



INNOVATION BRIEF FOR EDUCATION

Maximize the value of generative AI for your institution

Harness generative AI's full potential



Executive summary

A surge of excitement across students, administration, and institutional leaders—combined with broad, easy access to various technology—has led us to an inflection point for generative AI. Today, institutions are evaluating the impact of generative AI and exploring how they can take advantage of its transformative value.

Institutions recognize the importance of this moment and the need to quickly develop a strategy to explore generative AI. But for many, questions remain, including:

- How is generative AI different from previous generations of AI?
- What are the main use cases that will add value to the organization?
- Which institutions are already using it?
- How/where should I start?
- What are the risks, and how can they be mitigated?

This innovation brief provides an overview of generative AI, outlining its capabilities, use cases, and value for education institutions. It also offers valuable insights from AWS subject matter experts, leveraging our extensive knowledge and experience in AI and machine learning (ML) technologies.



Who is this content for?

This innovation brief is designed for institutional leaders seeking to better understand generative AI—and learn how they can use it to improve and deliver results for their organization.



Table of contents

Introduction: A new world of intelligence.....	4
Understanding generative AI.....	7
Generative AI capabilities for education	9
Generative AI capabilities for research	10
Organizational considerations for generative AI	11
How AWS can help you succeed with generative AI	13
Next steps	16



INTRODUCTION

A new world of intelligence

Students, faculty, and institutional leaders alike are fascinated by generative AI's ability to create sophisticated content, generate insights, answer questions, and more—all from simple natural language prompts, often within seconds.

While a lot of attention has been given to how students and educators are using generative AI, there is an even bigger opportunity in how institutions will use it to deliver extraordinary experiences for their students, educators, and administrative staff. The true power of generative AI goes beyond a search

engine or a chatbot and will transform every aspect of how institutions operate.¹

Awareness of technology is spreading and the education industry is changing at an unprecedented pace. According to Forbes, educators are already leveraging AI to save time and improve the quality of teaching and student engagements. Routine work is automated, teacher's roles are elevated, and the learning experience is becoming more interactive.²

Seizing the opportunity

Institutional leaders are racing to seize the opportunity to revolutionize education presented by generative AI. For institutions, the impact is not only about financial opportunity, but also options such as improvements in curriculum delivery or an increase in student retention. More importantly, it is about the changes in pedagogy altogether, such as educators changing from custodians of knowledge into facilitators of the learning process. For universities and colleges of every size, generative AI is a truly revolutionary technology that can drive considerable value for educational leaders and has the power to fundamentally transform the institutional landscape

According to research by Goldman Sachs, generative AI could increase global GDP by as much as 7%, or roughly \$7 trillion, over the next 10 years.³

These lofty financial projections are not driven by consumer interest alone. The potential of generative AI to improve productivity and output accounts for just as much, if not more, of the excitement and enthusiasm surrounding the technology.

For universities and colleges of every size, generative AI is a revolutionary technology that is beginning to drive considerable value for educational leaders—and has the power to fundamentally transform the institutional landscape.



³ "Generative AI could raise global GDP by 7%," Goldman Sachs, April 2023

⁴ "Generative AI Market," Polaris Market Research, January 2023

Global generative AI market

Market forecast to grow at a CAGR of 34.2%⁴



The strategic imperative

Institutions around the world are looking to use generative AI to reimagine student and faculty experiences, optimize recruitment efforts, personalize learning content, and improve student graduation rates.

