



ARTIFICIAL INTELLIGENCE:

Step-by-step guide for startup founders

A practical guide to help make
your AI vision a reality



INTRODUCTION

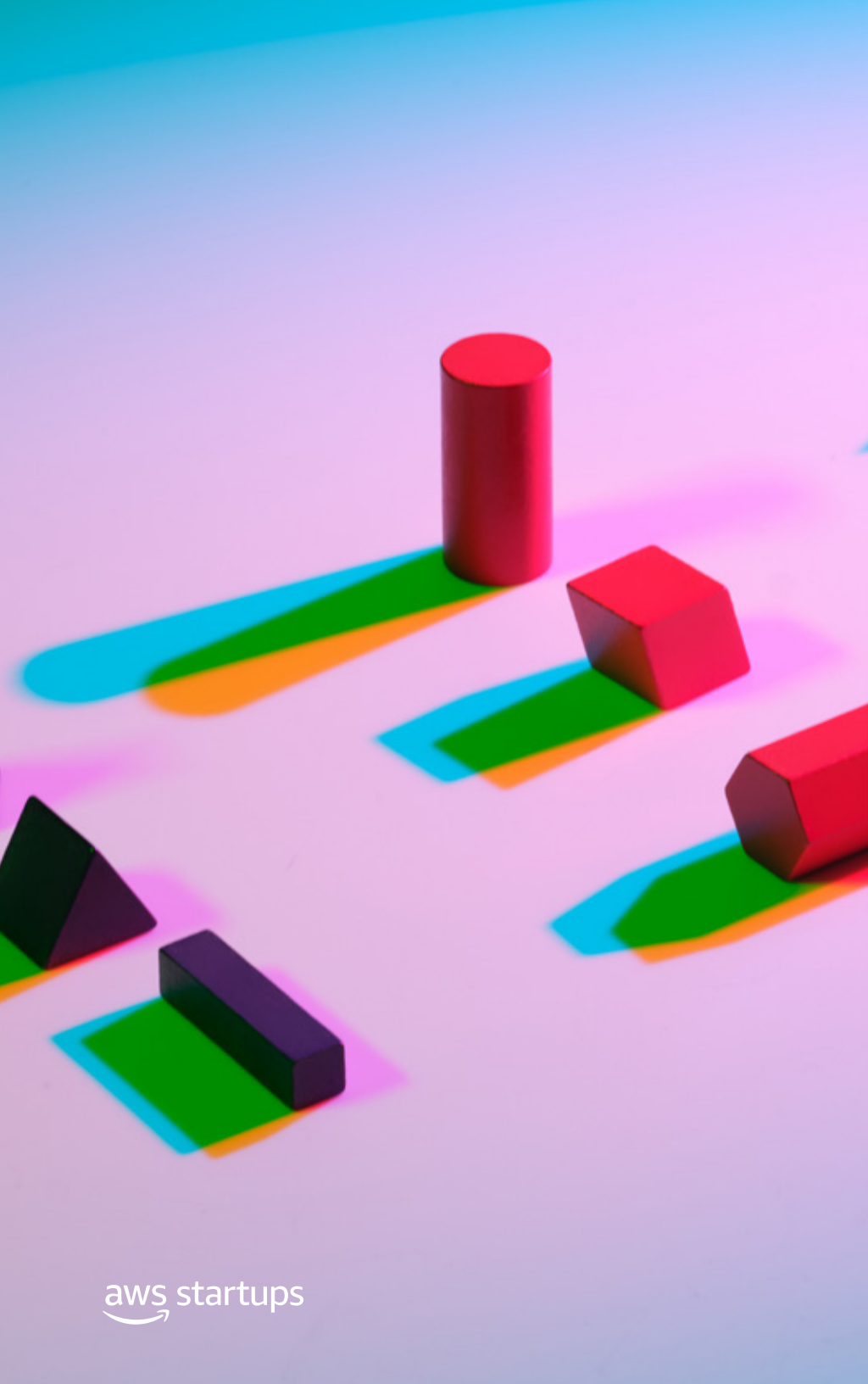
Forging ahead

When deployed with the right strategy, generative artificial intelligence (AI) has the potential to change the anatomy of work. By streamlining work activities that absorb 60–70 percent of employee time, startups can improve productivity down to the individual worker.¹ This can fast track product innovation, increase agility, streamline processes, boost revenue, and allow for better, faster decision making.

