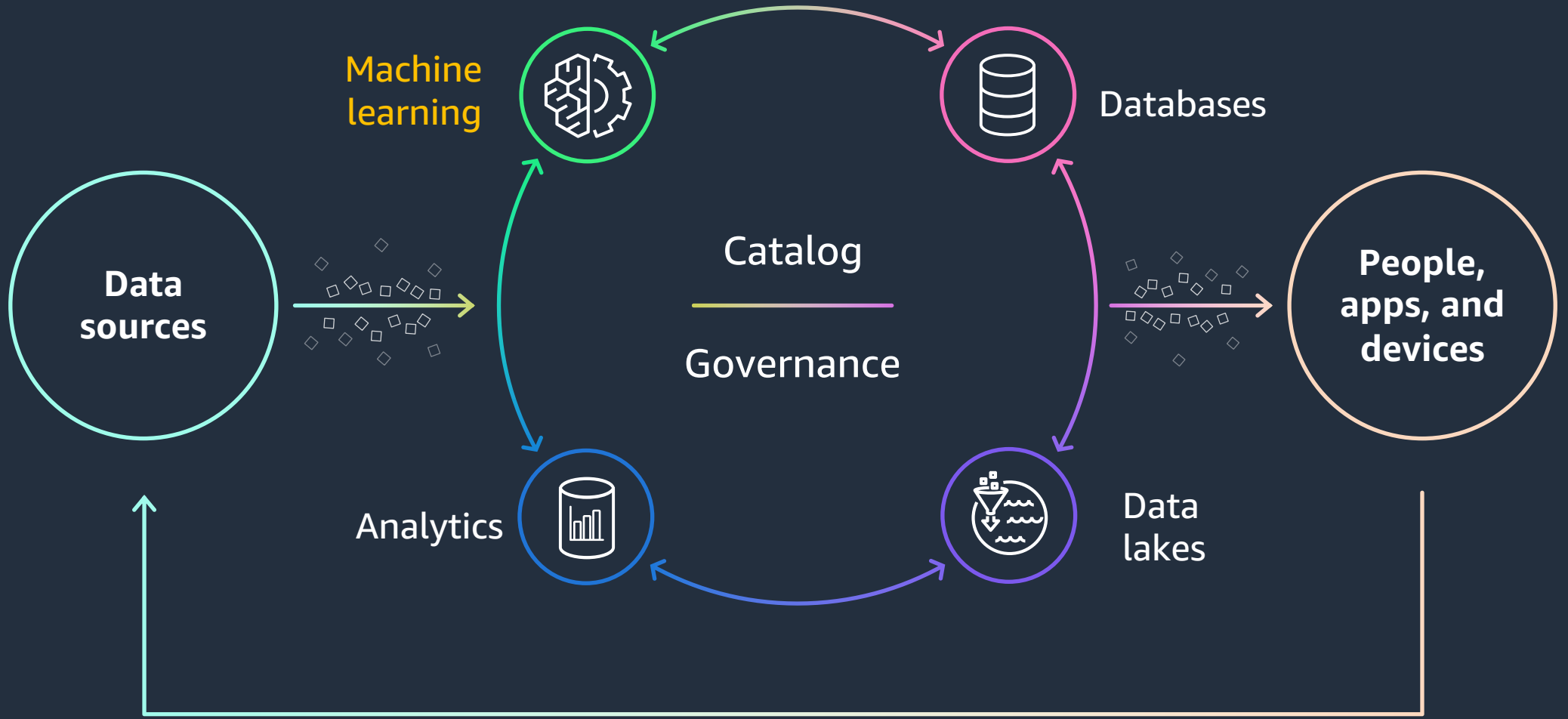
MIAMI LEARNING DAY
DATA AND ANALYTICS TRACK

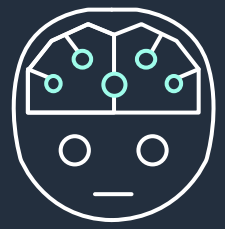
AI/ML for Data and Analytics

Hiren Deliwala
Solutions Architect Leader
State & Local Government/Education
hdaws@amazon.com

Modern data strategy in action



AI/ML/Gen AI Hierarchy



Artificial intelligence (AI)

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

Machine learning (ML)

