aws

User Guide

Amazon SageMaker

Amazon SageMaker: User Guide

Table of Contents

What is Amazon SageMaker?

Bringing together widely adopted artificial intelligence (AI) and analytics capabilities, the next generation of Amazon SageMaker delivers an integrated experience for analytics and AI with unified access to all your data. Collaborate and build in Amazon SageMaker Unified Studio using familiar AWS tools for SQL analytics, data processing, model development, and generative AI, accelerated by Amazon Q Developer. Access all your data whether it's stored in data lakes, data warehouses, or third-party or federated data sources, with governance built in to meet enterprise security needs.

The next generation of Amazon SageMaker overview

Guide to SageMaker

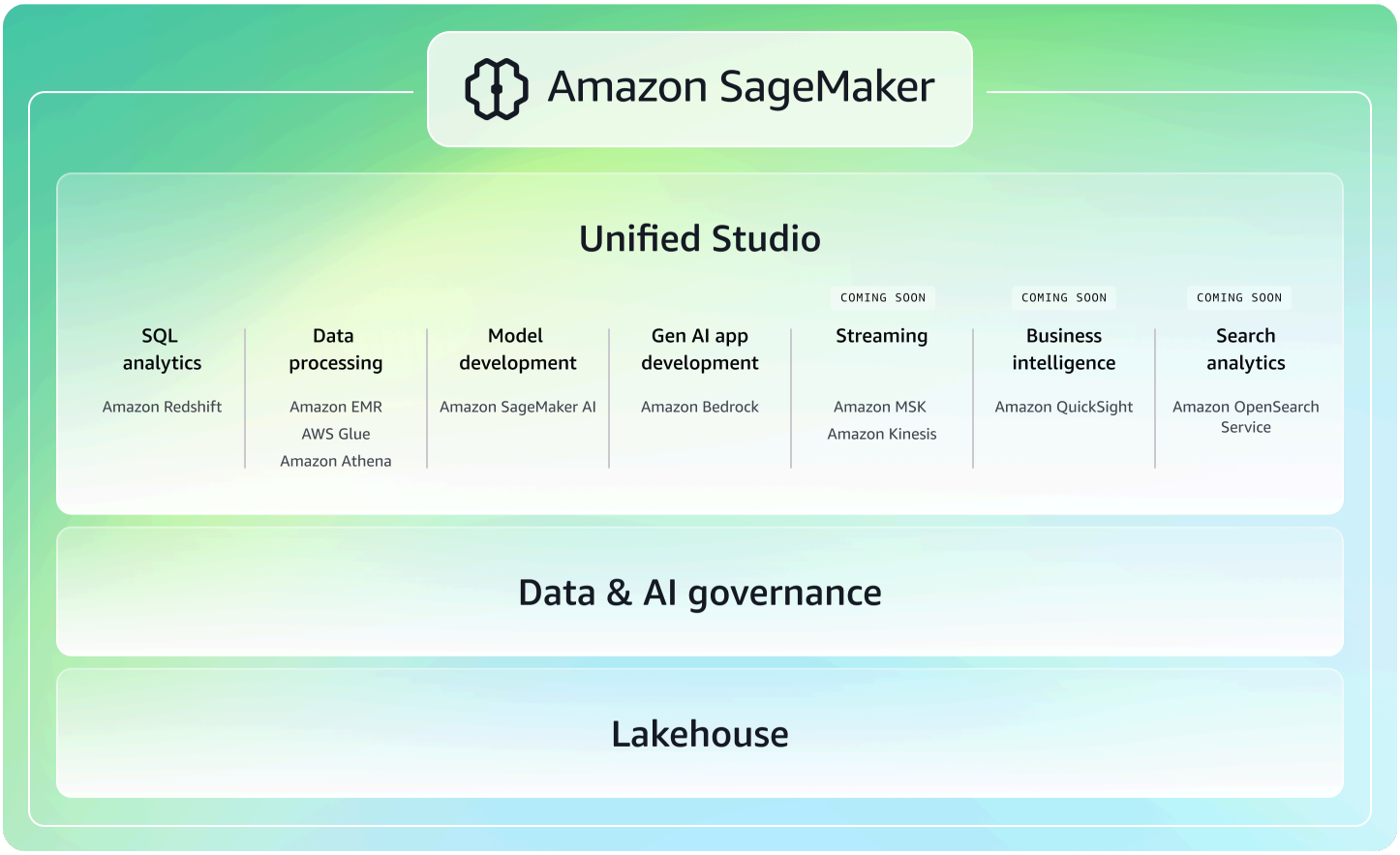
The next generation of Amazon SageMaker was announced at re:Invent 2024 serves as the center for all data, analytics, and AI. Analytics and AI workflows are converging, with organizations now using the same data sources for traditional analytics, machine learning, and generative AI. In response, AWS has created the next generation of SageMaker to serve as a unified platform for these workflows. The next generation of SageMaker brings together the purpose-built components needed for data exploration, preparation and integration, big data processing, SQL analytics, machine learning (ML) model development and training, and generative AI application development.