The lack of a single universally accepted playbook for generative AI success is keeping some founders on the sidelines, unsure of how to take the next (or even the first) step on the AI journey. By democratizing generative AI, Amazon Web Services (AWS) is helping startups at any stage overcome barriers to adoption and forge ahead with confidence. This eBook outlines a proven path—from taking the first step to measuring results—with insights from AWS best practices and its experience helping thousands of customers, including leading startups like yours, realize their own initiatives.

Generative AI could add
**\$2.6
TRILLION**
to
**\$4.4
TRILLION**
in economic benefits annually

The impact of generative AI on productivity could add trillions of dollars in value to the global economy, increasing the effect of all AI by 15%–40%¹

An abstract 3D graphic featuring several geometric shapes: a tall red cylinder, a red cube, a red hexagonal prism, a dark purple rectangular prism, and a dark purple pyramid. These shapes are set against a background with a vertical color gradient from light blue at the top to light purple at the bottom. Each shape casts a shadow on the surface below it, with the shadows having a multi-colored gradient of cyan, green, and orange.

What are artificial intelligence, machine learning, and generative AI?

AI is a way to describe any system that can replicate tasks that previously required human intelligence. Almost all AI systems are built using machine learning (ML). ML utilizes large amounts of data to create and validate decision logic, forming the basis of an AI model. The AI application then feeds input data into that model, and the model outputs human-like decisions. The rapid advancement of ML, massive proliferation of data, and easy availability of scalable compute capacity can help businesses accelerate how they use AI and, particularly, how they leverage generative AI.

Generative AI is a type of AI that can create new content and ideas, including conversations, stories, images, videos, and music. Like most AI, generative AI is powered by ML models—very large models that are pretrained on vast amounts of data and commonly referred to as foundation models (FMs).

Why generative AI?

Before diving into the steps of your generative AI journey, let's explore why any startup should begin its journey in the first place. According to Goldman Sachs, generative AI could drive a 7 percent (or almost \$7 trillion) increase in global GDP and lift productivity growth by 1.5 percentage points over a 10-year period.² Globally, we have reached an inflection point where most customer experiences and applications will be reinvented with generative AI.

Developing a strategy that includes generative AI initiatives is imperative to ensure successful business outcomes. Even after completing the steps outlined in this eBook, a business will need to regularly remind itself what it's working toward—staying focused on the precise business benefits that can be unlocked by fully leveraging generative AI technology.



Startups are already realizing the impact of:

Ideation and developing new products faster

Generative AI applications can automate manual aspects of product development, such as routine design, analysis, and writing tasks. It can also generate dozens (and, in some cases, hundreds or even thousands) of ideas from simple text prompts, all within seconds, allowing you to fast-track the entire innovation process from ideation to go-to-market.

Boosting employee productivity

Generative AI is being leveraged for its transformative value to help AWS customers reach new levels of productivity. The latest advancements in generative AI can be used to increase employee productivity. The technology can help you accelerate your team's performance to make the most of headcount and limited resources.

Improving customer experiences

Startup founders can take advantage of generative AI to improve customer engagement, increase personalization, and attract new users through deeper

experiences. AWS customers have improved their own customer experiences through the effective use of chatbots, virtual assistants, personalization, and content moderation.

Transforming content creation for greater creativity

Startups, like game-industry-enabler Scenario, are using generative AI to turbocharge real-world production processes across multiple types of creative content—including art, music, text, images, animations, and video. The AI-powered capabilities for automating writing, media design, and character modeling allow for unprecedented creative exploration and speed.

Driving outcomes for real, competitive advantage

Regardless of a startup's size, field, and domain, generative AI should be embedded into every business plan to stay relevant and competitive. AI21 Labs (AI21) built a midsize state-of-the-art large language model (LLM), deploying **Amazon SageMaker** to train the FM in under 20 days.³ This midsize model would preserve the quality of text generation, making it nearly the same as AI21's largest model at a much lower inference cost to AI21 and its customers.

Now that we've outlined the *why* of generative AI, it's time to explore the *how*

The next sections will demonstrate the steps in the AI journey using AWS best practices and those of AWS startup customers to exemplify the necessary changes that must take place in order to successfully implement, deploy, and scale AI solutions.

