
- Knowledgebase search & summarization
- Doc summarization
- AI-optimized logistics
- Augmented product R&D



Here are a few key questions to help you get started

- ✓ Do you have significant bottlenecks or high-cost areas that could be automated or augmented by GenAI?
- ✓ Are there customer-facing processes or interactions that could be improved by GenAI's personalized content or enhanced responsiveness?
- ✓ Do you have specific types of data (e.g., customer feedback) that could be analyzed or transformed by GenAI to generate new product features, market insights, or revenue streams?



At this stage, you're also going to need to identify who will work on your AI efforts.

Form a dedicated AI application team, including cross-functional members such as developers, domain experts, data scientists, and IT specialists, to lead the initiative.

Remember, culture is key. GenAI moves quickly, so you'll need a group that is collaborative, agile, and adaptable. You'll also need to define clear roles and responsibilities so that the team can build, test, and deploy seamlessly.

Key roles and responsibilities²

Business leaders represent the people who use or are affected by the solution.

AI specialists tune, maintain, and update GenAI models.

Data scientists preprocess and provide correct, unbiased training data for models.

Ethics and compliance officers ensure that GenAI initiatives comply with regulations.

IT operations specialists integrate solutions with existing infrastructure and enforce security policies.

Development teams and communities need to be involved at the beginning to collaborate, create, share, and improve open source tools, frameworks, and best practices for AI adoption. This will ensure AI uses are connected to business value.

² Kearney. "[Standing up tiger teams to tackle generative AI complexity](#)," 15 Nov. 2023.

Here are a few key questions to help you get started

- ✓ What roles will be needed to build, deploy, and manage the use case(s) that you have defined?
- ✓ Should your AI team be centralized or distributed across different departments, and what are the advantages and disadvantages of each approach?
- ✓ How will you scale the team as your GenAI initiatives expand and evolve?



Step 3: Choose the right tools

Model Selection

With a task and team in place, it's time to get started with technology and specifics.

That begins with model selection.

Choose the appropriate AI model based on your identified use case. Whether it's an LLM for GenAI or a predictive model for data analysis, make sure the model's capabilities align with the goals of your use case. Consider factors like the model's complexity, scalability, and compatibility with your existing systems.

Here are a few key questions to help you get started

- ✓ What performance requirements and trade-offs (e.g., accuracy, speed, cost, scalability, security) are critical for the successful deployment of your use case(s)?
- ✓ Which model's architecture and training data can most effectively process your data, deliver the required results, and handle the complexity of the task at hand?
- ✓ What customization, fine-tuning, and ongoing support are required for the model to meet your ongoing business needs?



Infrastructure

You're also going to want to make sure that your infrastructure can handle the unique needs of GenAI workloads.

Specialized compute power like GPUs and other AI accelerators are the backbone of your GenAI efforts; make sure that they can handle the intense parallel processing demands of GenAI models. You'll need highly secure, high-performance data storage and retrieval systems so that you can minimize bottlenecks and maximize the efficiency of data-intensive GenAI training. And don't neglect your network: it must be robust and scalable, with high bandwidth and low latency to support distributed training and handle growing GenAI workloads.



Here are a few key questions to help you get started

- ✓ Is your infrastructure flexible and scalable enough to support future growth and innovation?
- ✓ How does the TCO of your GenAI infrastructure align with your budget and expected ROI for AI initiatives?
- ✓ What kind of security measures do you have in place, including data protection, disaster recovery, and compliance with relevant regulations, to minimize business risks?



Data

It's also important at this phase to consider what data you will use to train and enhance your models.

Consider synthetic data. This approach, using methods such as LLM teacher and LLM student, allows you to generate high-quality training data when real data is scarce or sensitive.

Synthetic data can help improve the model's robustness and performance without compromising privacy.

Here are a few key questions to help you get started

- ✓ Do you have specific data challenges where synthetic data can help you solve real, pressing business problems (such as data scarcity or sensitivity)?
- ✓ How will you ensure the regulatory compliance, security, and privacy of your synthetic data, especially if it's based on sensitive information?
- ✓ Have you accounted for any cost implications of generating and maintaining synthetic data?