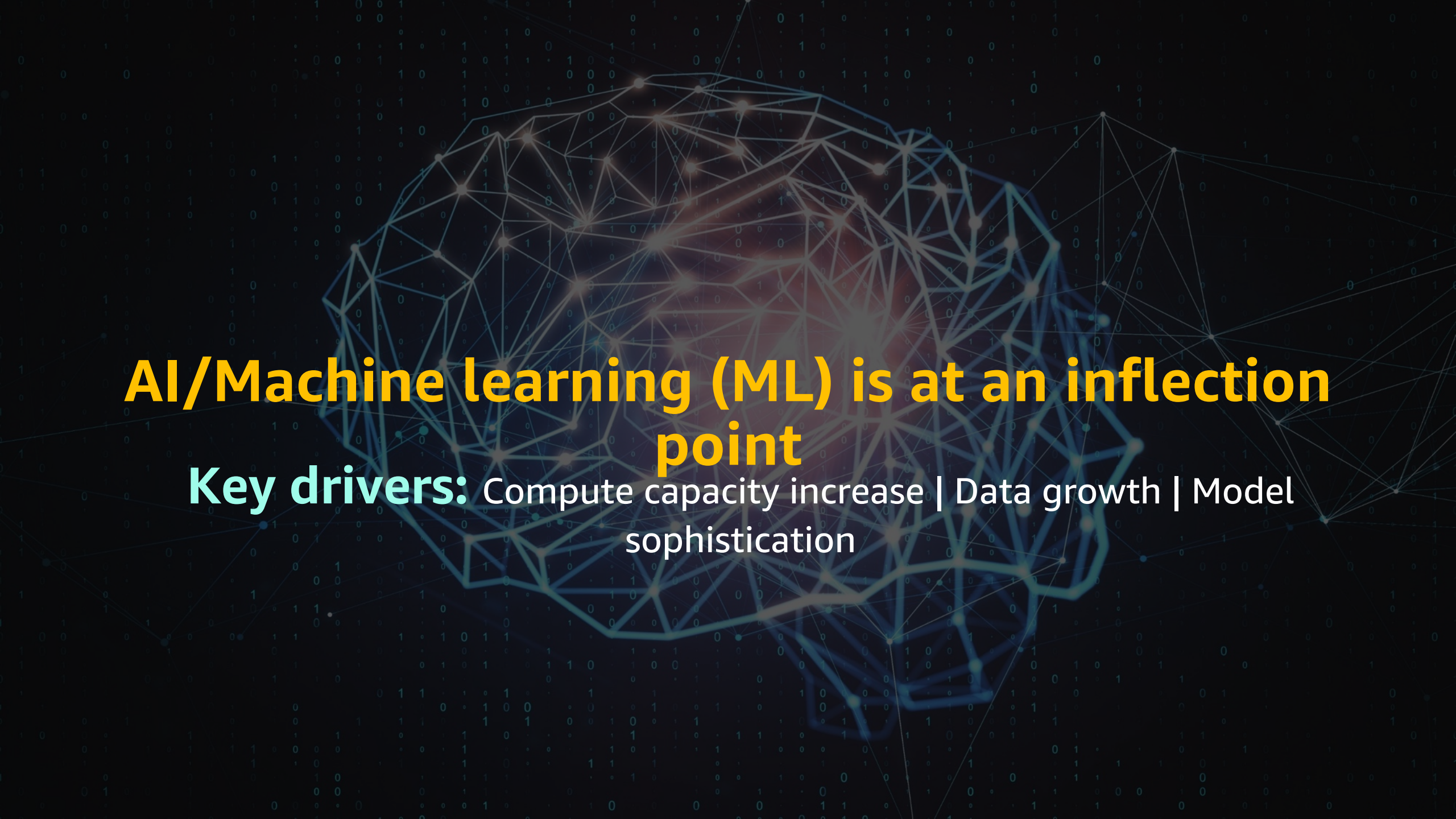
A subset of AI that uses machines to search for patterns in data to build logic models automatically

Deep learning (DL)

A subset of ML composed of deeply multi-layered neural networks that perform tasks like speech and image recognition

Generative AI

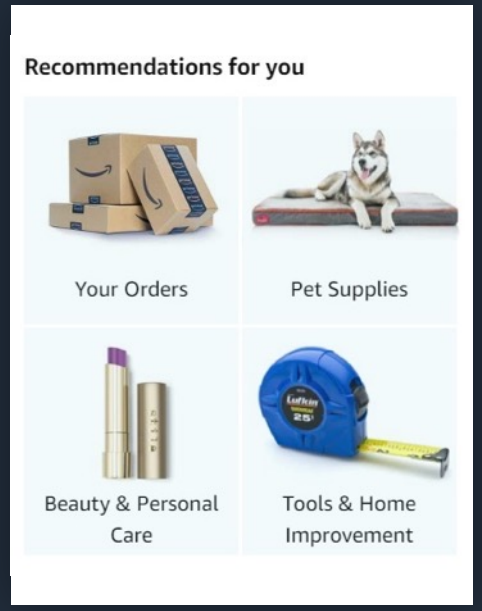
Powered by large models that are pretrained on vast corpora of data and commonly referred to as foundation models (FMs)