STEP 1

Make data your differentiator

Like with ML, data is an important piece of generative AI. While general LLMs can be used “out of the box” for some use cases, when you want to build generative AI applications that are unique to your business needs, your startup’s data is your strategic asset. Data is the difference between a generic generative AI application and one that truly knows your business and your customer in order to deliver a better, more differentiated experience.

While some companies will build and train their own LLMs with vast amounts of data, many more will use their organizational data to fine-tune FMs for their unique business needs or to add context to prompts through Retrieval Augmented Generation (RAG). For example, you can use data from sources such as your data lake, database, and data warehouse to create a chatbot that can provide technical support for your products or a model that provides marketing copy trained on your best-performing ads.

Because of this, success with generative AI requires relevant, high-quality data, which means that you need a strong data strategy in the cloud. According to **McKinsey Digital**, “...companies that have not yet found ways to effectively harmonize and provide ready access to their data will be unable to fine-tune generative AI to unlock more of its potentially transformative uses.”

The right data strategy for generative AI includes a comprehensive set of services to store and query data at scale, breaks down silos so you have ready access to all of your data to leverage for generative AI applications, and makes sure your data is secured and governed throughout the lifecycle of building generative AI applications.

“With tens of thousands of teams relying on Coda AI to not only finish tasks but make progress in their work, it’s important to bring the best of security, privacy, and AI quality to their enterprise data and user experience. Amazon provides secure storage in S3 all the way through serving AI responses with Bedrock, keeping our LLMs inside AWS, offering scalability and dependability. Coda AI can use this to create curated content and insights—and all of it stays within AWS’s secure platform.”

David Kossnick, Product Lead, Coda AI



How Easebuzz did it

Easebuzz provides a comprehensive online payment solution platform for individuals and businesses. With its expansion, it became essential to extract value from its data to provide data-driven features to customers. To accommodate the number of logs ingested daily, Easebuzz built a data lake on AWS.

Implementing a data lake reduced data storage expenses by 60 percent. Instead of paying an hourly rate, the company now only pays for the data it stores in **Amazon Simple Storage Service** (Amazon S3). The data lake delivers analytics services to customers, offering valuable insights into business performance and the ability to identify emerging trends. Customers can obtain a highly detailed report that contains three years of data in under three minutes. Report consumption has increased by 15 percent as customers are accessing on-demand daily dashboard reports 89 percent faster.

[Read the customer story ›](#)

“Our AWS data lake architecture has significantly added value to our platform. It raises our ability to derive actionable insights, optimize operations, and deliver a superior payment experience to our customers.”

Amit Kumar, CTO, Easebuzz

STEP 2

Finding the right business problem to address

One mistake startups often make in their AI journeys is to let their technical team work in silos to build models as proofs of concepts rather than solve real business problems. With no specific business problem to solve, your team will find it increasingly difficult to unlock the benefits of AI. This can stall or even stop progress on AI initiatives.

Here are some important questions you should ask before embarking on an AI journey:

1. Does it solve a real business problem?
2. Will the project benefit from AI?
3. Can it eventually be operationalized?

The [AWS AI Use Case Explorer](#) is a business outcome-centric search and navigation site that enables users to find the right AI use cases and discover relevant customer success stories.

In a successful AI journey, you should start by creating a dream team to address your specific business problem. This requires including both technical and domain experts within your team. While the technical experts will take on the brunt of model creation, they need the field knowledge of domain experts to define precise business challenges and identify the data most important to finding a solution.

Your team should also work through how to measure success. “Make sure you have very crisp and clear metrics as you embark on the machine learning journey,” Freshwater said. “Many times, your models are taking over for something existing, and you want to make sure that they’re actually better and that you can measure it.”

For more on measuring the success of AI initiatives, refer to [Step 5](#) in this eBook.

You can also leverage the [AWS Generative AI Innovation Center](#) to work backwards from business challenges and go step-by-step through the process of creating AI projects and initiatives.

STEP 3

Upskilling teams

In parallel with creating a comprehensive data strategy, today's startups must focus on arming their team with the right skills to succeed in the era of generative AI. However, startups across dozens of industries are growing increasingly aware of an expanding skills gap—the separation between technologies and the ability of internal technical and business specialists to take full advantage of them. It's an issue that should set off alarm bells in light of recent research from the World Economic Forum, which found that more than 75 percent of organizations plan to adopt AI in the next five years.⁴

To help individuals train for the AI jobs of the future, AWS released **on-demand skills training** to support those who want to understand, implement, and begin using generative AI. Amazon has also designed training courses specifically for developers who want to use **Amazon CodeWhisperer**.