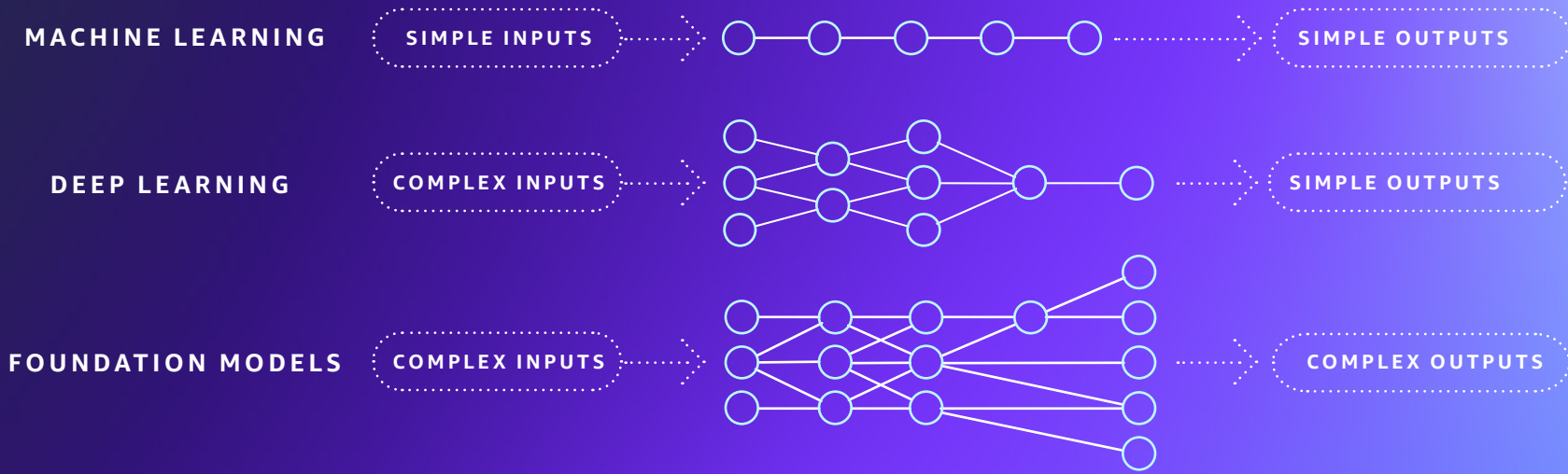
For most institutions, however, **the path** to achieving these benefits remains unclear.

Many acknowledge the need to invest in generative AI—and to do it quickly. However, few have succeeded in developing strategies on how to adopt the technology in order to achieve measurable outcomes in line with institutional goals.

Hear from Valerie Singer, general manager of global education at AWS, who has devoted her career to making the world more equitable and accessible through education, in the [AWS Behind the Cloud](#) podcast series.



[Watch the episode ›](#)



Understanding generative AI

Before your institution can fully unlock the value of generative AI, it's important to have a fundamental understanding of how the technology works.

Generative AI is a term used to describe algorithms that can create new content and ideas, including learning content, conversations, stories, images, videos, and music. Generative AI is powered by extremely large machine learning (ML) models that are pretrained on vast amounts of data. These are commonly known as **foundation models (FMs)**.

Traditional forms of ML allowed us to take simple inputs, like numeric values, and map them to simple outputs, like predicted values. With the advent of deep learning, we can take complicated inputs, like videos or images, and

map them to relatively simple outputs, for example, if the image contains a cat or not. With generative AI, we can leverage massive amounts of complex data to capture and present knowledge in more advanced ways—mapping complicated inputs to complicated outputs, like summarizing student data or research topics to extract key insights.

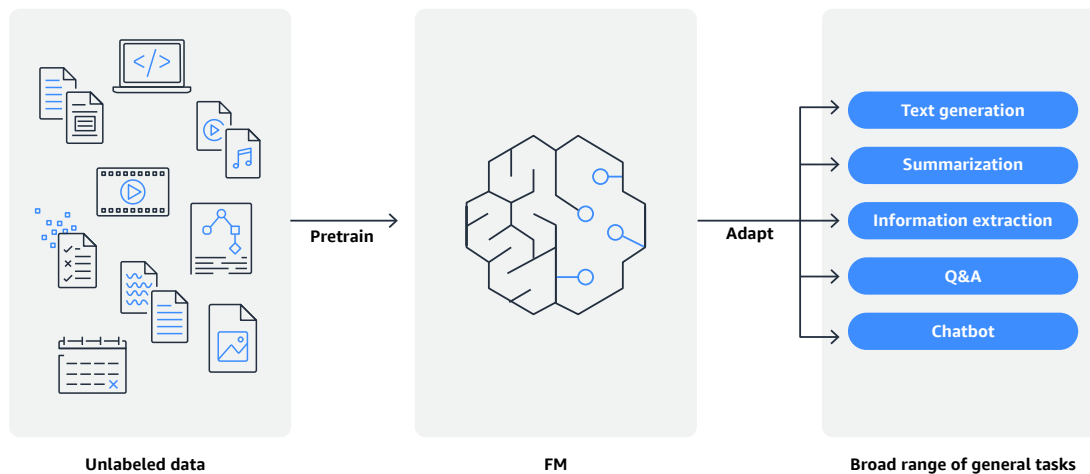
Text-based generative AI systems use a specific type of FM called a **large language model (LLM)**. LLMs can perform a wide range of tasks that span multiple domains, like writing code, composing music, writing an article, engaging in dialogue, and analyzing information in documents to answer questions.