AI/Machine learning (ML) is at an inflection point

Key drivers: Compute capacity increase | Data growth | Model sophistication

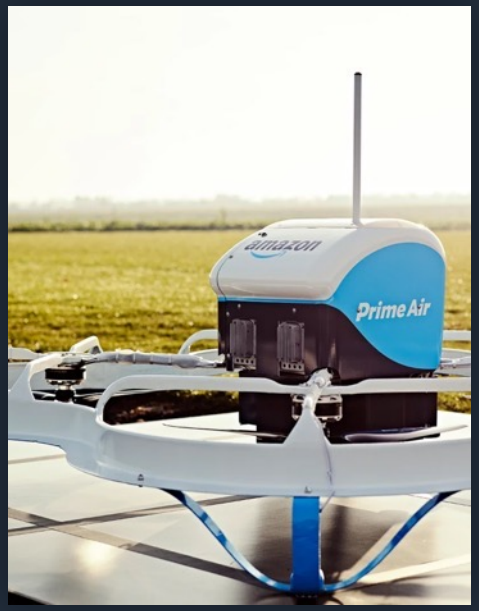
ML innovation is in Amazon's DNA



4,000 products per minute sold on Amazon.com

1.6M packages every day

Billions of Alexa interactions each week



First Prime Air delivery on **December 7, 2016**

More than 100,000 customers use AWS for ML

The Amazon Web Services AI/ML stack

Consumer

AI SERVICES

Generative AI

SCALING GENERATIVE AI
- Amazon Bedrock

AMAZON BUILT FOUNDATIONAL MODEL
- Amazon Titan

Specialized AI Services

BUSINESS PROCESSES
- Amazon Personalize
- Amazon Forecast
- Amazon Fraud Detector
- Amazon Lookout for Metrics

SEARCH
- Amazon Kendra

CONVERSATION
- Amazon Lex
- Contact Lens
- Voice ID
- Amazon Transcribe Call Analytics

CODE + DEVOPS
- Amazon CodeGuru
- Amazon CodeWhisperer
- Amazon DevOps Guru

INDUSTRIAL
- Amazon Monitron
- Amazon Lookout for Equipment
- Amazon Lookout for Vision

HEALTH
- Amazon HealthLake
- Amazon Comprehend Medical
- Amazon Transcribe Medical
- Amazon Omics

Core

TEXT
- Amazon Translate
- Amazon Comprehend

SPEECH
- Amazon Polly
- Amazon Transcribe

VISION
- Amazon Textract
- Amazon Rekognition
- Panorama
- AWS

Tuner

AMAZON SAGEMAKER

JUMPSTART
Use pre-built models in SageMaker

CANVAS
No-code ML for business analysts

STUDIO LAB
Learn ML

GROUND TRUTH
Label data

STUDIO IDE
- Prepare data
- Store features
- Geospatial ML
- Build with notebooks
- Train models
- Tune parameters
- Deploy in production
- Manage and monitor

CI/CD | GOVERNANCE | RESPONSIBLE ML

Provider

ML FRAMEWORKS AND INFRASTRUCTURE

PyTorch, TensorFlow | Amazon EC2 | CPUs | GPUs | AWS Inferentia | AWS Trainium | Habana Gaudi | FPGA

Challenges we are hearing from state and local government customers

Demand for government services is rising while resources and capacity to deliver them **aren't keeping pace**

Citizens increasingly expect government to **provide modern digital experiences** for conducting online transactions

Aging infrastructure for data capture, storage, and management **creates friction** for leveraging data for analytics and machine learning

Complex security, privacy, and compliance requirements create barriers to change and block adoption of many SaaS solutions

Risk averse culture and institutional inertia slow innovation

Machine learning is going mainstream in public sector

Emergency management
Emergency management, emergency response

Finance and administration
Tax, revenue, regulatory/compliance, fraud, budget, purchasing/procurement

Health and benefits
Local hospitals and clinics, public health, child welfare, homelessness, seniors/youth, health/food, environmental

Assessments and permitting
Planning and zoning, assessment, land admin/recorder of deeds, mapping, watershed, housing

Economic development
Grants and benefits administration, workforce development, corporate relations, land clearance/redevelopment

Constituent engagement
Contact center, website, mobile

Enforcement
Police, jails, courts, animal control

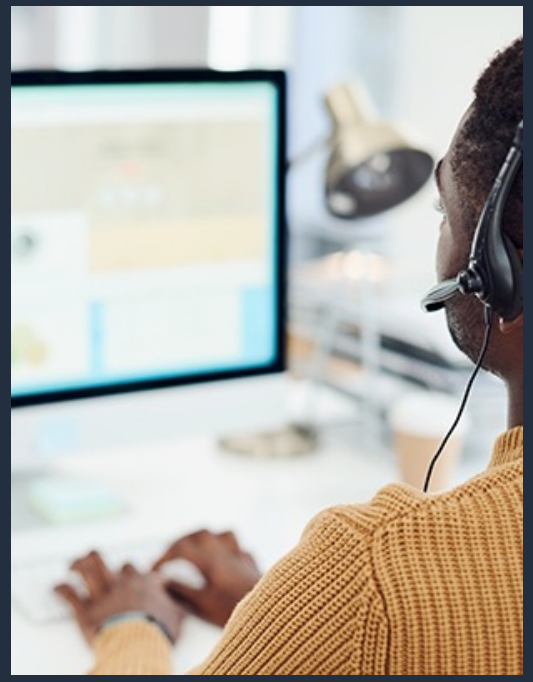
Elections
Registration, voter management, polling, voting management, candidate management