While there's no one-size-fits-all solution to the AI skills gap, there are proven methods that can maximize the abilities of your team, reducing the need to make large investments in buying or borrowing pretrained expert talent.





How Slalom did it

As a purpose-led, global business and technology consulting company, Slalom helps customers build practical end-to-end solutions that drive meaningful impact. To continue delivering powerful solutions and drive business growth, Slalom looked to **AWS Training and Certification** and **AWS Certified Machine Learning – Specialty**, which validate expertise in building and deploying ML models on AWS.

Slalom's experience using the AWS Training program has advanced its consultants' professional development and growth, which has also driven sales. In the six months after one learning cohort concluded in 2022, 50 percent of participants drove more sales than they did in the six months before.

[Read the customer story ›](#)

“Working with AWS Training and Certification empowers our team members to learn new, valuable skills that can help them grow and advance their career journeys. We try to align the needs of our consultants with where our customers’ strategic objectives are headed.”

Jeff Pearson, Managing Director, Slalom

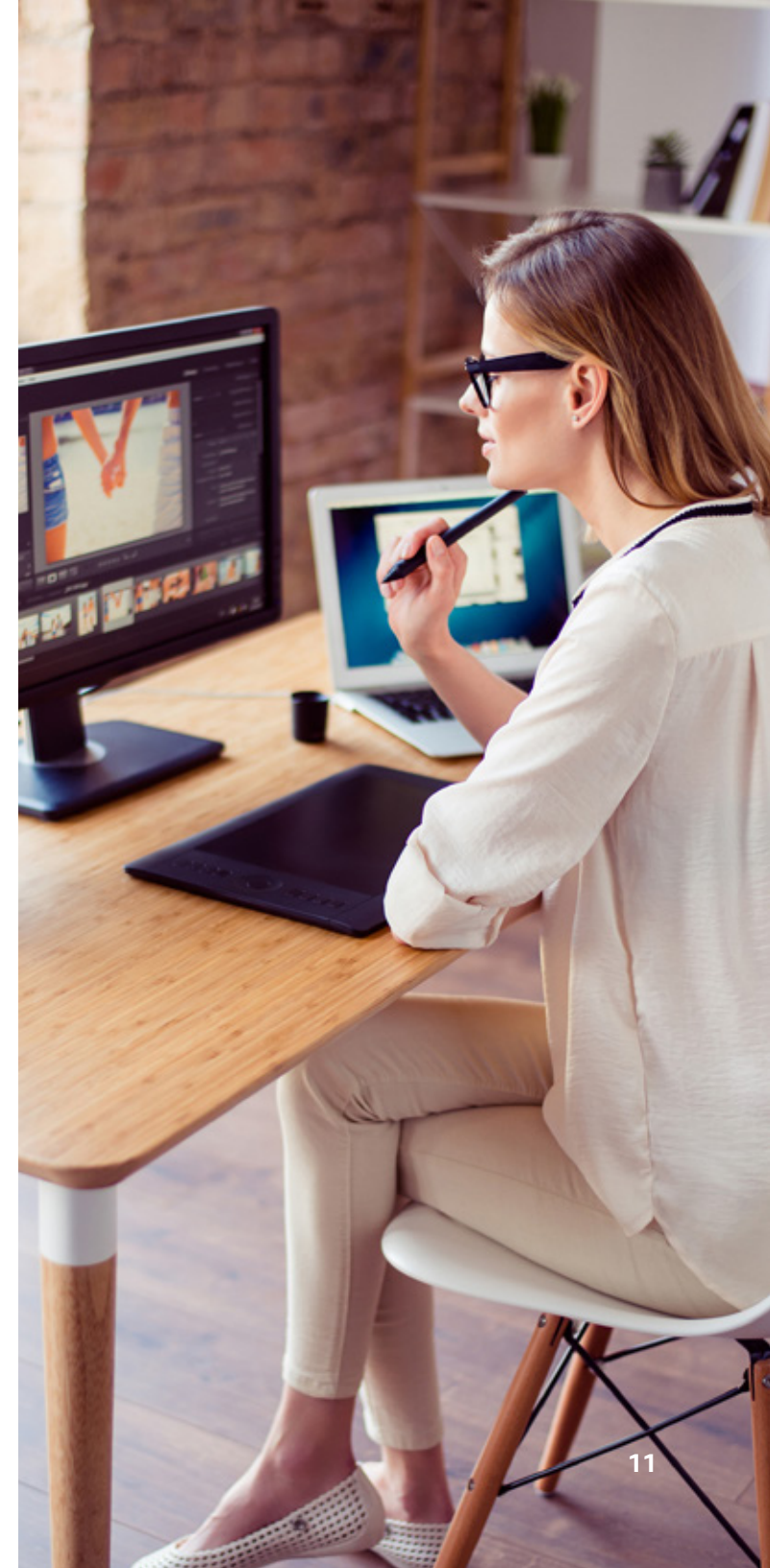
STEP 4

Scaling beyond pilot projects

After the first few successful pilots, you must take the next step on the journey: sustainably scaling AI across your business. This is both a technical and a cultural challenge. There are several ways startups approach the cultural shift necessary to scale AI. Some might find success by creating a center of excellence that rallies the community and continues to push for new initiatives. Or, like Amazon, you can make AI an integral part of your yearly planning processes, continuously bringing domain and technical experts together to brainstorm and determine their next steps.

Achieving scalability requires businesses to help their developers use AI. Building models at scale can be labor intensive, which can slow innovation. With **Amazon Bedrock** serverless experience, you can get started quickly to privately customize FMs with your own data and easily integrate and deploy them into your applications using AWS tools and capabilities you already know (including integrations with features like pipelines to manage your FMs at scale). Many startup customers are solving scalability with SageMaker to prepare data and build, train, and deploy models to and get them into production faster and at lower cost, enabling sustainable expansion of AI initiatives beyond pilot projects.

Many more businesses are scaling through **AWS AI services**, a set of pretrained and managed services that can be used as building blocks to address common use cases, including personalizing recommendations, improving safety and security, and increasing customer engagement.





How Fibe did it

