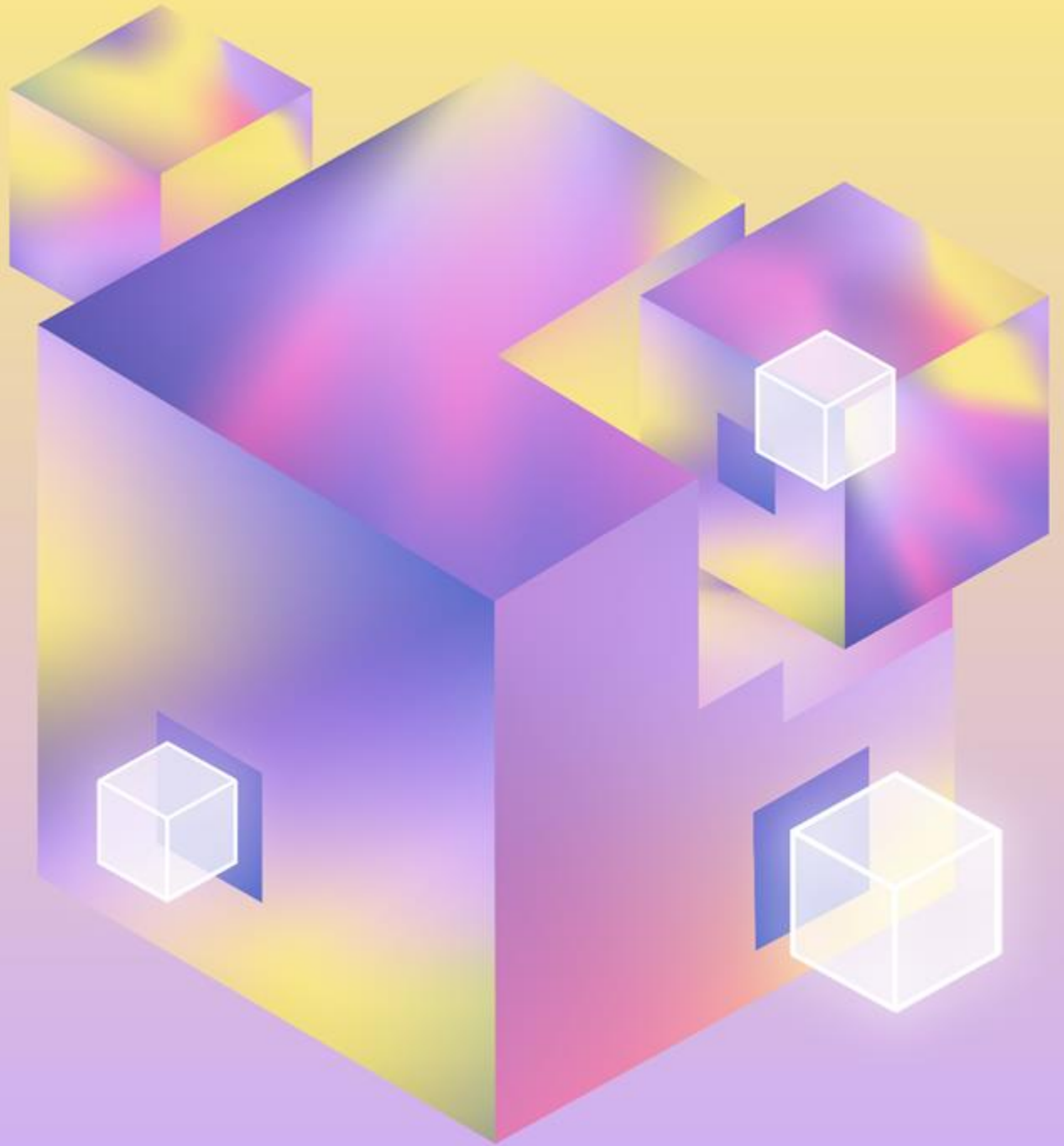




# Builders Online Series

31 JULY, 2025 | APJ



# AWS Builders Online Series

## Select the right AI/ML and generative AI tools for your use case

**Melwin Pais**

Senior Solutions Architect  
AWS



© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.



# Session agenda

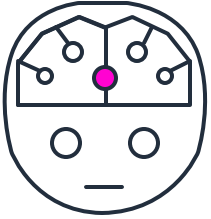
- AI overview and use cases
- Classification AI and predictive AI tools
- Generative AI tools
- Data, security and responsible AI
- Resources



# Expanding your AI toolbox on AWS

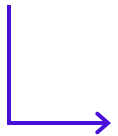


# AI, ML and generative AI



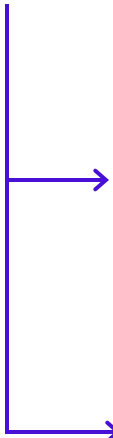
## Artificial intelligence (AI)

Any technique that enables computers to mimic human intelligence using logic, if-then statements, and machine learning



## Machine learning (ML)

A subset of AI that uses machines to search for patterns in data to build logic models automatically



## Classification AI and Predictive AI

A subset of ML that recognizes patterns to identify something (Classification AI) or predicts future trends based on statistical patterns and historical data (Predictive AI)



## Generative AI

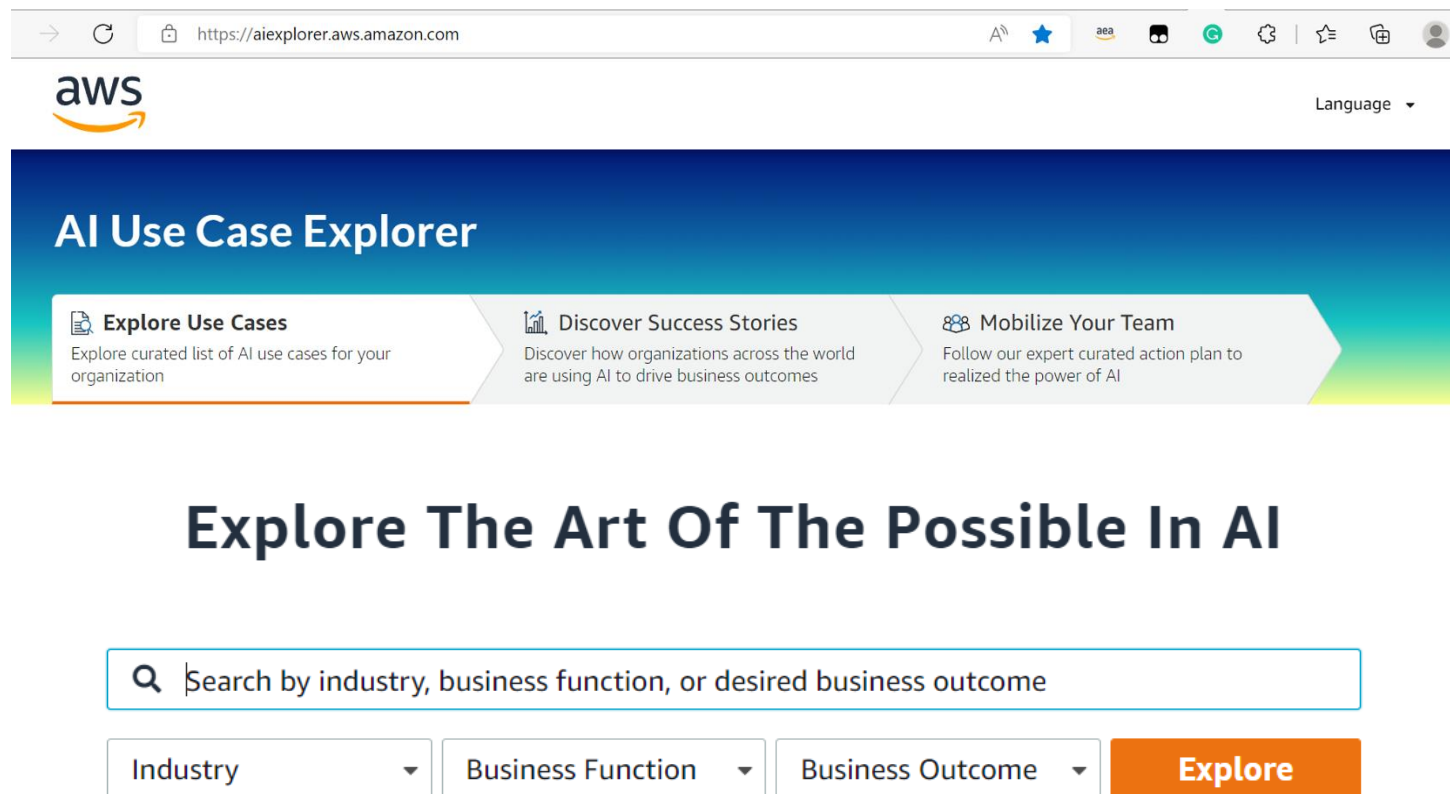
A subset of DL that can create new content and ideas powered by large, pretrained models called foundation models (FMs)

aiexplorer.aws.amazon.com



# AI Use Case Explorer

Easily find the most relevant AI use cases with related content and guidance to make them real