> **Note**
>
> The original Amazon SageMaker has been renamed SageMaker AI. It is available in the next generation Amazon SageMaker for those who wish to use it alongside additional capabilities, or as a standalone service for those who wish to focus specifically on building, training, and deploying AI and ML models at scale.

The next generation of Amazon SageMaker consists of two primary components:

1. Amazon SageMaker Unified Studio, which provides an integrated experience to use all your data and tools for analytics and AI
2. Data and AI governance, which applies enterprise-level security and data management with built-in governance throughout the entire data and AI lifecycle

Additionally, SageMaker is built upon an open lakehouse architecture that unifies access to all your data across Amazon Simple Storage Service (Amazon S3) data lakes, Amazon Redshift data warehouses, and other external sources



Unified Studio

Amazon SageMaker Unified Studio is a single data and AI development environment that brings together functionality and tools that AWS offers in Amazon EMR, AWS Glue, Amazon Athena, Amazon Redshift, Amazon MWAA, Amazon Bedrock, and Amazon SageMaker AI. From within the unified studio, you can discover, access, and query data and AI assets, then collaborate to build and share analytics and AI artifacts, including data, models, and generative AI applications.

Data & AI governance

The next generation of Amazon SageMaker simplifies the discovery, governance, and collaboration for data and AI. With Amazon SageMaker Catalog, users can securely discover and access approved data and assets using semantic search with generative AI–created metadata, or you could just ask Amazon Q Developer with natural language to find your data. Seamlessly share and collaborate on data and AI assets through publishing and subscribing workflows. With SageMaker, you can apply

Amazon Bedrock Guardrails to protect and filter your model outputs, helping ensure responsible gen AI application development. Build trust throughout your organization with data quality monitoring, sensitive data detection, and data and machine learning (ML) lineage.

Lakehouse architecture

The next generation of Amazon SageMaker is built on an open lakehouse architecture, fully compatible with Apache Iceberg. Unify all your data across Amazon S3 data lakes and Amazon Redshift data warehouses to build analytics and AI/ML applications on a single copy of data. The lakehouse gives you the flexibility to access and query your data with Apache Iceberg–compatible tools and engines. You can also connect to federated data sources such as Amazon DynamoDB, Google BigQuery, and Snowflake and query your data in-place. With zero-ETL integrations, you can bring data from operational databases and 3rd party applications into your lakehouse in near real-time. Integrated fine-grained access controls help you secure your data to ensure only the right people have access to the right data.

Capabilities of Amazon SageMaker Unified Studio

The next generation of Amazon SageMaker and its unified studio provide an integrated experience to use all your data and tools for analytics and AI. Discover your data and put it to work using familiar AWS tools for model development, generative AI, data processing, and SQL analytics. Work across compute resources using unified notebooks, discover and query diverse data sources with a built-in SQL editor, train and deploy AI models at scale, and rapidly build custom generative AI applications. Create and securely share analytics and AI artifacts such as data, models, and generative AI applications to bring data products to market faster.