Make data your differentiator

When you want to build generative AI applications that are unique to your needs, your institution's data is a strategic asset. FMs can be customized and fine-tuned with your institution's proprietary data to deliver a more differentiated experience compared to an "out-of-the-box FM." For example, having a solution where you can customize an FM to produce personalized learning recommendations aligned to the learner's interests but also with curriculum covered to date, enforcing a learning path.

Institutions can also use customized FMs to easily create unique learning content that addresses the individual student's needs based on university knowledge bases, referencing relevant white papers or university library collateral. For instance, research applicants could create a grant application by utilizing references and insights from research carried out at the relevant institution for the last five years.

Now that you have a general understanding of how the technology works, let's begin exploring how you can put generative AI to work for your institution.



Moments in generative AI history:

Today's FMs that are used to create generative AI applications are built atop a long history of AI innovation. Two of the earliest models with generative AI capabilities are the hidden Markov model (HMM) and the Gaussian mixture model (GMM), both developed in the 1950s. HMMs use known data to make educated guesses about unknown data (for example, predicting whether a card player is cheating based on their results). GMMs can examine a group of data (such as a music playlist) and subgroups within that data (for example, genres) to infer unknown information (such as "this is a rap song"). Both are still used today.

Gaussian mixture model

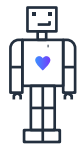
Generative AI capabilities for education

Below are several ways institutions can leverage generative AI to enhance productivity and drive impact:



Personalization

Create personalized learning pathways for students



Virtual assistants

Enhance the student experience with humanlike responses



Enhance experience

Reduce call wait times and decrease time to resolution with contact centers



Design and creativity

Get suggestions, generate prototypes, and explore innovative concepts



Productivity

Transform information into sample test questions and accelerate grading



Conversational search

Extract insights from all your institution's data



Content generation and summarization

Generate course materials, images, videos, and music. Summarize research documents, lecture transcripts, and class notes

Moments in generative AI history:

Another early example of generative AI is ELIZA, a chatbot (or "chatterbot," as they were previously known) developed by an MIT professor from 1964 to 1966. Like its namesake, Eliza Doolittle of *Pygmalion* and *My Fair Lady*, the program grew more sophisticated by "learning" from human interactions. ELIZA was most famously used to mimic the behavior of a therapist conducting an initial psychiatric interview, with the user playing the role of the patient.



Generative AI capabilities for research

Generative AI is uniquely changing the research landscape with opportunity to reduce research times with solutions ranging from analyzing images (e.g., satellite images and diagnostic imagery), to generating novel chemical compounds and medical treatments. Many research institutions are utilizing their knowledge and experience of ML to work on development of new LLMs built for particular languages enhancing cultural awareness or fine-tuning LLMs to perform specialized tasks.



Core benefits for research universities

From fact checking, data analysis, and summarization, generative AI is reshaping the way research is being conducted.

- Improve accuracy and reliability of response returns in line with cultural, language, or specialist topic requirements
- Boost productivity with analysis of large-scale datasets
- Save research time through automated grant application process
- Optimize business process with streamlined clinical research protocols



Technology Innovation Institute trains the state-of-the-art Falcon LLM 40B foundation model on Amazon SageMaker

"We proudly announce the official open-source release of Falcon-40B, the world's top-ranking open-source language model. Falcon-40B is an exceptional open-source model with 40B parameters, specifically designed as a causal decoder-only model. It was trained on a vast dataset of 1,000B tokens, including RefinedWeb enhanced with curated corpora ... By releasing Falcon-40B as an open-source model, we provide researchers, entrepreneurs, and organizations with the opportunity to harness its exceptional capabilities and drive advancements in AI-driven solutions from healthcare to space, finance, manufacturing to biotech; the possibilities for AI-driven solutions are boundless..."

