databricks

EBOOK

How to Transform the Public Sector with Generative AI

PUBLIC SECTOR

Contents

databricks

Generative AI Gaining Momentum Across the Public Sector

From optimizing supply chain management to improving citizen service and support, the emergence of generative artificial intelligence (AI) has transformed data-driven innovations across the government. Despite budget constraints, there is an undeniable paradigm shift taking place as every forward-thinking organization is increasing its investment in AI. In a recent joint research study with MIT, we discovered that 88% of executives are investing in generative AI, with more than a quarter already adopting it, and 99% of CIOs and CTOs are optimistic that data and AI investments will yield efficiencies, savings and innovations in the next two years. This resounding commitment to AI investment showcases a strong belief in the potential of generative AI to drive innovation and generate value.

But with all the buzz around generative AI technologies such as OpenAI's GPT-3, Google Bard or MosaicML MPT, how does anyone know where to begin? What is generative AI and how does it apply to the government?

At the highest level, generative AI is a subset of AI.

ARTIFICIAL INTELLIGENCE (AI)

MACHINE LEARNING (ML)

DEEP LEARNING (DL)

GENERATIVE AI
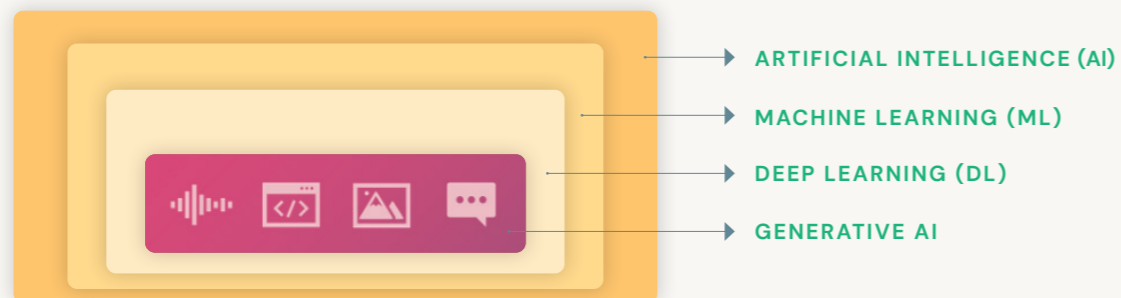
> **ARTIFICIAL INTELLIGENCE** uses computer science to create systems capable of emulating and surpassing human intelligence.

> **MACHINE LEARNING (ML)** is a branch of AI that trains computers to learn from existing data and make predictions without being explicitly programmed.

> **DEEP LEARNING** is a subset of ML that enables computers to process information and learn in a human-inspired way.

> **GENERATIVE AI** trains computers and systems to use generative models to produce new content like images, text, audio and music, video, code, 3D objects, and synthetic data.

These generative models utilize deep learning to discern patterns within large data sets like images, text and sound. Once the models complete the learning process, they can be employed to accomplish various tasks, including:

- Generating synthetic images and text
- Translating and question/answer generation
- Summarizing text
- Interpreting the intent or meaning behind text
- Converting audio snippets into text
- Transcribing music

databricks

Generative AI vs. LLMs: What Does It Mean?

You have most certainly heard a lot in recent months about large language models (LLMs) along with news and media about generative AI. Before diving into the rest of this eBook, we wanted to share a quick overview of what LLMs are versus generative AI. Large language models are a specific type of generative AI that are designed to understand and generate language. They use relationships in language data to predict and generate sequences of words and are built using deep learning techniques. These models can understand context, generate coherent text, and perform various natural language processing tasks like translation, summarization or answering questions. LLMs like ChatGPT have garnered a lot of publicity due to their powerful ability to understand and interact with users in a simplified manner.

In summary, while the terms are often used interchangeably, generative AI is a broader concept that encompasses various AI techniques for generating content, one of which is LLMs. In the rest of this eBook, we will primarily reference the umbrella category of generative AI.