Some common capabilities of Amazon SageMaker Unified Studio include the following:

SQL analytics

Leverage SageMaker's SQL analytic capabilities across all of your unified data through Amazon SageMaker's lakehouse architecture. Users have the flexibility to use Athena or Redshift query engines to support their analytical workloads. Query your data in open formats stored on Amazon S3 with high performance through Amazon Athena, eliminating the need to move or duplicate data between your data lakes and data warehouse. Include your Redshift data as part of the lakehouse architecture, leveraging the Redshift query engine for SQL workloads on structured data.

Data processing

Prepare, orchestrate, and process your data with capabilities in SageMaker, allowing you to run Apache Spark, Trino, and other open-source analytics frameworks in a unified data and AI development environment. Process your data, wherever it lives, with connectivity to hundreds of data sources with Amazon Athena, Amazon EMR, and AWS Glue.

Data integration

You can use data integration capabilities in Amazon SageMaker to connect to and act on all your data. With AWS data integration capabilities, you can bring together data from multiple sources, operationalize it, and manage to deliver high quality data to your lakehouse architecture, across your data lakes and data warehouses.

> **Note**
>
> *What data sources am I able to integrate with Amazon SageMaker?*
> You are able to unify all your data across Amazon Redshift data warehouses and Amazon S3 data lakes, including S3 Tables, with SageMaker's lakehouse architecture. Bring your operational databases and 3rd party application data like Salesforce and SAP to the lakehouse in near real time through zero-ETL integrations. You can use hundreds of connectors to integrate data from various sources. Additionally, you can access and query data in-place with federated query capabilities across third-party data sources.

Machine learning and model development

Amazon SageMaker AI is a fully managed service that brings together a broad set of tools to enable high-performance, low-cost machine learning (ML). Most capabilities of SageMaker AI are available as part of Amazon SageMaker Unified Studio, in addition to being available in Amazon SageMaker Studio. With SageMaker AI, you can build, train and deploy ML models at scale using tools like notebooks, debuggers, profilers, pipelines, MLOps, and more—all in one integrated development environment (IDE).

> **Note**
>
> *When should I use SageMaker Unified Studio instead of SageMaker AI studio?*

Currently, SageMaker Unified Studio should be used when you are looking to unify and share your data as a single integrated experience across analytics, ML, and gen AI workloads. You are able to eliminate data silos with an open lakehouse architecture to unify access to data lakes, data warehouses, third-party or federated data sources, and meet all enterprise security needs with built-in data and AI governance.
If you want to solely focus on the purpose-built tools to perform all machine learning (ML) development steps, from preparing data to building, training, deploying, and managing your ML and gen AI models, SageMaker Studio remains a great choice. Additionally, use SageMaker Studio when there are requirements for RStudio, Canvas, real-time collaboration via shared spaces, and Feature Store.

Generative AI application development

Access Amazon Bedrock's capabilities through SageMaker Unified Studio to quickly build and customize your generative AI applications. This intuitive interface lets you work with high-performing foundation models (FMs) from leading companies like Anthropic, Mistral, Meta, and Amazon, and use advanced features like Amazon Bedrock Knowledge Bases, Amazon Bedrock Guardrails, Amazon Bedrock Agents, and Amazon Bedrock Flows. You can develop generative AI applications faster within SageMaker Unified Studio's secure environment, ensuring alignment with your requirements and responsible AI guidelines.

Note

When should I use Bedrock in SageMaker Unified Studio versus the standalone Amazon Bedrock service?
Amazon Bedrock's capabilities in Amazon SageMaker Unified Studio are ideal for enterprise teams who need a governed low-code/no-code environment for collaboratively building and deploying generative AI applications, alongside unified analytics and machine learning capabilities.
Customers can use the standalone Bedrock service from the AWS Management Console or Bedrock API when they want to leverage the full feature set of Bedrock including the latest agents, flow and guardrail enhancements, and the Bedrock SDK.