Utilities
Electric/gas/water/sewer, waste management, generation/distribution

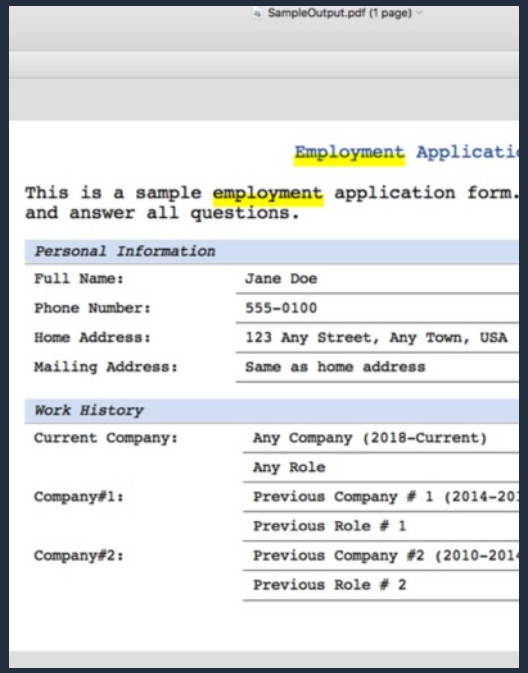
Education and library
K–12 and early childhood, adult education, library

Transportation
Traffic, mass transit, parking, airports, ports

Top AI/ML use cases for state and local government



Speech and language

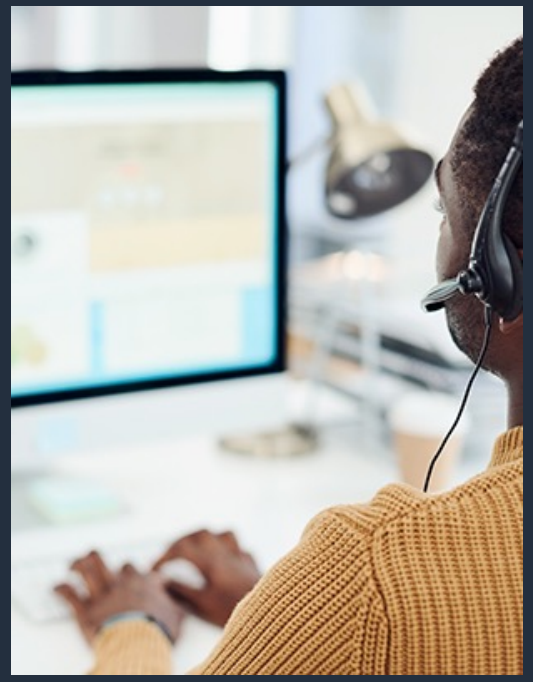


Intelligent document processing

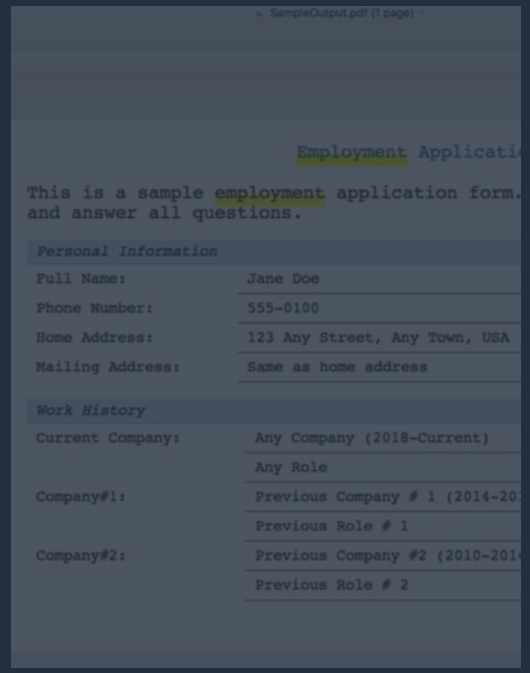
Computer vision

Predictions and insights

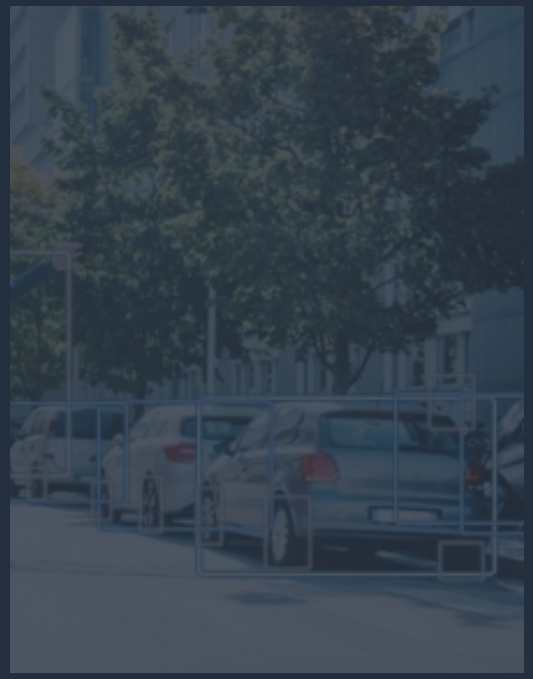
Top AI/ML use cases for state and local government