Why All the Buzz Around Generative AI?

Massive, diverse, high-quality and open source data sets are driving the recent rise of generative AI applications and solutions. Through the use of data, AI models can be trained to learn patterns, correlations and characteristics of these large data sets. The result is pretrained, state-of-the-art models that can be readily accessed and tailored for various use cases across industries.

Significant advancements in computing are also making it easier for organizations to harness the capabilities of generative AI. Improvements in hardware, including graphics processing units (GPUs), robust cloud computing and open source AI software like Hugging Face, are fueling the shift toward generative AI. Now that data teams have the necessary tools and starting points, they can explore and experiment with generative AI.

Lastly, innovative deep learning models are gaining new enhancements on top of improved computational power — processing information more efficiently and exhibiting a greater capacity to respond like humans.

"

Generative AI evolves the possibility and promise of AI exponentially. You can transform the conversation between the creator and the computer.

CYNTHIA STODDARD
Senior Vice President and Chief Information Officer, Adobe

databricks

Realizing Generative AI's Potential for the Public Sector

The accuracy and effectiveness of generative AI models have hit a tipping point, making once dreamlike use cases a reality. Many open source and customizable models are now accessible to organizations, driving the next phase in AI adoption. Use cases that weren't feasible just a year earlier are being fulfilled cost-effectively by nontechnical business users. Through a joint research study with MIT (MIT Technology Review Insights: The great acceleration: CIO perspectives on generative AI), Databricks found that 80% of executives expect AI to boost efficiency by at least 25% and one-third expect the gain to be at least 50% in two years. These lofty expectations speak volumes about the excitement of this technology and the perceived impact within their organizations.

Executives across industries believe that generative AI use cases can cut costs and create new revenue by improving efficiency, speeding time-to-insights, expanding visibility and driving data-based decision-making. Due to the high ROI, it's no surprise that 88% of executives say they are investing in generative AI, and more than a quarter are already adopting it.

Behind this optimism, however, are common themes that these leaders acknowledge are a must-have foundation for generative AI to perform at scale with a sustainable TCO:

Data Access
Ability to access all your data — any type, from any source, in real time.

Data governance
Governance is at the center of any successful AI strategy. Models are only as good as the secure, accurate and relevant data being sourced to them. Organizations need to ensure data pipelines and connections between sources remain resilient as updates are made and new sources are introduced. Retrieval augmented generation (RAG) is an AI framework that has gained much attention as it provides organizations with the ability to retrieve the most accurate and up-to-date data to ensure LLM models remain fresh and relevant. Resilience and scale aren't possible without governance, which covers data and the AI models themselves.

Data sharing
Finally, unleashing the impact of generative AI requires effectively leveraging the power of your entire ecosystem. By adding critical third-party data sources like the Databricks Marketplace, organizations can enrich models and get to better insights, faster — moving from a reactive to proactive posture.

databricks

DATA ACCESS, GOVERNANCE AND SHARING ON THE LAKEHOUSE

DELTA LAKE: Delta Lake is the only open format storage layer that can automatically and instantly translate across open formats. Delta Lake unifies all data types for transactional, analytical and AI use cases out of the box, with support for streaming and batch operations. Delta Lake offers industry-leading performance and is the foundation of a cost-effective, highly scalable lakehouse.

UNITY CATALOG: Unity Catalog is the industry's first unified governance solution for data and AI on the lakehouse. With Unity Catalog, organizations can seamlessly govern their structured and unstructured data, machine learning models, notebooks, dashboards and files on any cloud or platform. This unified approach to governance accelerates data and AI initiatives while ensuring regulatory compliance in a simplified manner.

DELTA SHARING: Databricks and the Linux Foundation developed Delta Sharing to provide the first open source approach to data sharing across data, analytics and AI. Customers can share live data across platforms, clouds and regions with strong security and governance.