Fibe, one of India's leading consumer lending apps, was looking to optimize its digital lending platform for the growing Indian market. To do this, the startup was looking to deeply integrate ML into its platform to offer faster and more convenient services to its customers.

To enhance the loan application process, Fibe implemented **Amazon Textract**, which utilizes optical character recognition (OCR) technology to extract text and data from Know Your Customer (KYC) documents. To reduce administrative complexity, Fibe implemented **Amazon Comprehend**, a natural language processing service, to extract relevant information from income statements and other application documents. Using **AWS Lambda**, Fibe streamlines development and maintenance processes. This service automatically handles code execution requests and simplifies processes like scaling instances. With AWS support, Fibe is spending less on IT, which directly enhances its profits.

[Read the customer story >](#)

“Our goal is to utilize AI more effectively in order to enhance the chatbot’s capabilities. We continue to rely on AWS for innovative solutions. ... Our relationship with AWS keeps our entrepreneurial spirit alive.”

Anil Sinha, CTO, Fibe

STEP 5

Measuring the results

When measuring the results of AI efforts, the traditional “project ROI” viewpoint—where a project has defined start and end points, a budget, and a return—is reductive and can be detrimental to the initiative’s success. If the project doesn’t generate a positive return within the given time frame, you and your team may decide to move on and miss out on critical opportunities down the line. Startups alike must measure AI efforts based on what success means for their business with regard to the processes being optimized. In addition, they must view AI efforts as long-term investments, acknowledging that a true “return” may not be realized for several years and after countless iterations.

When planning AI initiatives, it’s better to view the process through the lenses of agility, competitive advantage, or risk tolerance rather than “expected” return. A founder will have greater success if it disregards the question of “What will be my return on investment in X months?” in favor of something more like “If we don’t invest in this now, will we fall behind our competitors in X years when the technology matures?” While traditional ROI metrics may not be the best approach, the business impact of AI initiatives can still be measured—it just requires a different outlook. AI results can be measured through something resembling a “value tree,” where the main trunk of the tree represents the traditional revenue return and branches extending from the trunk recognize the value of other business outcomes. The specific branches of the value tree will depend on the business, the industry, and the initiative. But they might be things like time saved through automated processes, new leads, markets, and opportunities identified, customer service improvements, or increases in upsells.