**Enhance customer  
experience**



**Augment human  
ingenuity**



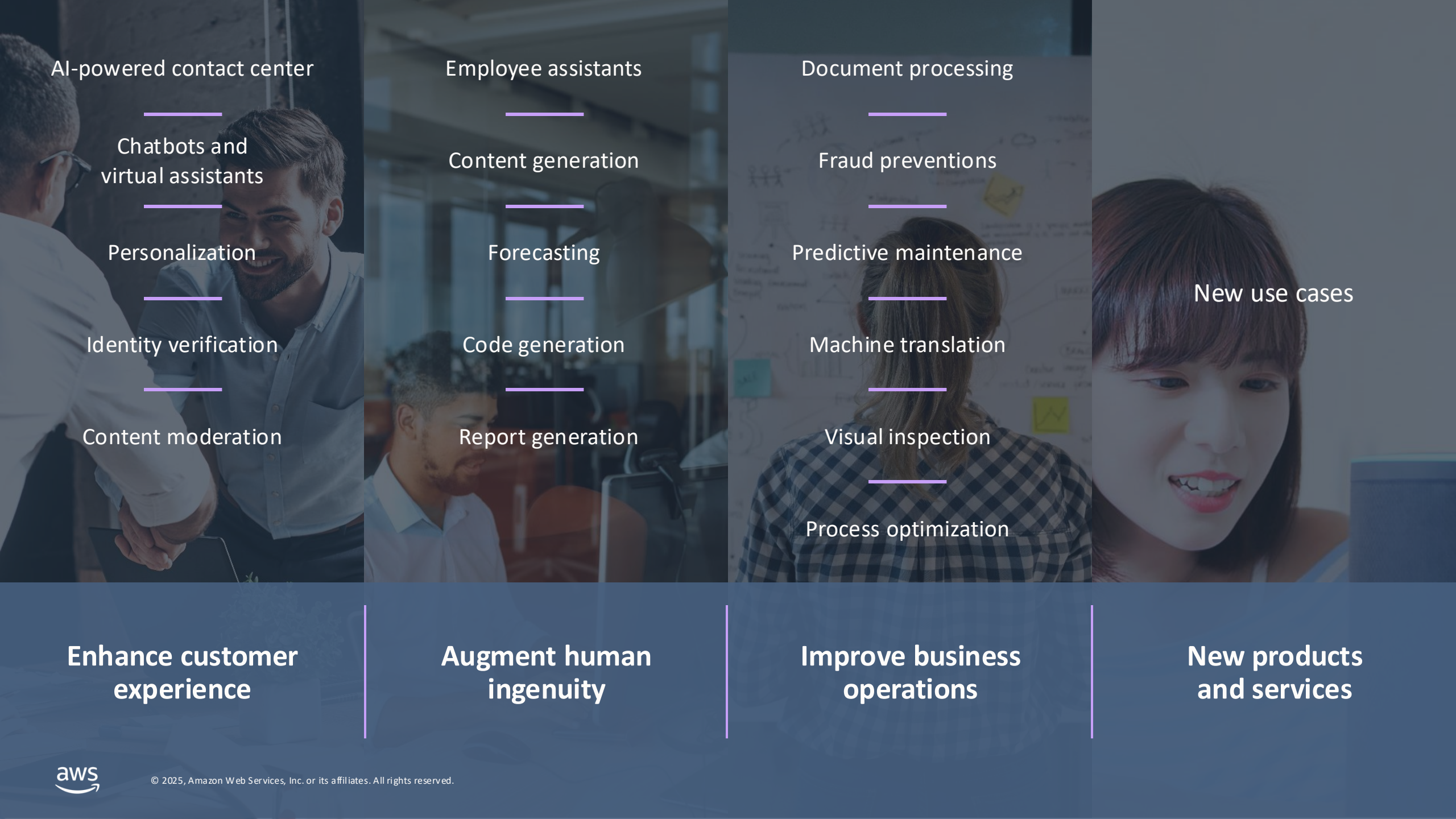
**Improve business  
operations**



**New products  
and services**







AI-powered contact center

Chatbots and  
virtual assistants

Personalization

Identity verification

Content moderation

Employee assistants

Content generation

Forecasting

Code generation

Report generation

Document processing

Fraud preventions

Predictive maintenance

Machine translation

Visual inspection

Process optimization

New use cases

**Enhance customer  
experience**

**Augment human  
ingenuity**

**Improve business  
operations**

**New products  
and services**





# Personalization

Generate personalized recommendations to increase customer engagement

## Requirements

- Deliver unique homepage experiences
- Help customers discover products faster
- Target customers more accurately
- Highlight new products, content, and promotion offerings

## Solutions

- Amazon Personalize for recommendations, personalized search, ranking, content generation and next best action
- Amazon SageMaker for advanced recommender systems and BYO models

## Customer references



## AI service



Amazon Personalize

## ML service



Amazon SageMaker

## Solution

Maintaining personalized experiences with machine learning

## Solution

Personalized recommendations

# Conversational AI

Voice and text-based interfaces to enhance end user experience and increase customer satisfaction

## Requirements

- Self-service bots
- Virtual agents and assistants
- Transactional bots
- Proactive help based on usage behavior

## Solutions

- Automatic Speech Recognition (ASR)
- Natural Language Understanding (NLU)
- Natural Language Processing (NLP)
- Language and named entity detection

## Customer references



## Underlying AI services



## Solution

QnABot on AWS

# Content moderation

Create safe online environments, protect your brand, and minimize moderation costs

## Requirements

- Improve safety for users and brands
- Streamline content moderation operations
- Support domains like gaming, social media, e-commerce, advertising
- Increase reliability and lower costs

## Solutions

- Image recognition and video analysis
- Contextual insights and sentiments
- PII and PHI detection
- Category detection & moderation
- Human in the loop

## Customer references

DeNA



DREAM11

FLIPBOARD

CoStar™

## Underlying AI services



Amazon  
Augmented AI



Amazon  
Rekognition



Amazon  
Transcribe



Amazon  
Comprehend



Amazon  
Translate

## Guidance

[AWS Solution Guidance](#)

# Identity verification

Enable secure and compliant digital experiences

## Use case categories

- Customer onboarding & verification
- Online proctoring
- Gig economy verification
- Event & airport check-ins
- Use challenges

## Solutions

- Detect and classify identity document
- Detect face liveness
- Validate and compare selfie with ID
- Detect duplicate users
- Signature detection

## Customer references



## Underlying AI services



Amazon  
Rekognition



Amazon  
Textract

## Guidance

[AWS Solution Guidance](#)



# Intelligent document processing

Make faster decisions by automatically extracting and analyzing data from documents

## Use case categories

- Insurance claims text extraction
- Mortgage applications data extraction
- Medical forms automation
- Financial invoices processing
- Legal contract validations

## Solutions

- Optical character recognition
- Form and table extraction
- Sensitive data redaction
- Invoices and receipts processing
- Human in the loop

## Customer references



## AI services



Amazon  
Comprehend



Amazon  
Textract



Amazon  
Augmented AI

## ML service



Amazon  
SageMaker

## Solution

Document understanding

## Solution

Extract and analyze  
data from documents

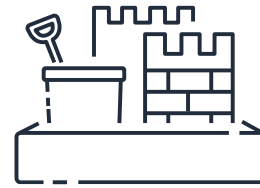




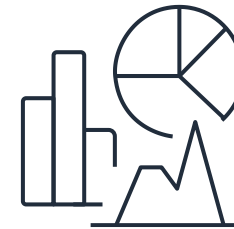
# Amazon SageMaker

Build, train, and deploy ML models for any use case with fully managed infrastructure, tools, and workflows

## End-to-end ML journey



**BUILD**



**TRAIN**



**DEPLOY**

aiexplorer.aws.amazon.com



# AI Use Case Explorer




Easily find the most relevant AI use cases with related content and guidance to make them real