Dr. Ebtesam Almazrouei
Executive Director–Acting Chief AI Researcher of the AI-Cross Center Unit and Project Lead for LLM Projects at Technology Innovation Institute

[Learn more ›](#)



Organizational considerations for generative AI

As you work to identify the capabilities of generative AI that are most impactful to your institution—and develop a strategy for implementing them into your everyday processes—you will need to determine which FMs to use in the creation of generative AI applications.

You should also carefully consider the infrastructure you will be using to support your FMs. Your models will benefit from a cost-efficient infrastructure that meets your requirements for performance.

When evaluating FMs used to create generative AI applications, look for models that offer:

1. Easy ways to build and scale generative AI applications with security and privacy built in
2. Performant, low-cost infrastructure to train your own models and run inference at scale
3. Ability to integrate your data to create and sustain your competitive edge



Responsible AI, security, and privacy

With their vast size and open-ended nature, FMs raise new issues in defining, measuring, and mitigating responsible AI concerns across the development cycle, such as accuracy, fairness, intellectual property (IP) considerations, hallucinations, toxicity, and privacy. For example, looking at the issue of fairness, if we ask an LLM to assign male and female pronouns to an academic researcher, will it do so at the same rate? Does that still apply if the prompt describes the academic researcher as being based in Oxford? What if we do the same for other professions like nurses, doctors, or Olympic participants? You can see that simply defining fairness in the context of an LLM is challenging and requires new approaches and solutions.

Generative AI technology and how it is used will continue to evolve, posing new challenges that will require additional attention and mitigation. To tackle these challenges and foster innovation, **academic, industry, and government partners** are working together to explore new solutions and concepts to ensure generative AI continues to evolve in a responsible, private, and secure way.

Data privacy and security are also critical to scaling generative AI responsibly. When it comes time to customize and fine-tune a model, institutions need to know where and how their data is being used. They need to be confident that their private data is not being used to train a public model and that student and financial data remains private. Institutions need security, scalability, and privacy to be baked in from the start to be viable for their applications.

Read the blog post *Responsible AI in the Generative Era*

[Learn more >](#)

Read *How McGraw Hill's innovation is transforming the way students are learning*

[Learn more >](#)



⁵ "UNESCO survey: Less than 10% of schools and universities have formal guidance on AI," UNESCO, September 2023

Creating guidance for the responsible use of AI

According to the recent UNESCO global survey of more than 450 schools and universities, less than 10 percent had institutional policies or any guidance around the use of generative AI while 81% of parents say a guidance will be helpful.⁵

The world economic forum released seven practical principles institutional leaders can consider to help create these guidance for learners and educators to improve learning results while promoting fairness in education.

1. Purpose
2. Compliance
3. Knowledge
4. Balance
5. Integrity
6. Agency
7. Evaluation

[Learn more >](#)

How AWS can help you succeed with generative AI

You can unlock the full value of generative AI for your institution with AWS. Create entirely new student and faculty experiences, reinvent your applications, drive unprecedented levels of productivity, and ultimately transform your institution.

Experience and expertise

One of the key advantages of AWS lies in a rich AI heritage built over two decades of focused investment. In fact, more than 100,000 customers currently use AWS for AI and ML.

Amazon, the driving force behind AWS, harnesses ML capabilities to power its e-commerce recommendations engine, optimize robotic picking routes in fulfillment centers, and much more. Further, ML informs Amazon's supply chain, forecasting, and capacity planning.

Deep learning is also employed in the Amazon Prime Air drone delivery system and the computer vision (CV) technology behind Amazon Go, the innovative retail experience that allows customers to select items and leave the store without traditional checkouts. And Alexa, which is supported by more than 30 different ML systems, helps customers with a wide array of tasks billions of times each week.