Get started with Amazon SageMaker

You can view demos of Amazon SageMaker and get started by setting up a domain and project.

View demos of Amazon SageMaker

To see Amazon SageMaker before using it yourself, you can review the following clickthrough demos:

- For an end-to-end demo, see the Amazon SageMaker detailed clickthrough experience. This demo includes Amazon SageMaker Lakehouse, Amazon SageMaker Catalog, and more in Amazon SageMaker Unified Studio.
- For a demo of Amazon SageMaker Lakehouse, see Amazon SageMaker: Access data in your lakehouse. This demo includes Amazon SageMaker Lakehouse in Amazon SageMaker Unified Studio, including adding a data source and querying data.
- For a demo of the Amazon SageMaker Catalog, see Amazon SageMaker: Catalog. This demo includes Amazon SageMaker Catalog in Amazon SageMaker Unified Studio, including browsing assets and subscribing to an asset.
- For a demo of generative AI, see Amazon SageMaker: Generative AI playground and Gen AI app development.

Get started with setting up Amazon SageMaker

To get started using Amazon SageMaker, go to *Setting up Amazon SageMaker* in this guide to set up a domain and create a project. This domain setup and project creation is a prerequisite for all other tasks in Amazon SageMaker.

Prerequisites for Amazon SageMaker

Complete the following prerequisite tasks before you can set up Amazon SageMaker and proceed with the use cases in this guide.

Topics

Sign up for an AWS account

If you do not have an AWS account, complete the following steps to create one.

To sign up for an AWS account

1. Open https://portal.aws.amazon.com/billing/signup.
2. Follow the online instructions.

   Part of the sign-up procedure involves receiving a phone call and entering a verification code on the phone keypad.

   When you sign up for an AWS account, an *AWS account root user* is created. The root user has access to all AWS services and resources in the account. As a security best practice, assign administrative access to a user, and use only the root user to perform tasks that require root user access.

AWS sends you a confirmation email after the sign-up process is complete. At any time, you can view your current account activity and manage your account by going to https://aws.amazon.com/ and choosing **My Account**.

Create a user with administrative access

After you sign up for an AWS account, secure your AWS account root user, enable AWS IAM Identity Center, and create an administrative user so that you don't use the root user for everyday tasks.

Secure your AWS account root user

1. Sign in to the AWS Management Console as the account owner by choosing **Root user** and entering your AWS account email address. On the next page, enter your password.

   For help signing in by using root user, see Signing in as the root user in the *AWS Sign-In User Guide*.

2. Turn on multi-factor authentication (MFA) for your root user.

   For instructions, see Enable a virtual MFA device for your AWS account root user (console) in the *IAM User Guide*.

Create a user with administrative access

1. Enable IAM Identity Center.

   For instructions, see Enabling AWS IAM Identity Center in the *AWS IAM Identity Center User Guide*.

2. In IAM Identity Center, grant administrative access to a user.

   For a tutorial about using the IAM Identity Center directory as your identity source, see Configure user access with the default IAM Identity Center directory in the *AWS IAM Identity Center User Guide*.

Sign in as the user with administrative access

- To sign in with your IAM Identity Center user, use the sign-in URL that was sent to your email address when you created the IAM Identity Center user.

  For help signing in using an IAM Identity Center user, see Signing in to the AWS access portal in the *AWS Sign-In User Guide*.

Assign access to additional users

1. In IAM Identity Center, create a permission set that follows the best practice of applying least-privilege permissions.

   For instructions, see Create a permission set in the *AWS IAM Identity Center User Guide*.

2. Assign users to a group, and then assign single sign-on access to the group.

   For instructions, see Add groups in the *AWS IAM Identity Center User Guide*.