A screenshot of the AWS AI Use Case Explorer website. The browser address bar shows the URL https://aiexplorer.aws.amazon.com. The AWS logo is in the top left, and a language dropdown is in the top right. The main header is a blue gradient bar with the text "AI Use Case Explorer". Below this is a horizontal navigation bar with three sections: "Explore Use Cases" (with a document icon), "Discover Success Stories" (with a bar chart icon), and "Mobilize Your Team" (with a group of people icon). Each section has a brief description. Below the navigation bar is a large heading "Explore The Art Of The Possible In AI". Underneath is a search bar with the placeholder text "Search by industry, business function, or desired business outcome". Below the search bar are three dropdown menus labeled "Industry", "Business Function", and "Business Outcome", followed by an orange "Explore" button.


aws

Language ▾

## AI Use Case Explorer

-  **Explore Use Cases**  
Explore curated list of AI use cases for your organization
-  **Discover Success Stories**  
Discover how organizations across the world are using AI to drive business outcomes
-  **Mobilize Your Team**  
Follow our expert curated action plan to realized the power of AI

## Explore The Art Of The Possible In AI

 Search by industry, business function, or desired business outcome

Industry ▾ Business Function ▾ Business Outcome ▾ **Explore**



# Generative AI application



# Generative AI stack

## Applications that leverage LLMs and other foundation models (FMs)



Amazon Q  
Business



Amazon Q  
Developer



Amazon Q in  
QuickSight



Amazon Q in  
Connect

## Tools to build with LLMs and other FMs



**Amazon Bedrock**

Guardrails

| Agents

| Studio

| Customization capabilities

| Custom Model Import

## Infrastructure for FM training and inference



GPUs



AWS Trainium



AWS Inferentia



Amazon SageMaker



Amazon EC2  
UltraClusters



Elastic Fabric  
Adapter (EFA)



Amazon EC2  
Capacity Blocks



AWS Nitro



AWS Neuron

# Reinvent work with Amazon Q

## BUSINESS

---

 Amazon Q  
Business

Knowledge search

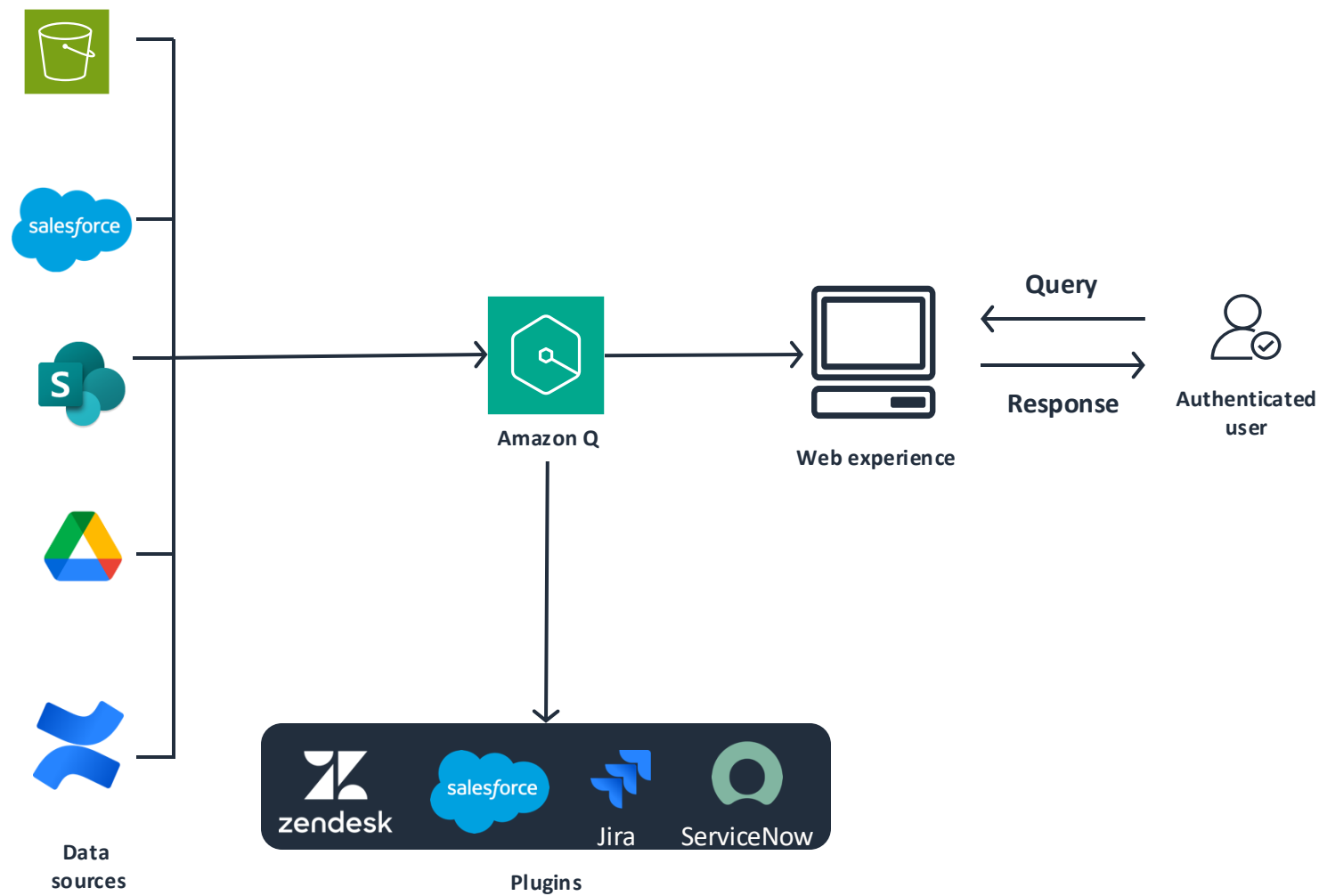
Summarization

Content creation

Extract insights

Research & analysis

# Amazon Q for business




# Amazon Q for business

Chat

Apps

Library

Info



## Amazon Q App Library

Get started with a collection of published Q Apps

All Q Apps

Finance

General

HR

IT

Legal

Marketing

Operations

Sales

Supply chain

Support

Outreach assistant

Generates tailored customer outreach emails

Marketing

Sales

Published on  
March 20, 2024

132

Open

Transformer troubleshooter

References the transformer troubleshooter guide

Operations

Support

Published on  
March 20, 2024

130

Open

Content creator

Crafts targeted marketing content

Marketing

Published on  
March 20, 2024

129

Open

HR Assistant

Advises on HR related topics

HR

Support

March 20, 2024

125

Open

Newsletter Generator

Generates tailored newsletters

General

Operations

Published on  
March 20, 2024

123

Open

Debugging Assistant

Assists and advises in debugging

IT

Published on  
March 20, 2024

122

Open

Task Manager

Prioritizes tasks

Incident Processing Planner

Processes incidents

Promo Doc Assistant

Generates content for promotion documents

© 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.



# Reinvent work with Amazon Q

## BUSINESS

 Amazon Q  
Business


Knowledge search

Summarization

Content creation

Extract insights

Research & analysis

 Amazon Q in  
QuickSight

Understand data

Build & refine visuals

Build calculations

Executive summaries

Create data stories



Data

×

Dataset100%

SPICE Q - SaaS Sales

Search fields

+ CALCULATED FIELD

Customer Location

- Country
- City
- Contact Name
- cstmr\_nm
- Customer ID
- Date Key
- Discount
- Industry
- License
- Order Date
- Order ID
- Product
- Profit
- Quantity
- Region
- Row ID
- Sales
- Segment
- Subregion