Speech and language



Intelligent document processing



Computer vision

Predictions and insights

AI/ML-Enabled Citizen engagement

Engage citizens and drive improvements in customer satisfaction

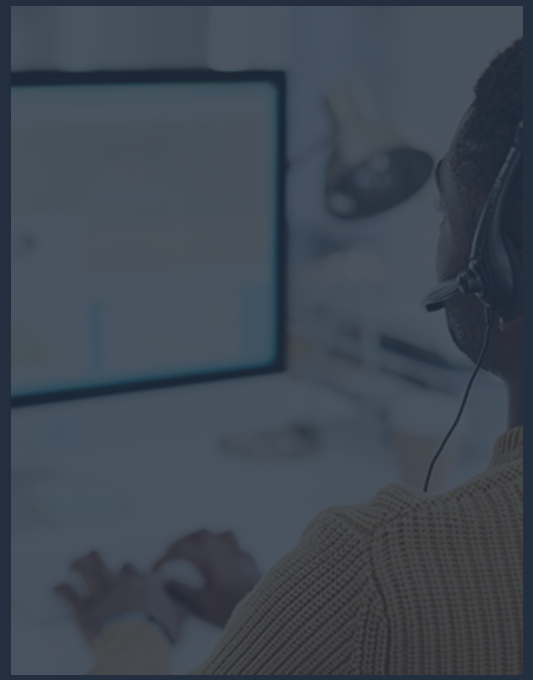
- Improve contact center agent effectiveness with real-time translation and decision support using Amazon Connect and Contact Center Intelligence
- Analyze call and text interactions with citizens to spot issues and trends and drive improvement
- Improve self service

Using AI to improve agent efficiency

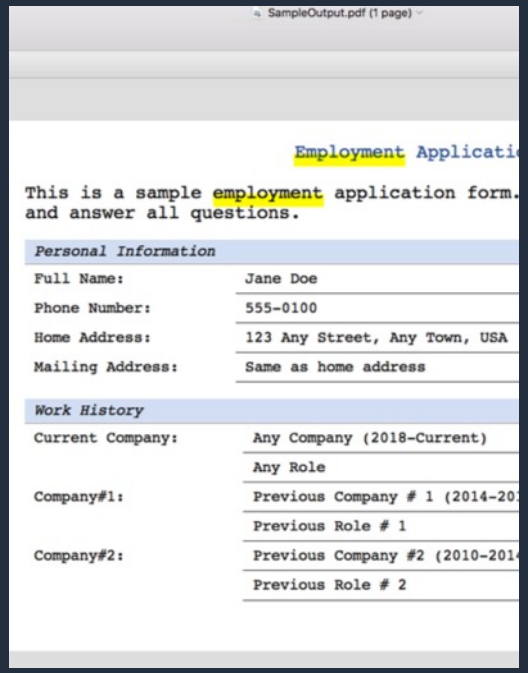
"During peak hours, previously you're 45-50 minutes on hold, and now that's has been reduced to about three and a half minutes. One of the other benefits we've gotten from Amazon Connect is sentiment analysis. On a call, we get real-time feedback on whether or not the customer was happy, frustrated, or angry…"

—Benny Chacko, Deputy General – LA County Internal Services Department

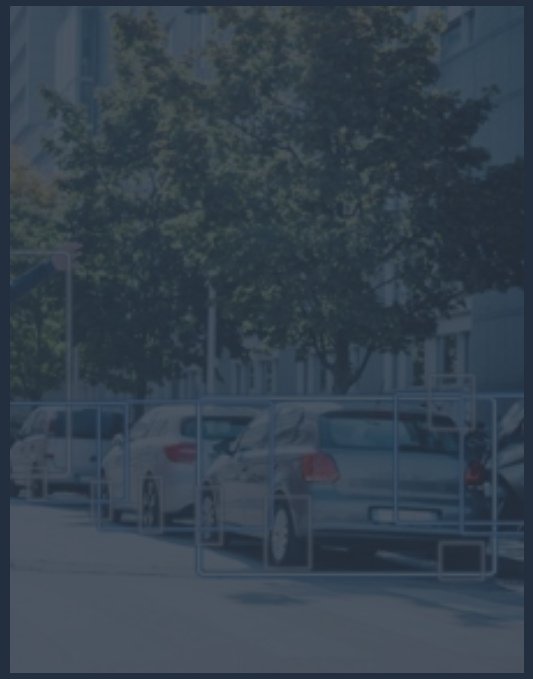
Top AI/ML use cases for state and local government



Speech and language



Intelligent document processing



Computer vision

Predictions and insights

Extract insights from unstructured content

Extract insights from unstructured documents and forms, like images, PDFs, and audio

- Analyze text with natural language processing (NLP) to identify topics, extract entities, understand sentiment, and classify documents with Amazon Textract, Amazon Rekognition, and Amazon Comprehend
- Translate content at scale with Amazon Translate

State Compensation Insurance Fund
Intelligent Document Processing (IDP) for Claims Forms

Business Challenge:

Customer processes 10 million documents per year using manual and semi-manual processes that are inefficient, inaccurate and inflexible. The current OCR platform are insufficient with an accuracy rate of approximately 65% resulting in significant manual intervention and costs.

Approach:

Develop a modern IDP solution based IDP Solutions Guidance to automatically ingest claims forms, properly classify them and ingest form data for processing for up to 6K claims forms per month and expand to support 10M documents per year. Trained models and deployed to Textract and Comprehend using Lambda to control process flow.