With thousands of dedicated ML engineers, AI and ML are deeply ingrained in the heritage of Amazon and AWS—continuing to shape the future.

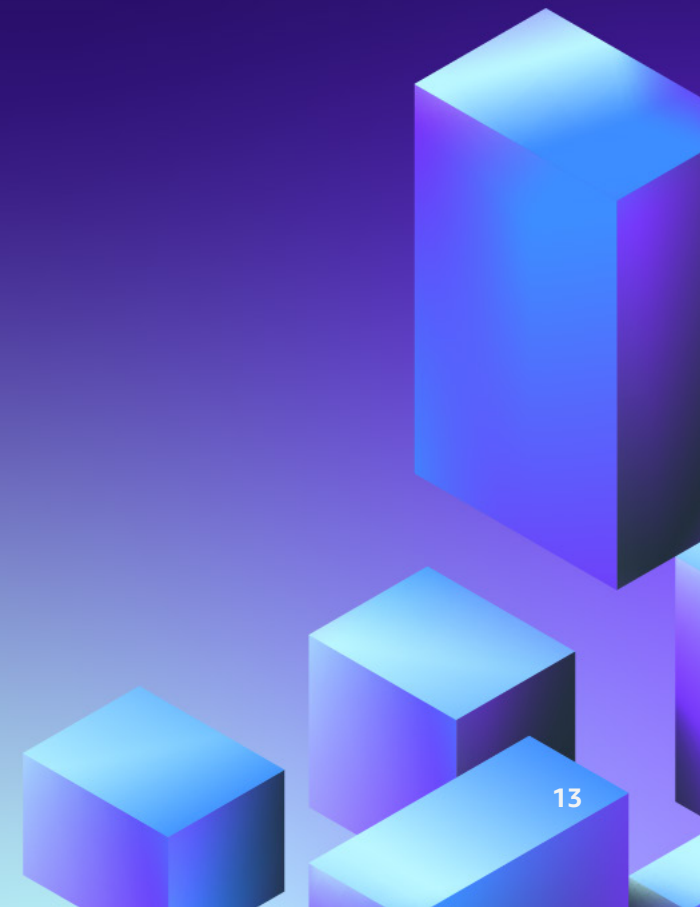
AWS Behind the Cloud:

Meet Valerie Singer, AWS education executive, uncover how generative AI is influencing the global education space from students to academic researchers; how Edtechs are leveraging and applying generative AI to advance the worldwide education community.

[Watch now >](#)



More than
100,000
customers currently
use AWS for AI and ML



Why build with AWS?

Institutions of all shapes and sizes choose to build generative AI and other AI and ML applications on AWS for many reasons. Here are some of the top advantages of building on AWS, according to our customers:

The easiest way to build and scale generative AI applications with security and privacy built in

Amazon Bedrock is the easiest way for customers to build and scale generative AI-based applications using FMs. Bedrock makes **Amazon Titan** FMs and models from leading AI companies such as AI21 Labs, Anthropic, Cohere, Stability AI, and Meta accessible via an API. Customers using Amazon Bedrock can harness the benefits of AWS, which is architected to be the most flexible and secure cloud computing environment available today. **Agents for Amazon Bedrock** is a fully managed capability that makes it easier for developers to create generative AI applications that can deliver up-to-date answers based on proprietary knowledge sources and complete tasks for a wide range of use cases.

The most performant, low-cost infrastructure for generative AI

For years, AWS has invested in developing silicon that delivers the highest levels of performance and cost optimization for AI and ML workloads. The results—**AWS Trainium** and **AWS Inferentia**—deliver the lowest costs for training models and running inference in the cloud. AWS has also developed **Amazon Elastic Compute Cloud** (Amazon EC2) instances to help you take advantage of these capabilities. For example, **Amazon EC2 Trn1** instances powered by Trainium save you up to 50 percent on training costs,⁶ while **Amazon EC2 Inf2** instances powered by AWS Inferentia2 deliver up to 40 percent lower cost per inference.⁷

Data as your differentiator

AWS simplifies the process of using your institution's data as a strategic asset to customize FMs and build your unique experience. Data is the difference between a general generative AI application and one that truly knows your institution and who you serve. And with the most comprehensive set of data and AI services, you can securely customize an FM on AWS with your data and build a model that is an expert on your purpose, your data, and your users (e.g., students, educators, and administrative staff).