Visuals

×

+ ADD

Select or add a visual



Sales Overview

+

Properties

×

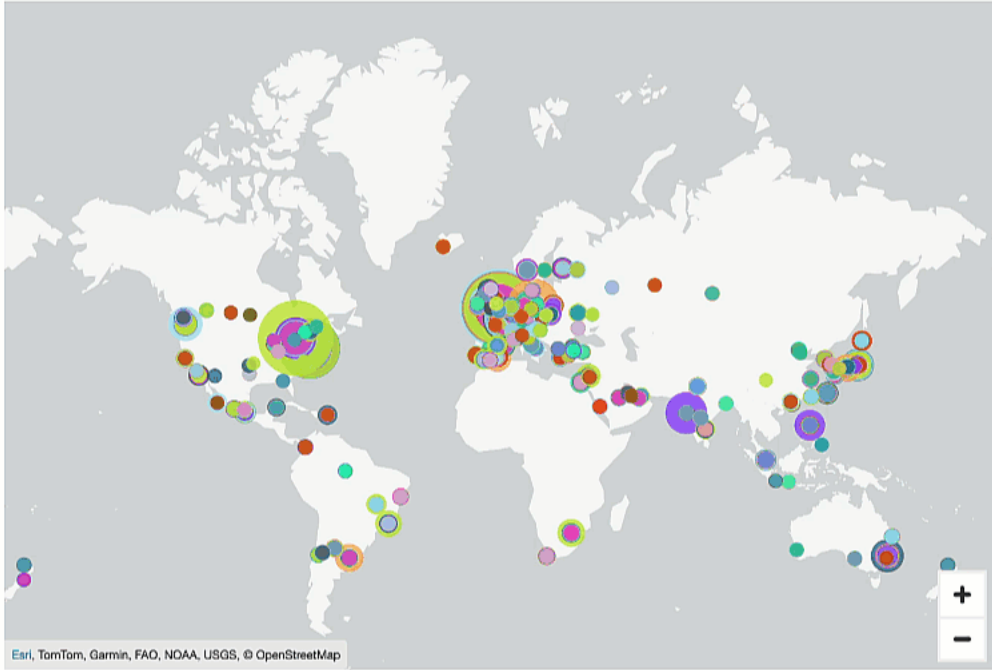
Select a visual or control



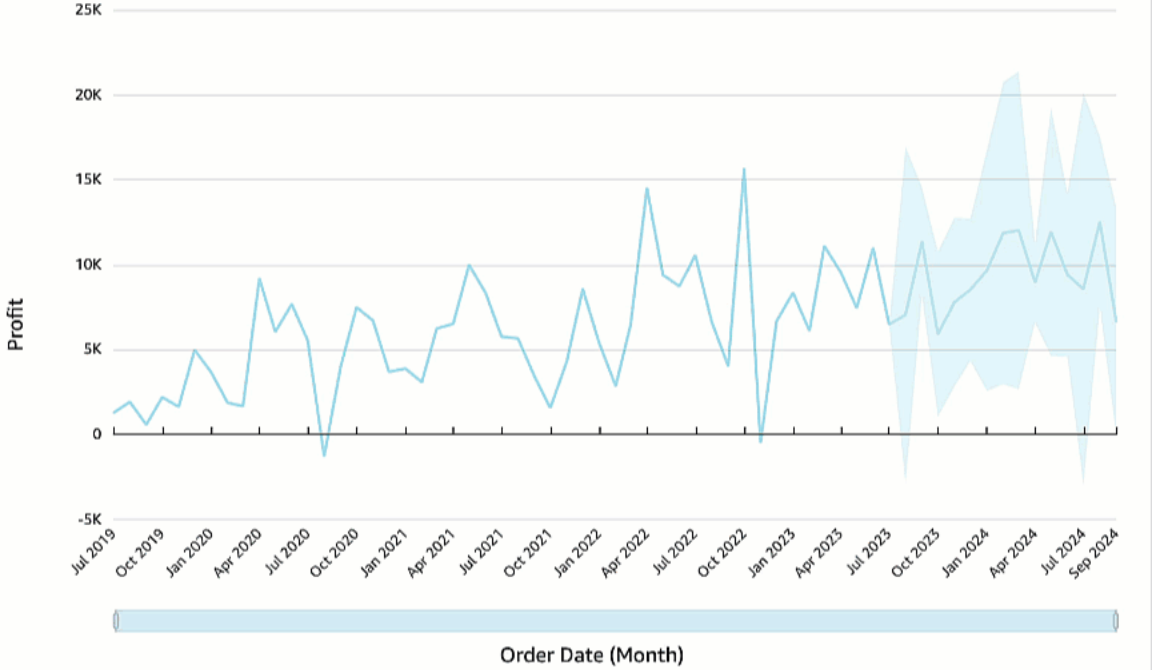
🗨️ Ask a question about SaaS Sales

Total Sales by City and Product

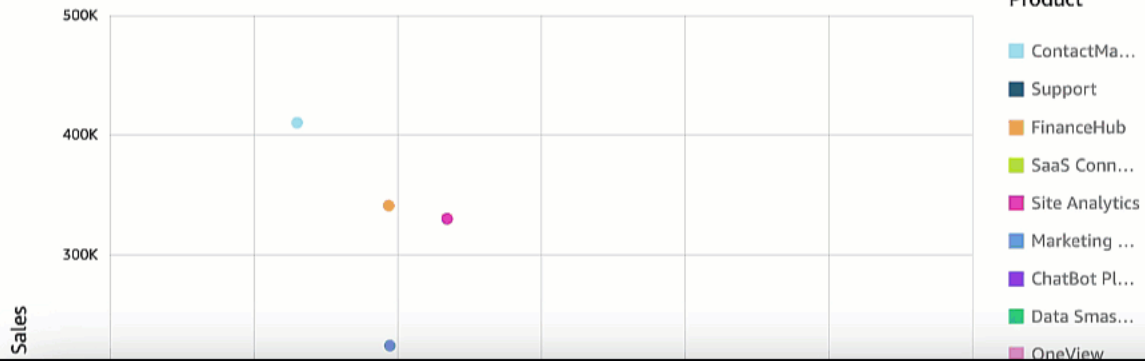
SHOWING TOP 262 IN CITY AND TOP 14 IN PRODUCT



Forecast of Total Profit by month







Profit Ratio and total Sales by Product



Maximum Sales by Region

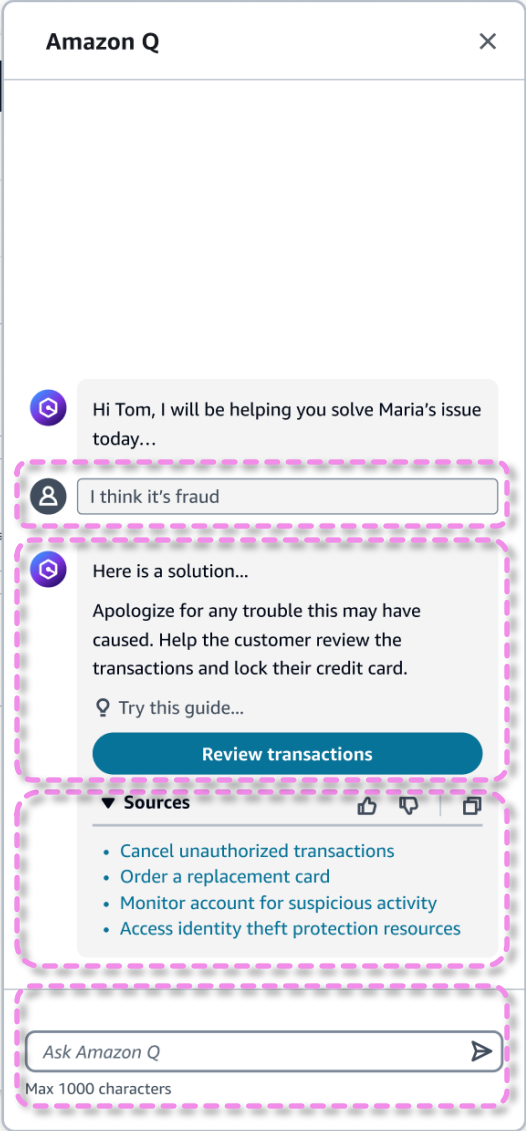
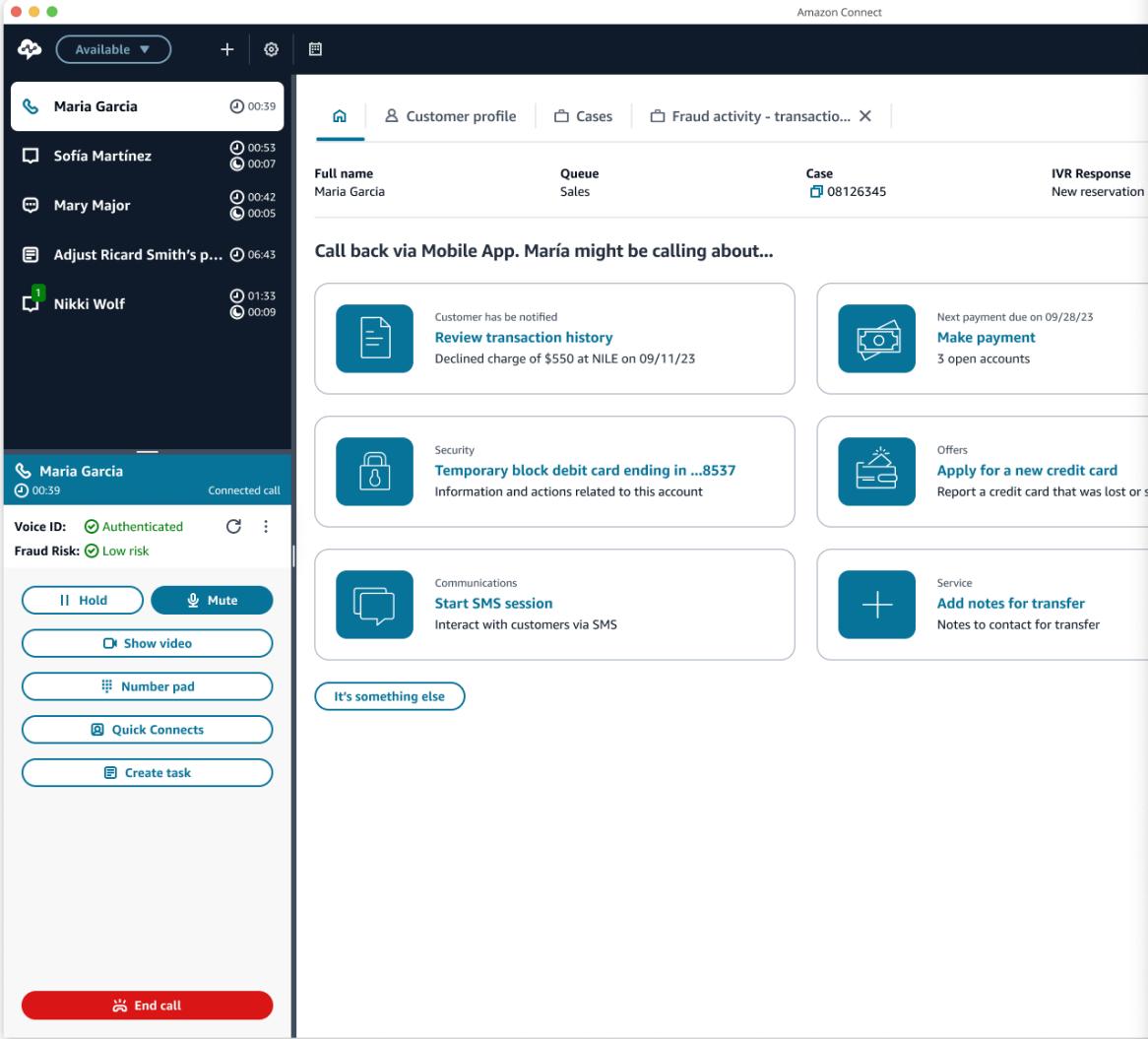