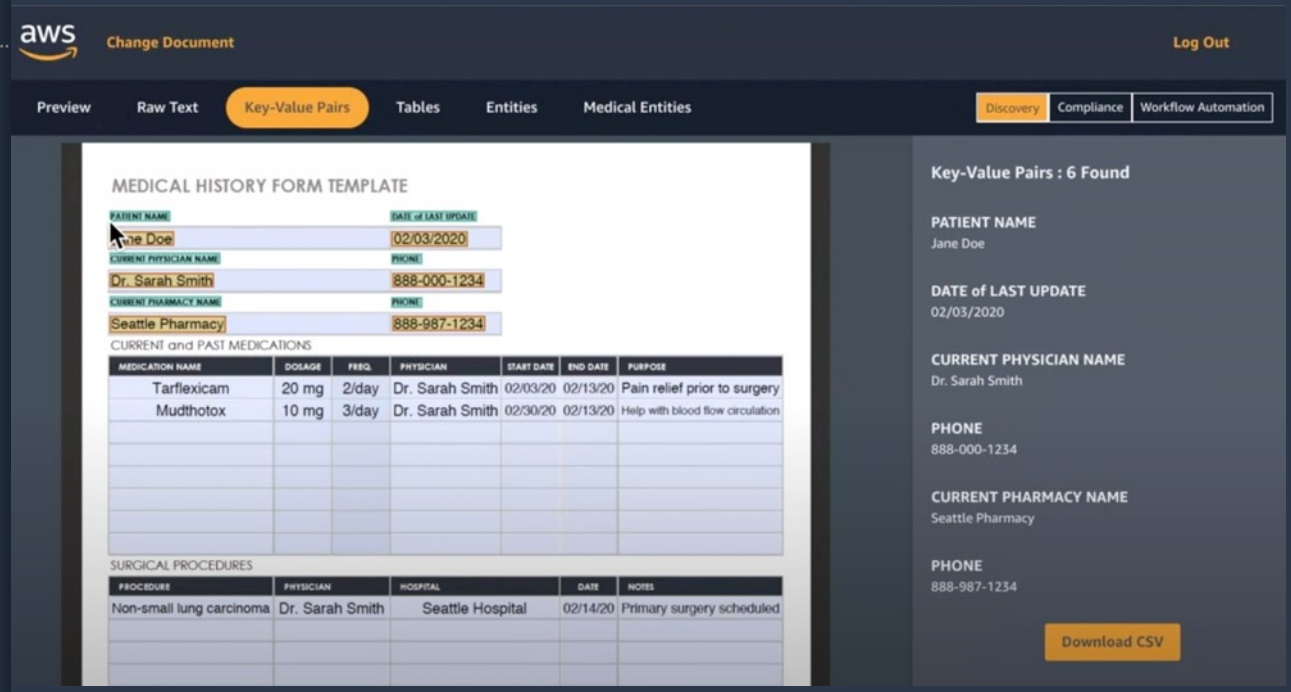
Result:

Successfully demonstrated the ability to identify & classify individual (7) claim form types with 100% accuracy and ingest data into a usable format for customer while routing low confidence (<.9) forms for human-in-the-loop processing at a processing rate of 1,578 per hour. Demonstrated cost/performance benefits of batch vs on-demand processing.

Automate form-based workflows

Reduce human effort for data entry

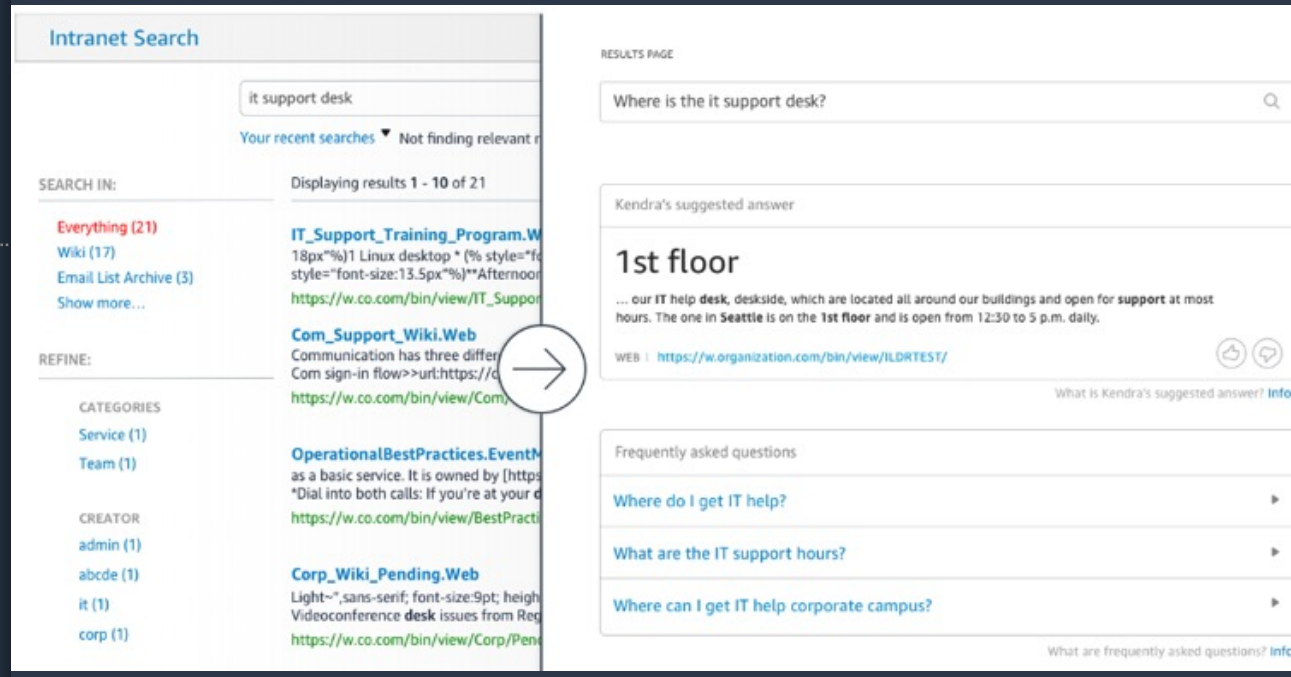
- Extract data and key value pairs from documents and forms with Amazon Textract
- Automate repeatable workflows to improve efficiency
- Redact PII/PHI from sensitive documents with Amazon Comprehend