As generative AI steps into the limelight, its potential for creating new opportunities for organizations has swiftly turned from speculation into reality. The anticipation surrounding generative AI illustrates the limitless potential and transformative impact within the government..

databricks

Lakehouse AI: A Data-Centric Approach to Generative AI

Achieving meaningful results with generative AI hinges upon the ability of AI models to comprehend unique jargon and data sources. That means different agencies and organizations all have to think about generative AI differently and how their data sets can contribute to generative AI solutions. To truly differentiate in the market, organizations need to leverage their unique data in building generative AI solutions.

With Lakehouse AI and its unique data-centric approach, we empower the public sector to develop and deploy AI models with speed, reliability, and full governance. Organizations can start with their own data to quickly build and deploy generative AI solutions that are unique to their need and domain — whether they want to build Q&A search bots using open source LLMs and

retrieval augmented generation (RAG) together with their data, or create custom models by fine-tuning or training with their data. Lakehouse AI comes with built-in governance and monitoring to ensure that the models are being built securely, with data privacy, permissions and compliance taken care of.
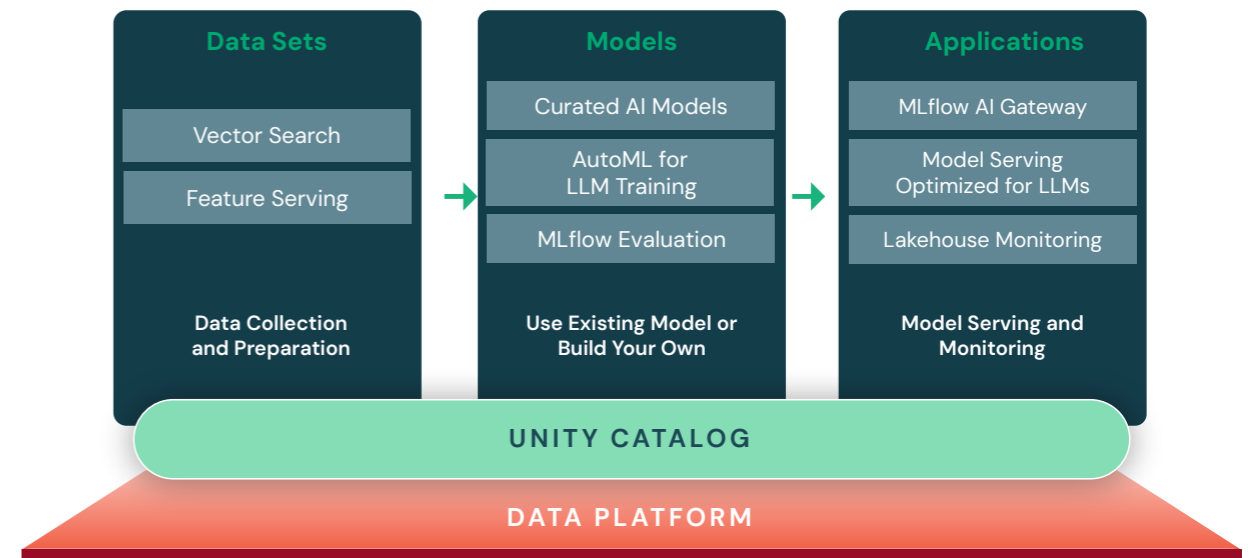
> Generative AI is driving disruption across every industry, and CIOs recognize that leveraging AI is no longer a nice-to-have but is imperative to remain competitive. To support this, we'll see data-driven companies extend and commercialize models with their own data, integrating them into customized applications that make sense for their business.

MATEI ZAHARIA
Co-founder and Chief Technology Officer, Databricks, and
Associate Professor of Computer Science, University of California, Berkeley

Lakehouse AI — Optimized for Generative AI

Data Sets	Models	Applications
Vector Search	Curated AI Models	MLflow AI Gateway
Feature Serving	AutoML for LLM Training	Model Serving Optimized for LLMs
	MLflow Evaluation	Lakehouse Monitoring
Data Collection and Preparation	Use Existing Model or Build Your Own	Model Serving and Monitoring

UNITY CATALOG

DATA PLATFORM

With MosaicML, organizations can quickly, securely and cost effectively build large AI models on their own data in their private environment.

databricks

Generative AI Use Cases

With the ability to generate content, images, text and even entire scenarios, generative AI has unlocked a multitude of applications.