# Reinvent work with **Amazon Q**

BUSINESS		SPECIALIZED USERS
 Amazon Q Business	 Amazon Q in QuickSight	 Amazon Q in Connect
Knowledge search	Understand data	Agent assist
Summarization	Build & refine visuals	
Content creation	Build calculations	
Extract insights	Executive summaries	 Amazon Q in AWS Supply Chain
Research & analysis	Create data stories	Supply chain



# Amazon Q in Connect



Detected or provided issue






Generated solution

Documents used to generate the response & solution

Agents can chat with Amazon Q for additional information



# Reinvent work with Amazon Q

BUSINESS		SPECIALIZED USERS		DEVELOPERS	
 Amazon Q Business	 Amazon Q in QuickSight	 Amazon Q in Connect	 Amazon Q Developer		
Knowledge search	Understand data	Agent assist	Plan application		
Summarization	Build & refine visuals		Code generation		
Content creation	Build calculations		Unit testing		
Extract insights	Executive summaries	 Amazon Q in AWS supply chain	Security scanning		
Research & analysis	Create data stories	Supply chain	Code remediation		
			Code migration		
			Troubleshooting		
			Developer knowledge		

# Amazon Q Developer supports developers across the SDLC

## Plan

- Amazon Q Developer in the AWS Management Console (best practices, AWS Well-Architected Framework, Amazon EC2 instance optimization)
- Business-specific application
- Explain code with **conversational coding**

## Create

- Inline coding companion in IDE and CLI
- Software development
- **Conversational coding**

## Test and secure

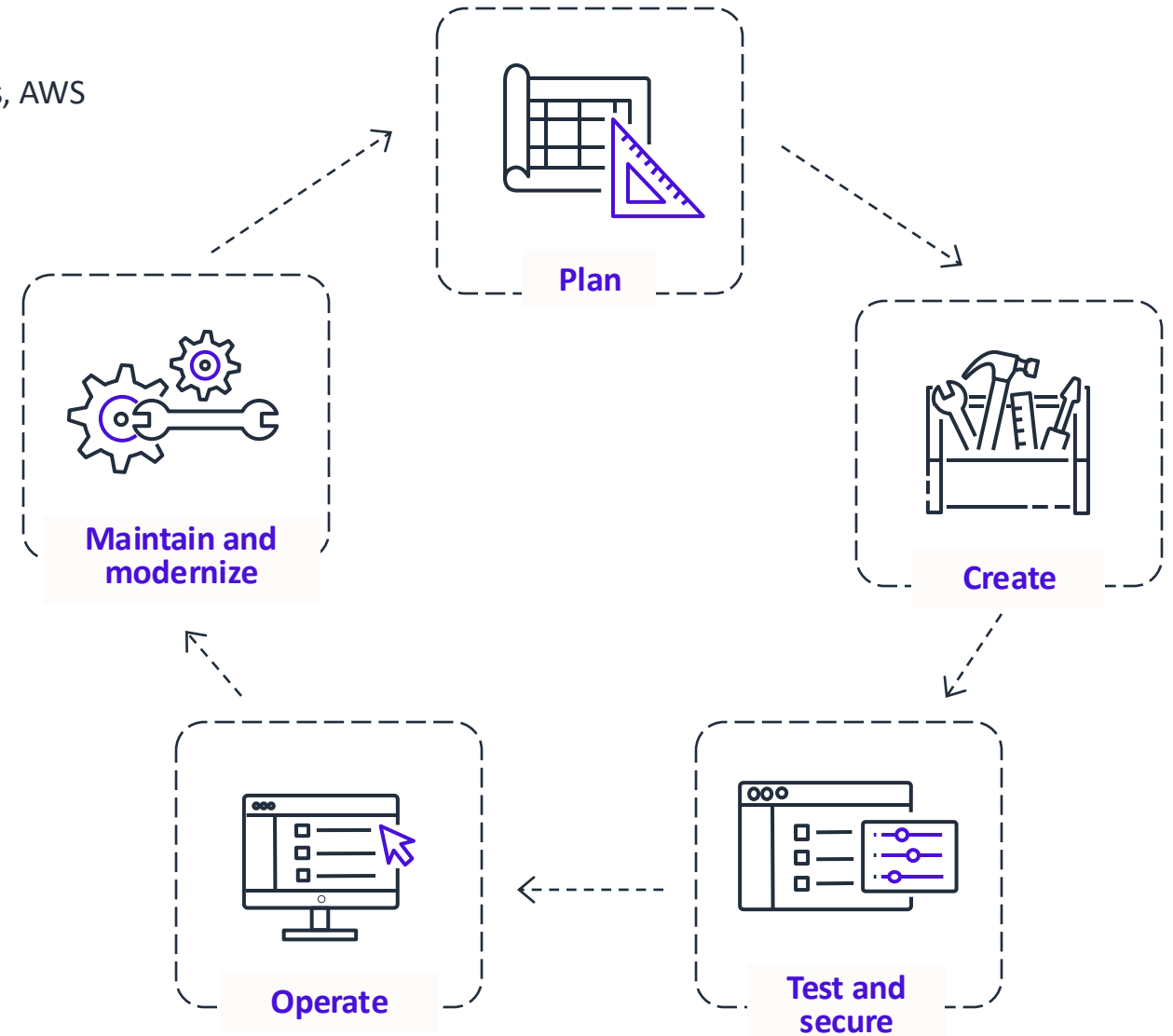
- Unit test generation
- Security scanning and remediation

## Operate

- Troubleshoot errors (Amazon S3, AWS Lambda, Amazon EC2, Amazon Elastic Container Service)
- VPC Reachability Analyzer
- Debug and optimize code with **conversational coding**
- **Helps you optimize your AWS resources and costs [NEW]**