Setting up Amazon SageMaker

Complete the following tasks to set up Amazon SageMaker.

Topics

Step 1 - Create an Amazon SageMaker unified domain

Complete the following procedure to create an Amazon SageMaker unified domain with the Quick setup option.

> **Important**
>
> Note that there is an additional charge for any VPC or resources that AWS sets up if you chose the Quick setup option for domain creation.

1. Navigate to the Amazon SageMaker management console at https://console.aws.amazon.com/datazone and use the region selector in the top navigation bar to choose the appropriate AWS Region.
2. Choose **Create a Unified Studio domain** and then choose **Quick setup**.

   With this option, you're choosing to create an Amazon SageMaker unified domain and you're letting Amazon SageMaker configure your domain with the following default capabilities that you can customize later:

   - Data analytics, machine learning, SQL, and generative AI
   - Data and AI governance
   - Generative AI app development using Amazon Bedrock serverless models
   - Amazon Q - Free tier
   - Authentication via AWS IAM or AWS IAM Identity Center

3. If you see the following note **No VPC has been specifically set up for use with Amazon SageMaker Unified Studio**, you can use the **Choose VPC** or **Create VPC** buttons to **Create a new VPC (recommended)** or choose an existing properly-configured VPC.

   If you plan to choose your own VPC, Amazon SageMaker Unified Studio enables you to choose VPCs within the same account as well as shared VPCs from other member accounts of the AWS organization. For more information, see Share your VPC subnets with other accounts.

   > **Note**
   >
   > If you choose to create a new VPC, note that the VPC template with which it is created is not intended for production use. You can use this template as a start and modify it for your organization's purposes.

4. If you see the following note **No models accessible**, you can use the **Grant model access** button to grant access to Amazon Bedrock serverless models for use in Amazon SageMaker.
5. Expand the **Quick setup settings** section and review the specified configurations for the domain. Leave these defaults and then choose **Continue** to proceed with creating you domain.

   > **Note**
   >
   > For more information, see IAM roles for Amazon SageMaker Unified Studio.

6. On the **Create IAM Identity Center user** page, create a new or select an existing SSO user that you want to enable to log in to Amazon SageMaker Unified Studio. This is done because IAM roles that are used to create Amazon SageMaker unified domains cannot log in to Amazon SageMaker Unified Studio. The SSO user specified here is used as the administrator in Amazon SageMaker Unified Studio.
7. Choose **Create domain**.

Step 2 - Create a new project

In Amazon SageMaker, projects enable a group of users to collaborate on various business use cases. Within projects, you can manage data assets in the Amazon SageMaker catalog, perform data analysis, organize workflows, develop machine learning models, build generative AI apps, and more.

Project name and description

To begin creating a project, navigate to the Amazon SageMaker landing page and choose **Create project**.

The project name and description includes the following fields:

- Project name - the name of your project. Enter a name here. The name of the project can not be edited after the project is created.
- Description - an optional description of your project. You can edit this later.
- Project profile - project profiles define which resources and tools should be provisioned in the project. These include tools and compute resources for SQL, data science, data engineering, and machine learning development. Project profiles can include resources and tools from Amazon Redshift, Amazon SageMaker AI, and other AWS services. To complete the use cases in this getting started guide, choose the **All capabilities** project profile.

Choose **Continue** to review parameters.

Review parameters

On the next page of project creation, you can review and optionally edit the names and values for different resources that are created when the project is created. You can leave all the defaults and then choose **Continue**.

Review

Use the last page of project creation to review the configurations you have selected. When everything is configured as desired on the project creation review page, choose **Create project**.

You are then redirected to the project home page. The project will start building and a progress bar will appear with the status.

Document history for the Amazon SageMaker User Guide

The following table describes the documentation releases for Amazon SageMaker.

Change	Description	Date
Initial release	Initial release of the Amazon SageMaker User Guide	June 13, 2025