DATABRICKS SOLUTION ACCELERATORS

Throughout this eBook we will be sharing links to relevant Databricks Solution Accelerators. Accelerators are purpose-built guides — fully functional notebooks and best practices — built to speed up results across your most common and high-impact use cases.

Check them out to learn more and share with your teams to save hours of discovery, design, development and testing — go from idea to proof of concept (POC) in as little as two weeks.

Optimizing citizen experience

According to Databricks and MIT research, the most valuable generative AI use cases over the next two years will be focused on citizen experience and personalization. With generative AI, organizations can summarize, translate and generate new text for humanlike interactions that adapt and match different languages, speaking styles and emotional tones. These models are also being applied to search engines to improve generated answers for faster resolutions.

With citizen-specific data in your models, you can quickly, accurately and — with fewer resources — recognize, anticipate and satisfy the needs of individual constituents, patients, etc.

Customize interactions and solve requests immediately
Whether it's paying taxes or signing up for services, all citizens should each get what they want — automated solutions for scheduling appointments, processing returns, and registering for services.

Generate written responses without humans
Automatically create new or update existing documents like finanial statements, threat assessment reports and citizen demographics

Continuously optimize constituent experience
Generative AI models constantly use feedback from prior experiences to influence future interactions with specific citizens.

DATABRICKS SOLUTION ACCELERATOR
Elevating customer satisfaction with LLM-powered chatbots

LEARN MORE →

databricks

Bolstering cybersecurity

Protecting mission-critical data from cyberattacks and accidental deletions, natural disasters, and hardware failures is paramount to keeping your organization running in today's cybersecurity landscape. Generative AI simplifies and reinforces cybersecurity responsibilities to control costs and deliver protection.

Many enterprises start their generative AI journey to increase profits. Prioritizing cybersecurity use cases is essential to avoid threats to national security and other risks.

Automate cybersecurity for better protection
Reduce the likelihood that an incident will occur due to human error — the number one cause of data loss — by automating threat detection, alerting and escalation, real-time responses, and various security-related tasks.

Multiply the force of cybersecurity analysts
A unified, data-centric AI platform democratizes generative AI for lower-skilled analysts to accelerate time-to-insights, reduce labor costs and allow higher-skilled analysts to focus on innovation.

Proactively fix security flaws in code
Scan infrastructure code, identify potential issues, and automate the appropriate reaction to maintain the quality of your code and prevent cybersecurity vulnerabilities.

Streamlining supply chains

Generative AI helps organizations understand and integrate data across the supply chain to match upstream expectations with downstream output. As a result, traditional supply chain operations are modernizing.

With generative AI, the supply chain can apply automation, operational data and high-visibility reporting to reduce disruptions, lower labor costs, increase speed, and predict with precision.

Identify complex patterns in demand and forecasting: Pair historical data and market trends with inventory levels, supplier lag time, and output to uncover intricate correlations and opportunities that legacy forecasting methods might overlook.

Select the best suppliers and manage with oversight: Utilize performance metrics, quality records and pricing to identify suppliers based on predefined criteria and support the relationship with risk and opportunity insights.

Enhance predictive maintenance: Prepare for evolving equipment needs and automate your response with generative AI's constant feedback loop that adjusts for model variations, changing environmental impacts and different types of machine use.

databricks

Industry-Specific Generative AI Applications

Organizations in the public sector face the added challenge of operating within the confines of state, federal, international and industry-specific regulations. Generative AI connects the complex legalese typically used to set standards with enterprise data across processes to ensure compliance. Leveraging these use cases, public sector enterprises can align with regulations, demonstrate proof to oversight committees through reporting, and avoid steep out-of-compliance fines and penalties.