## Maintain and modernize

- Update code with Amazon Q Developer agent for code transformation



# Generative AI stack

## Applications that leverage LLMs and other foundation models (FMs)



Amazon Q  
Business



Amazon Q  
Developer



Amazon Q in  
QuickSight



Amazon Q in  
Connect

## Tools to build with LLMs and other FMs



**Amazon Bedrock**

Guardrails

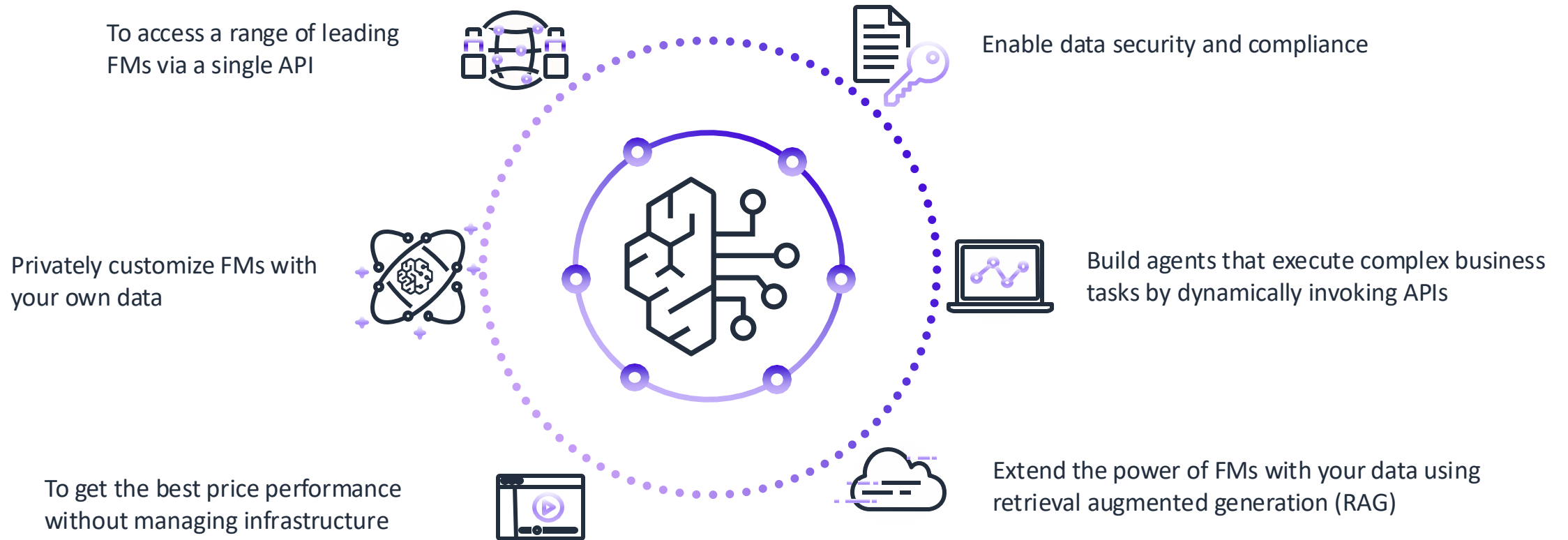
| Agents

| Studio

| Customization capabilities

| Custom model import

# When to use Amazon Bedrock?



# Amazon Bedrock

BROADEST SELECTION OF FULLY MANAGED MODELS FROM LEADING AI COMPANIES

**AI21labs**

Effective reasoning & rapid analysis for long context windows

JAMBA

**amazon**

Frontier intelligence & industry leading price performance, high-quality image & video generation

AMAZON NOVA

**ANTHROPIC**

Extended thinking & coding, including computer use skills

CLAUDE

**cohere**

Multimodal search & advanced retrieval powering multilingual knowledge agents

COMMAND  
EMBED  
RERANK

**deepseek**

Advanced reasoning, comprehension, scientific & mathematical analysis, & code generation

DEEPSEEK-R1

**Luma**

High-quality video generation with natural, coherent motion & ultra-realistic details