Measuring the success of AI through a more holistic and long-term model will keep startups focused on the best outcomes for their business future.





How AI21 Labs did it

AI21 Labs (AI21) is a leader in generative AI and LLMs. Initially, the company released two models: one with seven billion parameters and another with 178 billion parameters. But its team saw an opportunity to offer customers an LLM of 17 billion parameters to bridge the gap between the existing sizes.

AWS trained the FM in under 20 days using SageMaker—saving several weeks of time compared with AI21’s previous training methods. “Because Amazon SageMaker handles node failures, restarts elegantly, and orchestrates large, distributed runs, the team working on pre-training the model can focus on core tasks,” says Dan Padnos, vice president of platform at AI21.

[Read the customer story >](#)

“We have a really good relationship with the AWS team—[its members] have gone deep into the technical details with us and collaborated on challenging tasks. Throughout the process, the AWS team has been creative and has had awareness about our challenges and goals.”

Dan Padnos, VP of Platform, AI21 Labs

Taking the next step with AWS

No matter where your startup is in its AI adoption journey, you can take the next step with AWS solutions built on the most comprehensive cloud platform and optimized for generative AI with high performance computing (HPC), security, and analytics. Featuring the world's broadest and deepest set of AI services, over 100,000 customers are running their AI workloads on AWS. Generative AI can help you unlock new business value for your startup. From reinventing customer experiences to enhancing productivity and accelerating growth, generative AI holds the power to help you transform your business.

Generative AI with AWS, by the numbers

100,000+

customers using AWS for their AI workloads

20+ YEARS

of building experience at Amazon

Improve customer experiences



Chatbots and virtual assistants: Streamline customer self-service processes and reduce operational costs by leveraging generative AI-powered chatbots, voice bots, and virtual assistants to automate responses to customer queries.



Agent assist and conversational analytics: Enhance agent performance to improve first contact resolution and augment tasks such as knowledge search, call summarization, and problem-solving. Managers can extract valuable insights to improve customer experience, monitor agent performance, and boost business performance.

Boost employee productivity



Conversational search: Improve employee productivity by quickly and easily finding accurate information and summarizing content through a conversational interface.



Code generation: Accelerate application development with code suggestions based on the developer's comments and code.



Automated report generation: Generate financial reports, summaries, and projections, saving time and reducing errors.

Enhance creativity and content creation



Marketing: Create engaging marketing content, such as blog posts, social media updates, and email newsletters, to save time and resources.



Sales: Generate personalized emails and messages based on prospect's profile and behavior, improving response rates. Generate sales scripts or talking points based on the customer's segment, industry, and product or service.



Product development: Generate multiple design prototypes based on certain inputs and constraints, speeding up the ideation phase or optimizing existing designs based on user feedback and specified constraints.

Accelerate process optimization



Document processing: Improve business operations by automatically extracting and summarizing data from documents and insights through generative AI-powered question and answering.



Data augmentation: Generate synthetic data to train ML models when the original dataset is small, imbalanced, or sensitive.

Solving the biggest artificial intelligence challenges

Most startups have made investments and progress in their AI journeys and are exploring the possibilities of generative AI. But many find themselves hitting stopgaps along the way, worried that costs and complexities will grow too high as they progress.

Throughout this eBook, we explored the steps to forge ahead and realize the full power of generative AI. To recap, let's look at the biggest challenges we identified along the way, along with a brief recommendation of how your startup can solve them.

To learn more about how you can overcome obstacles and accelerate your AI journey, visit the [AWS AI Resource Hub](#).

[Get started ›](#)

[Explore generative AI ›](#)

Challenge	Solution
Discouraging failures	Developing a fault-tolerant culture
Siloed, unprocessed data	Creating a modern data strategy that includes data lakes
Finding the right business problems	Building blended teams that include both technical and domain experts
The AI skills gap	Adopting new organizational models, processes, and team management philosophies
Sustainably scaling beyond pilot projects	Leveraging end-to-end tools like Amazon Bedrock and SageMaker to build and scale generative AI applications
Measuring the results	Forgoing traditional ROI metrics in favor of agility, competitive advantage, and risk tolerance using the value tree model