Use cases

Optimize human resources across government agencies: Manage the challenges of an extensive employee base by automating résumé screenings, matching candidates to job descriptions, and analyzing feedback to improve hiring and workforce engagement.

Simplify technical documents for faster resolutions: Replace internal documentation with summarizations that can automatically be analyzed and made available for accurate search results and report creation.

Assist in training and education with a central knowledge base: Analyze audio recordings of lectures and training, syllabi, and lab instructions so students and trainees can search and automatically source answers about coursework.

Modernize legacy code bases and accelerate migration: Take advantage of innovations in code — including code converters, code copilots, and code explainers or summarizers that streamline and organize migration activities.

Identify foreign threats with intelligence analysis: Use text processing and extraction, multilingual analysis and translation, pattern recognition, data fusion, and adversarial intent detection for comprehensive intelligence gathering and discovery.

Assess data privacy and disclosure risk: Evaluate potential privacy risks, disclosure concerns, anonymization techniques and cybersecurity readiness, and suggest privacy-preserving methods to protect customers while maximizing data utility.

Manage assets with predictive maintenance: Extend the life of assets, avoid downtime, allocate resources, track asset inventory, and produce reports per regulations and risks.

databricks

USE CASE SPOTLIGHT
Regulatory compliance assistance

Navigating intricate legal frameworks and staying updated on evolving requirements is key for any organization to achieve regulatory compliance — whether it's enforcing regulations related to environmental protection or regulating transportation sectors such as aviation, railways and roadways to ensure proper safety standards are met. Due to the complexities of laws and evolving regulations, there are massive troves of documentation that need to be reviewed to ensure businesses and individuals adhere to legal standards.

With generative AI models, any business or organization can quickly interpret legal jargon, abstract clauses and complex terminologies, translating them into comprehensible insights that facilitate swift comprehension of not only compliance obligations but what is required to meet them.

USE CASE SPOTLIGHT
Call center supports/chatbots for constituent engagement

Effective call center support stands as a linchpin of successful governance and service delivery. Constituents, ranging from citizens to businesses, rely on government agencies to provide accurate information, assistance and solutions promptly. Whether it's a citizen inquiring about a tax filing deadline or a business seeking permits, delivering the right information at the right time cultivates a positive perception of public services and reinforces the vital relationship between the government and its constituents.

By powering sophisticated chatbots and virtual assistants, generative AI models are enabling government bodies to efficiently support their constituents across a wide array of topics — serving as expert guides providing reliable and relevant information.

DATABRICKS SOLUTION ACCELERATOR
Natural language processing (NLP) for customer service analytics

LEARN MORE →

databricks

Lakehouse for Generative AI: Invest in a Platform, Not Just a Model

Generative AI will have a transformative impact on every organization. Databricks has been pioneering AI innovations for a decade, actively collaborating with thousands of customers to deliver AI solutions, and working with the open source community on projects like MLflow, with 11 million monthly downloads. With Lakehouse AI and its unique data-centric approach, we empower customers to develop and deploy AI models with speed, reliability and full governance.