LUMA RAY 2

**Meta**

Advanced image & language reasoning

LLAMA

**MISTRAL AI**

Knowledge summarization, expert agents, & code completion

MISTRAL  
MIXTRAL

**poolside**

Software engineering AI for large enterprises

MALIBU  
POINT

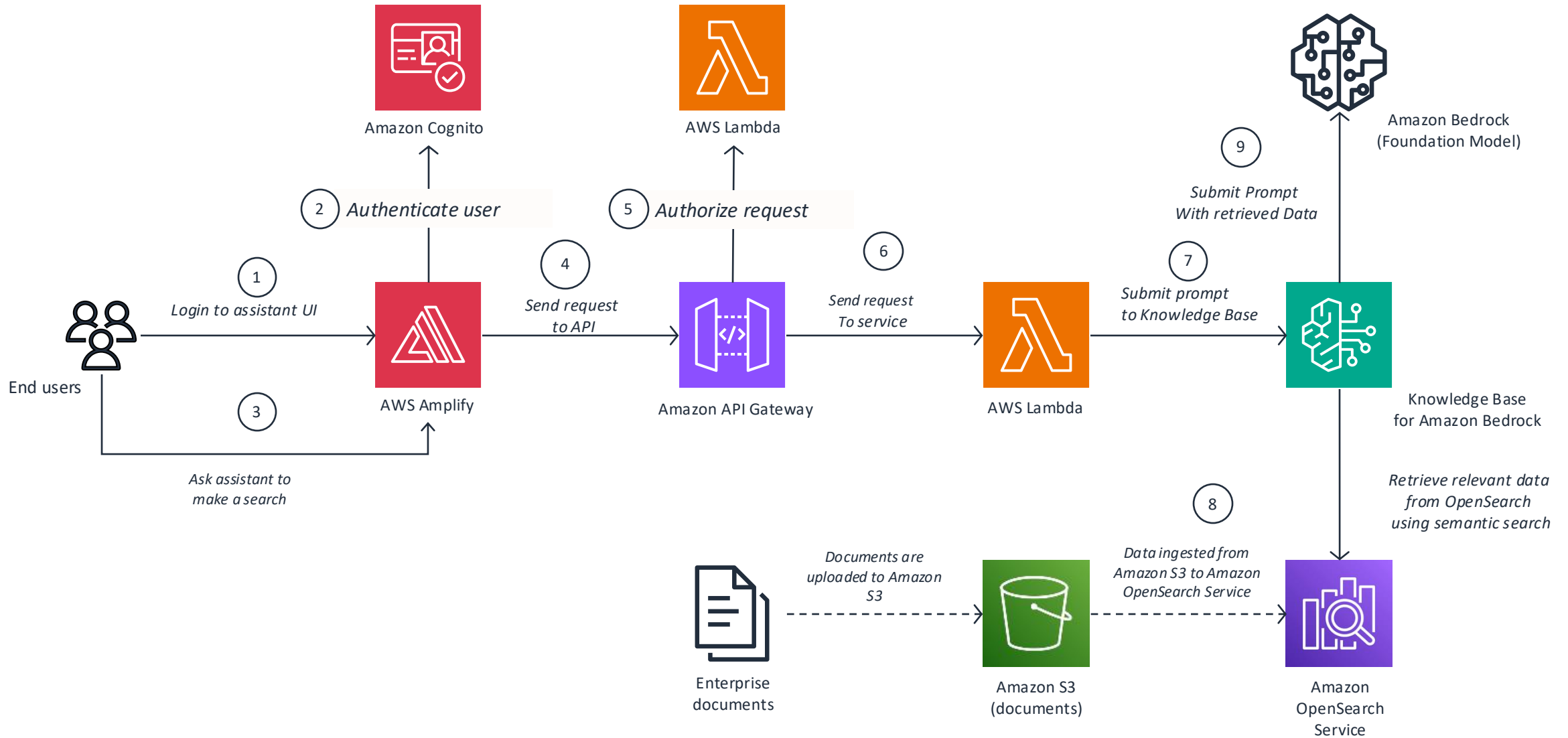
**stability.ai**

High-quality AI image generation, easily deployable at scale

STABLE DIFFUSION

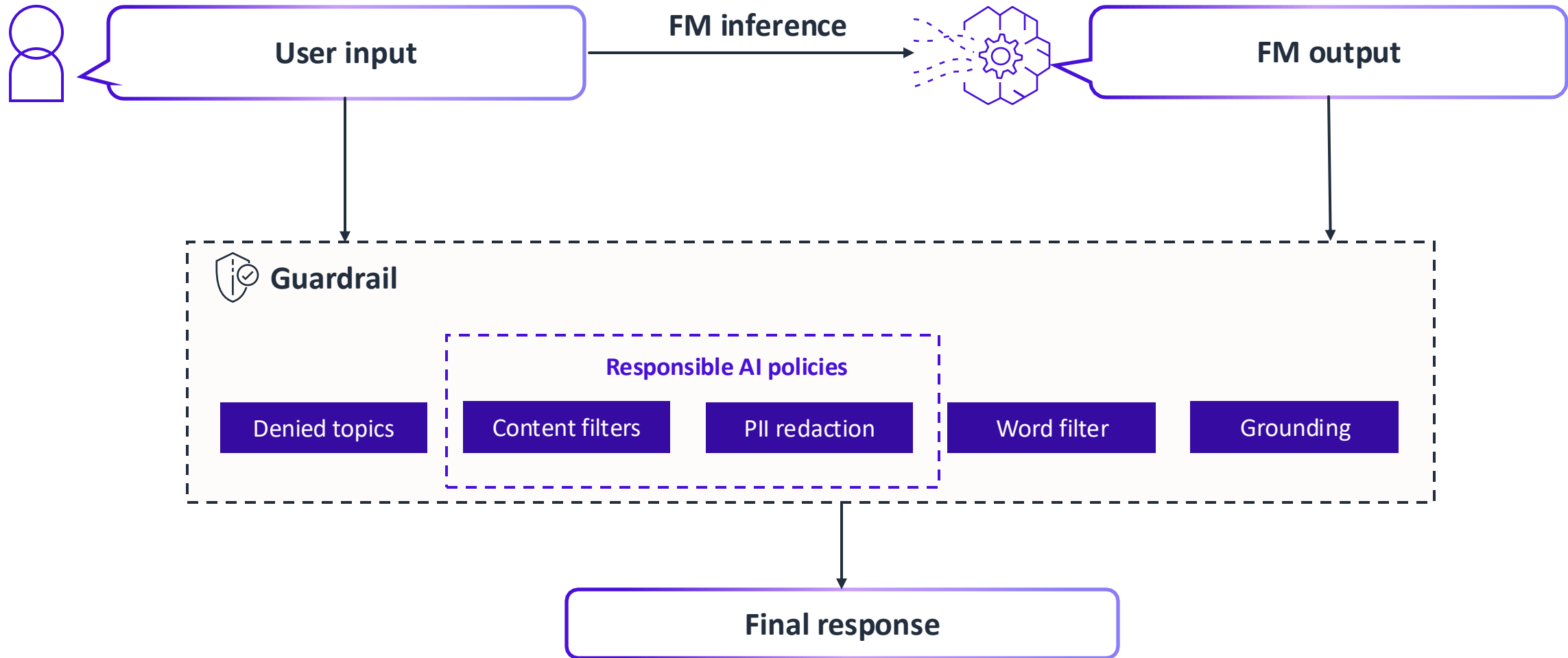


# Amazon Bedrock – Digital assistant example





# Amazon Bedrock - Guardrails



# Generative AI stack

## Applications that leverage LLMs and other foundation models (FMs)



Amazon Q  
Business



Amazon Q  
Developer



Amazon Q in  
QuickSight



Amazon Q in  
Connect

## Tools to build with LLMs and other FMs



**Amazon Bedrock**

Guardrails

| Agents

| Studio

| Customization capabilities

| Custom model import

## Infrastructure for FM training and inference



GPUs



AWS Trainium



AWS Inferentia



Amazon SageMaker



Amazon EC2  
UltraClusters



Elastic Fabric  
Adapter (EFA)



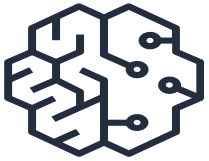
Amazon EC2  
Capacity Blocks



AWS Nitro



AWS Neuron



































# Amazon SageMaker

Build, train, and deploy ML models  
at scale, including FMs

- Access the latest and publicly available FMs
- Build FMs from scratch
- Customize FMs
- Run inference
- Implement FMOps and governance

# Build with the broadest selection of foundation models

Available in Amazon SageMaker JumpStart

  <b>HuggingFace</b> Explore hundreds of popular and trending models from HuggingFace. <a href="#">View 368 models &gt;</a>	  <b>Meta</b> Explore popular and trending models from Meta including Llama, Code Llama, and more. <a href="#">View 44 models &gt;</a>	  <b>AI21</b> Explore popular and trending models from AI21 Labs including Jurassic and more. <a href="#">View 6 models &gt;</a>	  <b>stability ai</b> Explore popular and trending models from Stability.ai including Stable Diffusion and more. <a href="#">View 11 models &gt;</a>
  <b>cohere</b> Explore popular and trending models from Cohere including Command, Rerank, and more. <a href="#">View 12 models &gt;</a>	  <b>TensorFlow</b> Explore popular and trending models from TensorFlow for computer vision and NLP tasks. <a href="#">View 319 models &gt;</a>	  <b>PyTorch</b> Explore popular and trending models from PyTorch for computer vision and NLP tasks. <a href="#">View 34 models &gt;</a>	  <b>upstage</b> Explore popular and trending models from Upstage including Solar mini chat model and more. <a href="#">View 4 models &gt;</a>
  <b>LightOn</b> Explore popular and trending models from LightOn including mini-instruct models. <a href="#">View 2 models &gt;</a>	  <b>NCSOFT</b> Explore popular and trending models from NCSOFT including VARCO LLM models. <a href="#">View 3 models &gt;</a>	  <b>LG AI Research</b> Explore popular and trending models from LG CNS including EXAONE Atelier and more. <a href="#">View 1 models &gt;</a>	  <b>Jina AI</b> Explore popular and trending models from Jina AI including Jina Embeddings model and more. <a href="#">View 1 models &gt;</a>
  <b>VOYAGE AI</b> Explore popular and trending models from Voyage AI including Voyage-2 Embedding Model and more. <a href="#">View 3 models &gt;</a>	  <b>Nomic</b> Explore popular and trending models from Nomic including Nomic Embed Model and more. <a href="#">View 2 models &gt;</a>	  <b>EvolutionaryScale, PBC</b> Explore frontier, generative language models for biology from EvolutionaryScale, including ESM3. <a href="#">View 1 models &gt;</a>	  <b>aws</b> Explore popular and trending models from AWS for computer vision, NLP, and tabular tasks. <a href="#">View 36 models &gt;</a>

# Generative AI stack

## Applications that leverage LLMs and other foundation models (FMs)



Amazon Q  
Business



Amazon Q  
Developer



Amazon Q in  
QuickSight



Amazon Q in  
Connect

## Tools to build with LLMs and other FMs



**Amazon Bedrock**

Guardrails

| Agents

| Studio

| Customization capabilities

| Custom model import

## Infrastructure for FM training and inference



GPUs



AWS Trainium



AWS Inferentia



Amazon SageMaker



Amazon EC2  
UltraClusters



Elastic Fabric  
Adapter (EFA)