Step 4: Continuously improve

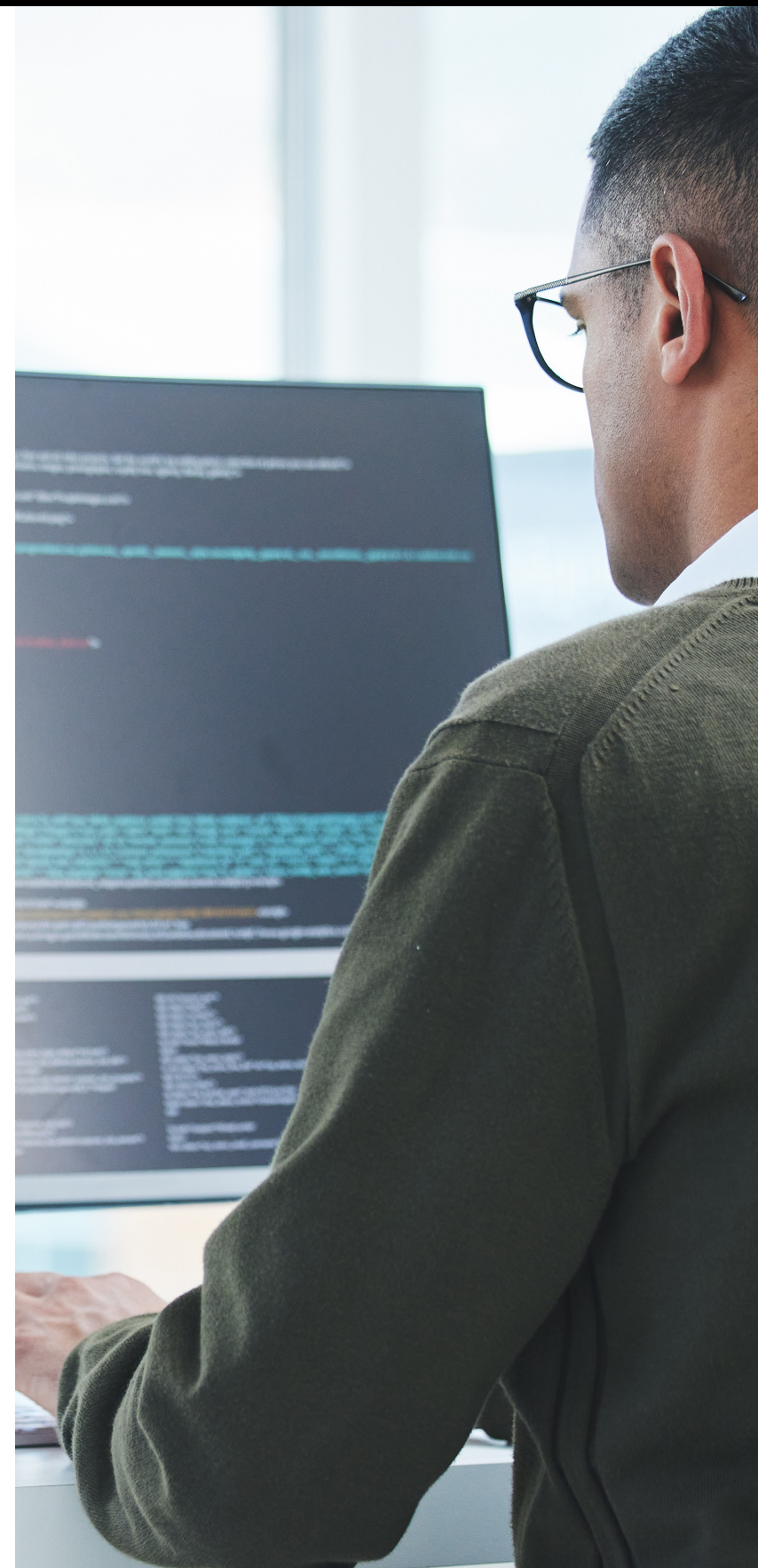
Testing and validation loops

As with every strategic effort, it's critical to set benchmarks for success and frequently measure against them.