Generative AI-powered applications transform how work gets done

AWS is building powerful new applications that transform how institutions get work done with generative AI. Boost productivity with purpose-built conversational agents that streamline coding in the enterprise with **Amazon CodeWhisperer** and simplify business intelligence with **Amazon QuickSight Generative BI**. With security, privacy, and responsible AI at the forefront, easy customization, and integration into your existing data sources and applications, institutions can quickly take advantage of generative AI without the heavy lifting.

Further reading on responsible AI:

[AWS responsible AI resource hub ›](#)

[EBook: Democratized, Operationalized, Trusted: The 3 Keys to Successful AI Outcomes ›](#)



⁶ Over other comparable Amazon EC2 instances

⁷ Compared to prior generation AWS Inferentia-based instances

AWS generative AI services

Facilitate your generative AI applications with a range of AWS technologies, including:



Amazon Bedrock >

Build and scale generative AI applications with FMs. Amazon Bedrock supports a variety of FMs, including:

- **Amazon Titan:** For text summarization, generation, classification, open-ended Q&A, information extraction, embeddings, and search
- **AI21 Labs Jurassic-2 Multilingual LLMs:** For text generation in various languages
- **Anthropic Claude 2:** LLM for conversations, question answering, and workflow automation based on research into training honest and responsible AI systems
- **Stability AI Stable Diffusion:** Generates unique, realistic, high-quality images, art, logos, and designs
- **Cohere Command + Embed:** Text generation model for business applications and embeddings model for search, clustering, or classification in over 100 languages
- **Meta Llama 2:** Pre-trained and fine-tuned LLMs for natural language tasks like question and answering and reading comprehension

AWS Trainium: Train models faster with up to 50 percent cost savings⁸ using this ML model accelerator⁹.

AWS Inferentia2: Run high-performance FM inference with up to 40 percent lower cost per inference using this accelerator.

Amazon CodeWhisperer: Enjoy 57 percent faster application development¹⁰ while helping to ensure security with this AI coding companion, which is at no cost for individual use.

Amazon QuickSight Generative BI: Transform traditional multistep business intelligence (BI) tasks into intuitive and powerful natural language experiences with generative BI capabilities in Amazon QuickSight.

Amazon SageMaker: Build your own FMs with managed infrastructure and tools to accelerate scalable, reliable, and secure model building, training, and deployment.

Amazon SageMaker JumpStart: ML hub that provides access to algorithms, models, and ML solutions so you can quickly get started with ML. With SageMaker JumpStart, ML practitioners can choose from a broad selection of **publicly available FMs**. ML practitioners can deploy FMs to dedicated SageMaker instances from a network-isolated environment and customize models using SageMaker for model training and deployment.

Next steps

Now that you have a better understanding of generative AI, what it can do, and its potential organizational benefits, the next step is to clearly define your objectives and identify use cases for taking advantage of it. It's best to start with smaller experiments and simple, precise goals. Once you've achieved some quick wins, you can begin scaling your efforts upward and outward.

Collaboration with experts is highly recommended to ensure you consider factors such as data availability, data quality, and ethical implications related to generative AI. Furthermore, infrastructure considerations should not be an afterthought, as they can significantly impact costs, scalability, and energy consumption. Engaging with AWS experts can provide valuable guidance throughout the decision-making process and stages of implementation.

The time is now

The dramatic rise of generative AI brings us to a tipping point. FMs grow more sophisticated and powerful every day. For institutions, it's the power to create entirely new student and faculty experiences and drive unprecedented levels of efficiency and productivity.

All of which leads to an indisputable fact: To compete in this new era of profound technological advancement, every institution needs to consider making generative AI a part of its innovation road map.

With the most cost-effective cloud infrastructure for generative AI; a host of AI products, services, and solutions; and years of trusted AI expertise, AWS can help turn the promise of generative AI into results for your institution.



Partner with AWS to accelerate your generative AI journey today.

[Get started >](#)