Databricks Lakehouse AI unifies the data and AI platform so government organizations can develop their generative AI solutions faster and more successfully — from using foundational SaaS models to securely training their own customs models with their enterprise data. Lakehouse AI lets you operate the right model for the right use case, from ML to AI. Organizations can accelerate their generative AI journey by bringing together data, AI models, LLM operations (LLMOPs), monitoring, and governance on the Databricks Data Intelligence Platform.
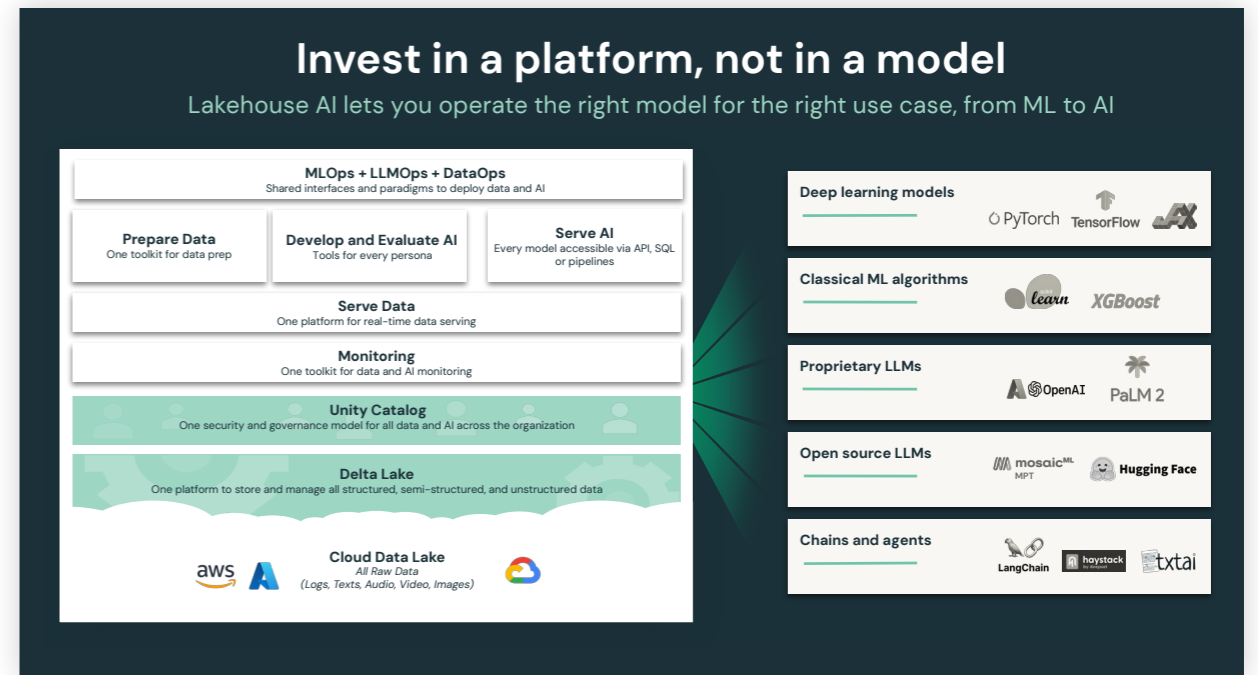
Lakehouse AI: The first AI platform built directly into the data layer.

> We found a very significant improvement in the accuracy and predictability of using that AI model to a point where there could be a significant reduction in mortality. For this to work, we have to build trust and help healthcare workers understand, hey, this is what AI can do. The technologies that we're putting in place are enabling physicians to be a part of the development of AI, and because of the level of validation involved, I think there will be more trust in the models that we develop.