Amazon EC2  
Capacity Blocks



AWS Nitro



AWS Neuron



# Generative AI application



# Generative AI application

## Data foundation

Storage

Governance  
& compliance

Databases,  
analytics,  
& data lakes

Data integration



# End-to-end data foundation



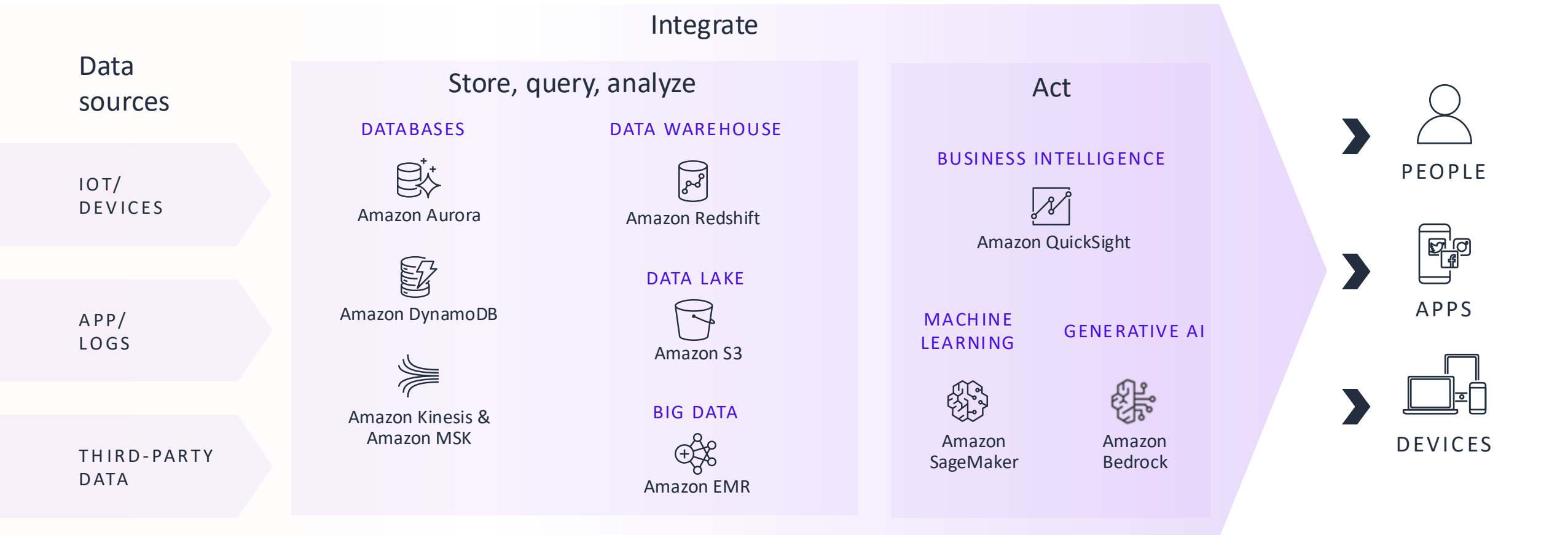
Comprehensive



Integrated



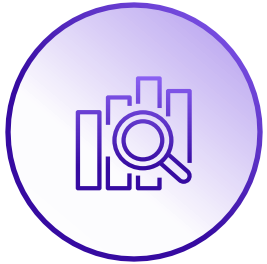
Governed



Catalog and govern | AWS Lake Formation, Amazon DataZone



# Enabling vector support across our services



AMAZON OPENSEARCH  
SERVICE



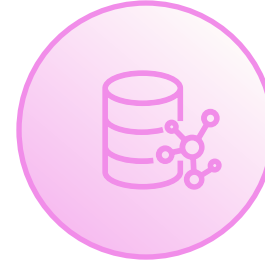
AMAZON AURORA  
POSTGRESQL



AMAZON RDS  
FOR POSTGRESQL



AMAZON MEMORYDB



AMAZON NEPTUNE



AMAZON DOCUMENTDB

# Responsible AI

## **FAIRNESS**

Considering impacts on different groups of stakeholders

## **EXPLAINABILITY**

Understanding and evaluating system outputs

## **CONTROLLABILITY**

Having mechanisms to monitor and steer AI system behavior

## **SAFETY**

Preventing harmful system output and misuse

## **PRIVACY & SECURITY**

Appropriately obtaining, using and protecting data and models

## **GOVERNANCE**

Incorporating best practices into the AI supply chain, including providers and deployers

## **TRANSPARENCY**

Enabling stakeholders to make informed choices about their engagement with an AI system

## **VERACITY & ROBUSTNESS**

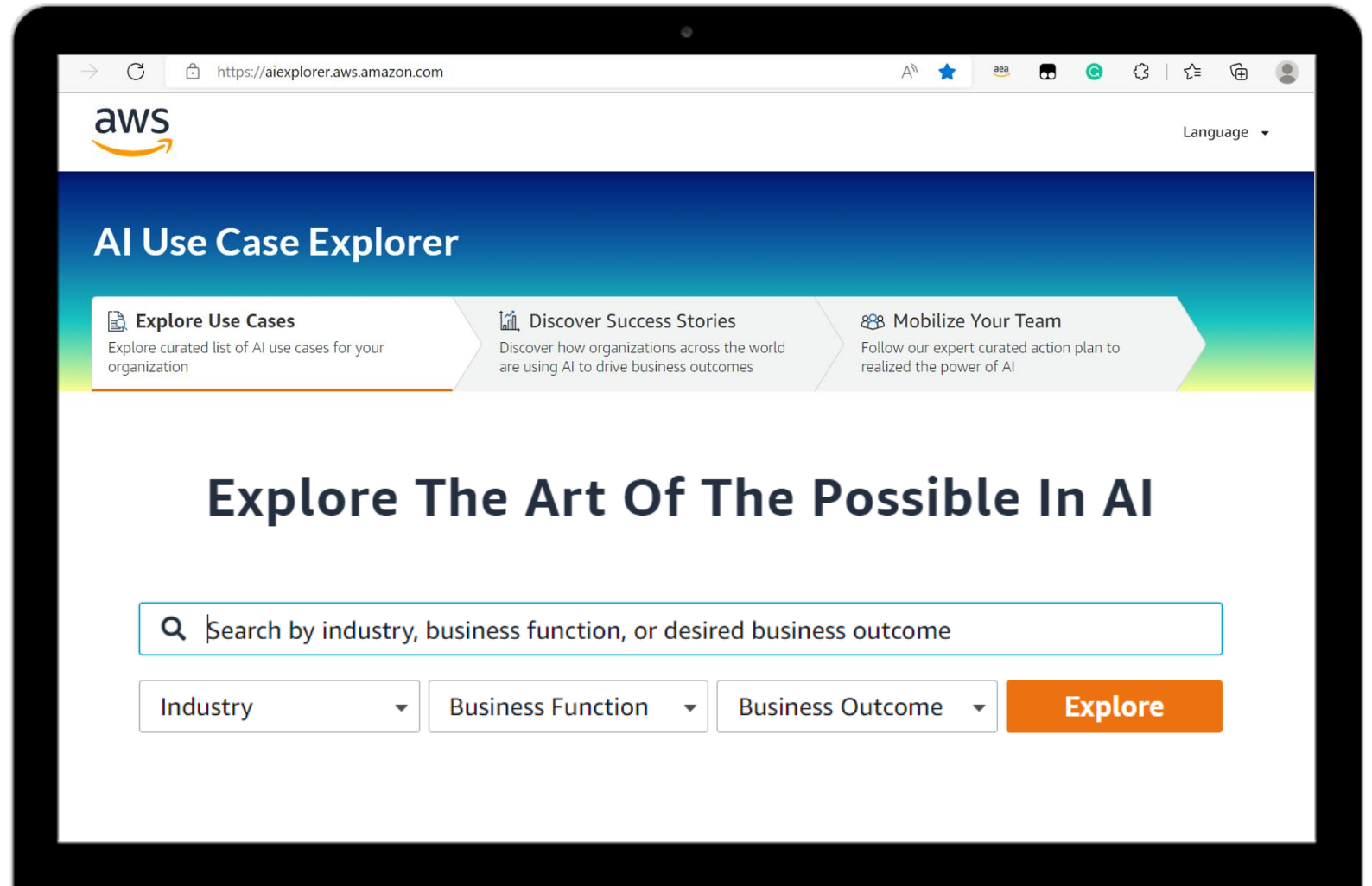
Achieving correct system outputs, even with unexpected or adversarial inputs

aiexplorer.aws.amazon.com



# AI Use Case Explorer

Easily find the most relevant AI use cases with related content and guidance to make them real



# Dive deeper with AWS migrate and modernize resources

Unlock the full potential of the cloud with AWS—  
learn how to seamlessly migrate your workloads and  
modernize with tools, best practices, and expert support.

[Visit the Resource Center via event platform homepage »](#)



# AWS Training & Certification

Access 600+ free digital courses with AWS Skill Builder

Focus on the cloud skills and services that are most relevant to you across 30+ AWS solutions, including digital self-paced learning plans and ramp-up guides

- Build your future in the AWS Cloud at your own pace
- Advance your skills and knowledge with learning plans
- Validate your cloud expertise with AWS Certification



<https://shorturl.at/22v8J>

Learn your way, explore [skillbuilder.aws](https://skillbuilder.aws) >>



# Thank you for attending **AWS Builders Online Series**

We hope you found it interesting! A kind reminder to complete the survey.  
Let us know what you thought of today's event and how we can improve the event experience for you in the future.



[aws-apj-marketing@amazon.com](mailto:aws-apj-marketing@amazon.com)



[twitter.com/AWSCloud](https://twitter.com/AWSCloud)



[facebook.com/AmazonWebServices](https://facebook.com/AmazonWebServices)



[youtube.com/user/AmazonWebServices](https://youtube.com/user/AmazonWebServices)



[linkedin.com/company/amazon-web-services](https://linkedin.com/company/amazon-web-services)



[twitch.tv/aws](https://twitch.tv/aws)



# Thank you!

**Melwin Pais**

Senior Solutions Architect  
AWS