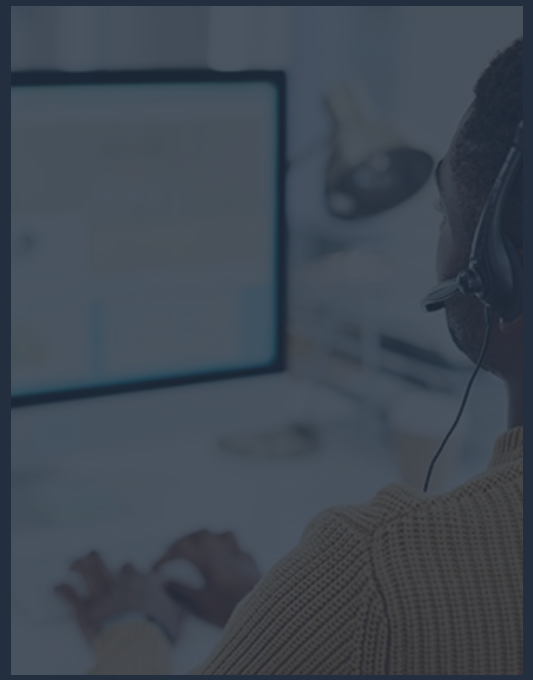
Machine learning-powered enterprise search

Increase employee productivity by quickly and easily finding accurate information and answers with Amazon Kendra

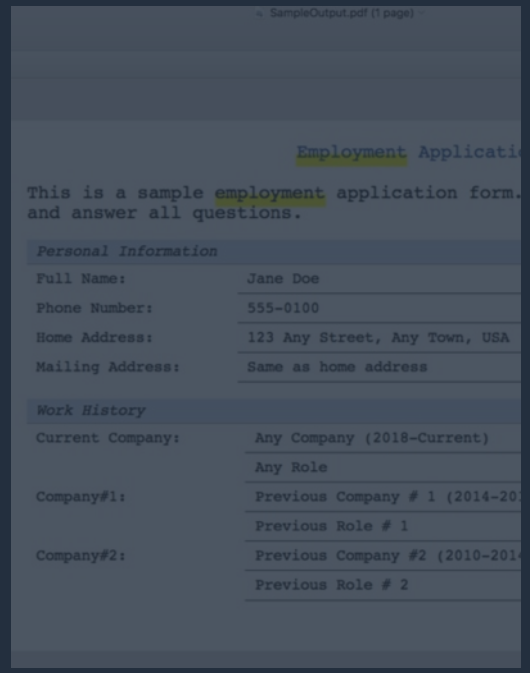
- Understand the intent of queries, not just the keywords
- Return answers, not links and documents
- Index documents and unstructured content at scale