RICHARD SPENCER SCHAEFER
Chief Health Informatics Officer, Kansas City VA Medical Center
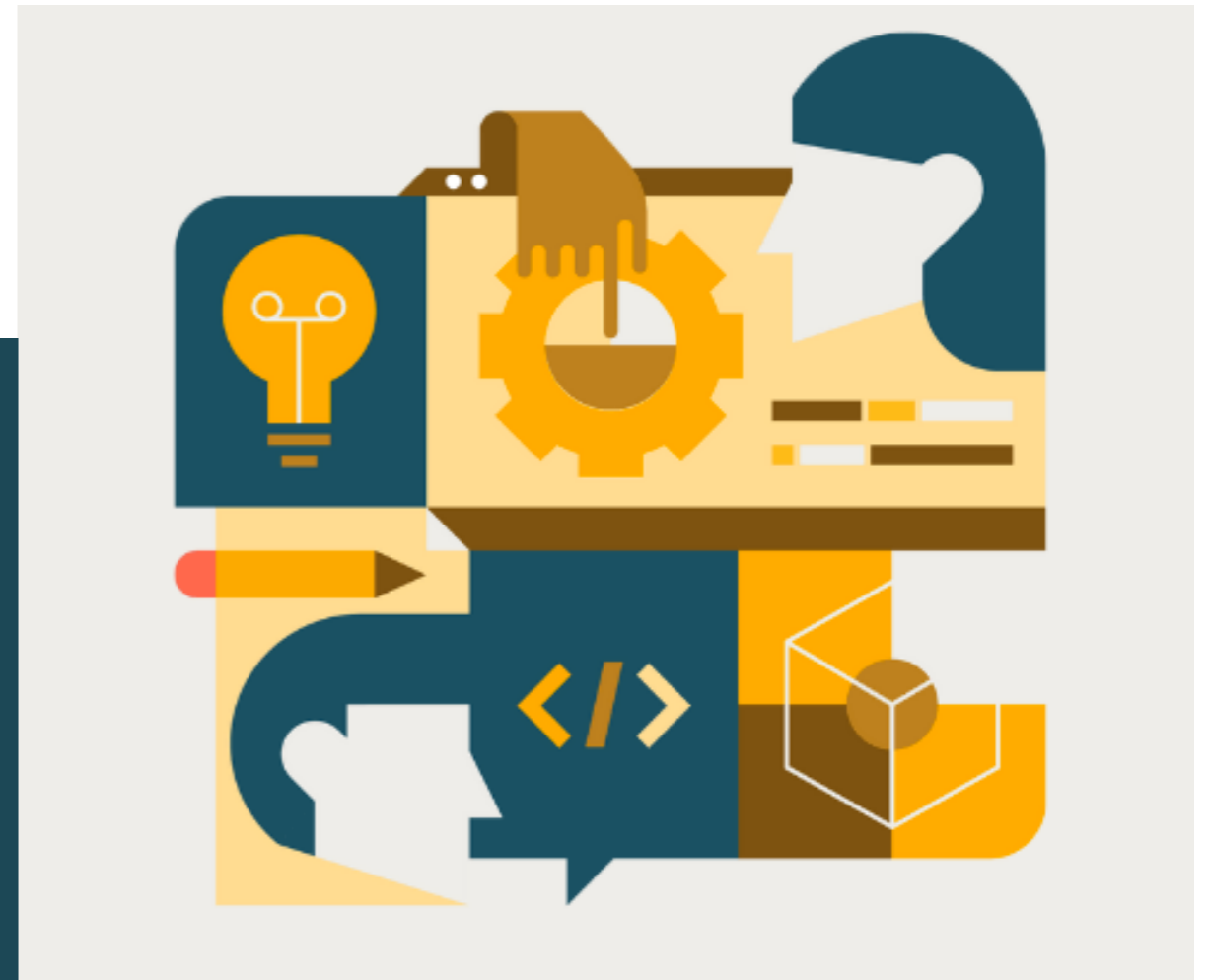
databricks

Start Your Teams' Journey With Generative AI Today

With Data Intelligence Platform as the cornerstone of your AI strategy, the full potential of generative AI is at your fingertips. Start on your generative AI journey with Databricks and redefine what's possible for your organization.

Tap the full potential of generative AI in this thorough yet concise eBook on the latest breakthroughs in NLP and LLM. Get up to speed on Lakehouse with free **Databricks Lakehouse Fundamentals** training.

Or, let's connect to discuss generative AI and your organization.

CONTACT US →

databricks

About Databricks

Databricks is the data and AI company. More than 10,000 organizations worldwide — including Comcast, Condé Nast, and over 50% of the Fortune 500 — rely on the Databricks Lakehouse Platform to unify their data, analytics and AI. Databricks is headquartered in San Francisco, with offices around the globe. Founded by the original creators of Apache Spark™, Delta Lake and MLflow, Databricks is on a mission to help data teams solve the world's toughest problems. To learn more, follow Databricks on Twitter, LinkedIn and Facebook.

START YOUR FREE TRIAL

Contact us for a personalized demo
databricks.com/contact