Your next step is to set up testing and validation loops.

Define clear success criteria for your AI implementation, such as performance metrics, accuracy rates, or business objectives.

Establish testing and validation loops to continuously assess your model's effectiveness. Regular feedback from these loops will help fine tune the model and keep your AI journey on track.



Here are a few key questions to help you get started

- ✓ What specific metrics will define success, and how will you measure them? In other words, what does “good” look like?
- ✓ How will you mitigate bias and ensure fairness in your models?
- ✓ How will you ensure ongoing reliability, maintain performance, and detect problems before they impact your business?



Model tuning

Now, it's time to make sure that your model works hard for you.

Through model tuning, customize your selected model using the real or synthetic data you've identified.

This tuning process involves feeding the model relevant data to improve its accuracy and relevance to the use case(s) that you've defined. Fine tuning makes sure that the model adapts to your organization's unique context and needs.

Here are a few key questions to help you get started

- ✓ How will the tuning strategies your team employs directly impact the key performance metrics you've defined and drive meaningful results for your organization?
- ✓ What are the trade-offs between model performance, computational cost, and deployment time? Are you sacrificing speed for the sake of accuracy or vice versa?
- ✓ Are you ensuring that the tuning process doesn't introduce bias, and that the model will be able to handle real-world variations and cyberattacks?



Drift monitoring

Drift monitoring provides general, content-based monitoring. Rather than structured configuration management, drift monitoring tracks changes to content on the local file system.

Drift monitoring can help you detect and address any deviations or declines in the model's accuracy over time. Continuous monitoring ensures that the model remains effective and relevant as conditions change.



Here are a few key questions to help you get started

- ✓ What thresholds and alerts will trigger intervention when model drift exceeds acceptable limits, and how do those thresholds align with your business risk tolerance?
- ✓ How will you ensure that your systems are able to detect performance, concept, and data drift, and how will you differentiate between them to inform response strategies?
- ✓ How will you investigate and address drift and ensure that the processes you establish maintain or improve the fairness and robustness of the model?



Your path to an AI-enabled future

AI is a game-changing technology, and getting started can seem daunting.

It doesn't have to be. Move forward with Cisco and Red Hat.

Together, we clear your path to AI with a secure, unified, enterprise-ready private cloud platform. Driven by consistent cloud-native management, we simplify IT complexity, accelerate AI innovation with efficient fine-tuning and optimized inference, and extend diverse edge operations securely. Our validated solutions reduce costs, enhance protection, and transform your hybrid cloud.



Accelerate time to market for AI-powered applications and unlock faster time to value.

Pre-validated, full-stack Cisco AI PODs ensure streamlined deployment and reduce integration complexities, while Red Hat OpenShift AI's MLOps capabilities and stable OpenShift Kubernetes foundation accelerate AI model fine-tuning and deployment from months to days.



Optimize GPU utilization and reduce operational overhead with purpose-built, high-performance AI infrastructure and efficient model serving.



Ensure secure and compliant AI operations with integrated security and governance across the fine-tuning and inference lifecycle, while maximizing throughput and minimizing latency for inference through the integrated OpenShift environment.

Free your teams
to innovate.

Exceed your customers'
expectations.

Prepare for future
growth and scale.

We've cleared your secure path
to AI adoption. The open road awaits.

[Learn more](#)



Clear path. Open road.

For more than a decade, Cisco and Red Hat have partnered to build highly secure, powerful solutions that accelerate digital transformation and IT modernization. It's what you need in today's hyper-competitive landscape: proven partners who will work with you to build AI-ready data centers, future-proofed workplaces, and digital resilience.

What does that look like?

- Flexible, simplified management of hybrid cloud infrastructure and cloud-native applications.
- Virtualization that doesn't break the bank...and positions you for future growth.
- Automation that frees your team to innovate.
- And AI capabilities that can give them wings.

We've cleared your secure path to digital transformation. The open road awaits.