Top AI/ML use cases for state and local government



Speech and language



Intelligent document processing

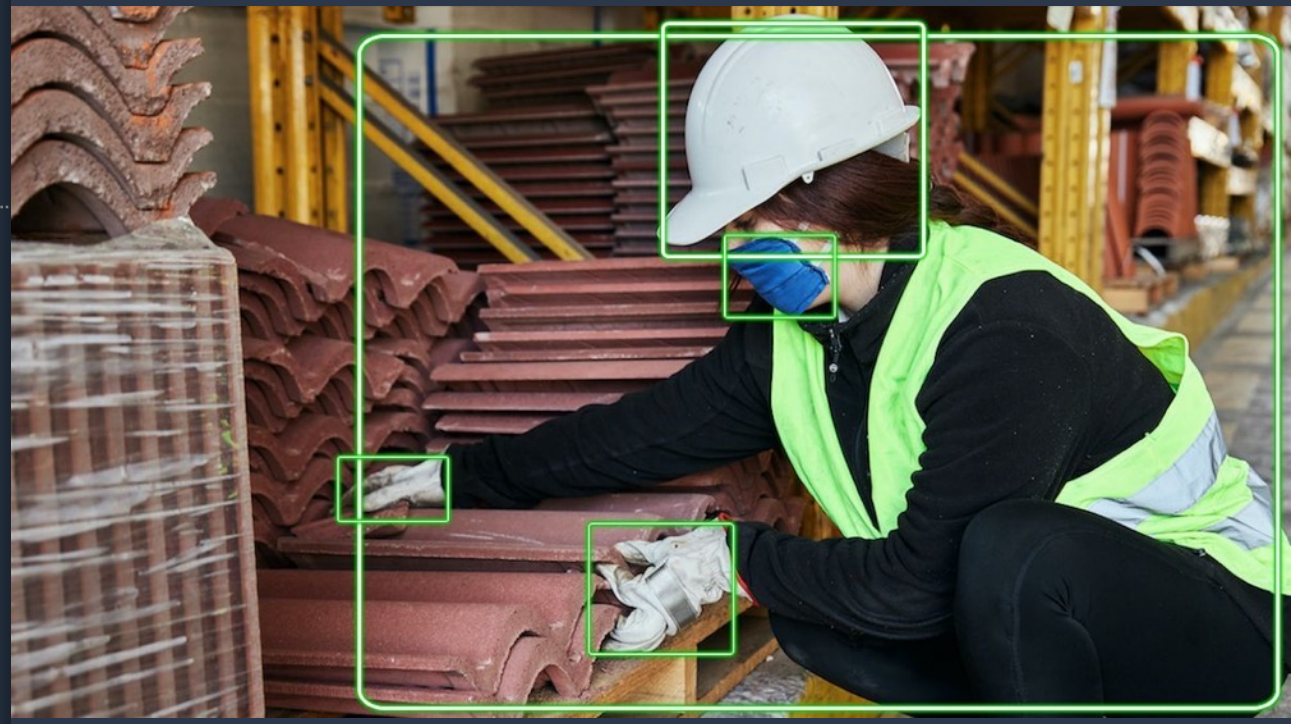
Computer vision

Predictions and insights

Content analysis and object detection

Extract insights and identify objects of interest from large volumes of images and videos with Amazon Rekognition

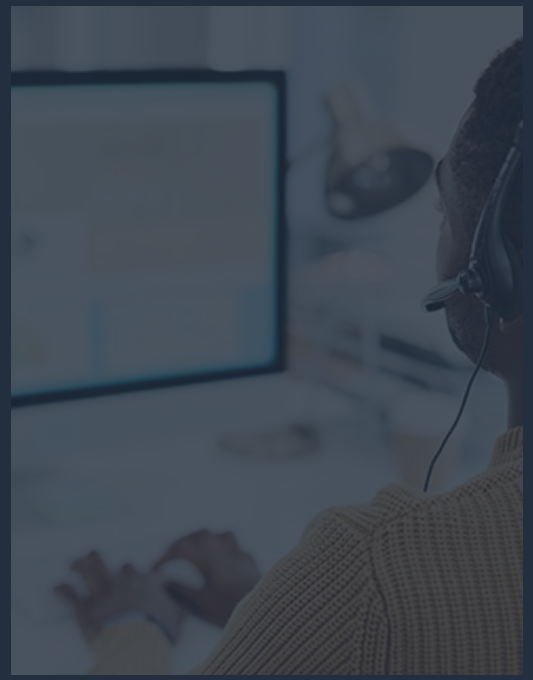
- Detect personal protective equipment (PPE) to improve worker safety
- Analyze vehicle traffic and pedestrian and bicycle safety
- Detect objects of interest in video and reduce human effort required to review footage



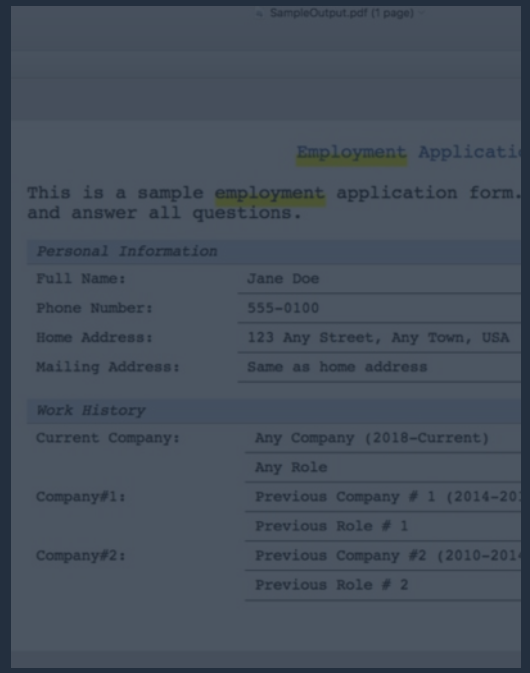
Assessing damage from natural disasters

EagleView runs deep learning models on AWS to make quicker, more accurate assessments of property damage within 24 hours of a natural disaster. Amazon Elastic Inference makes those workflows more cost effective at scale.

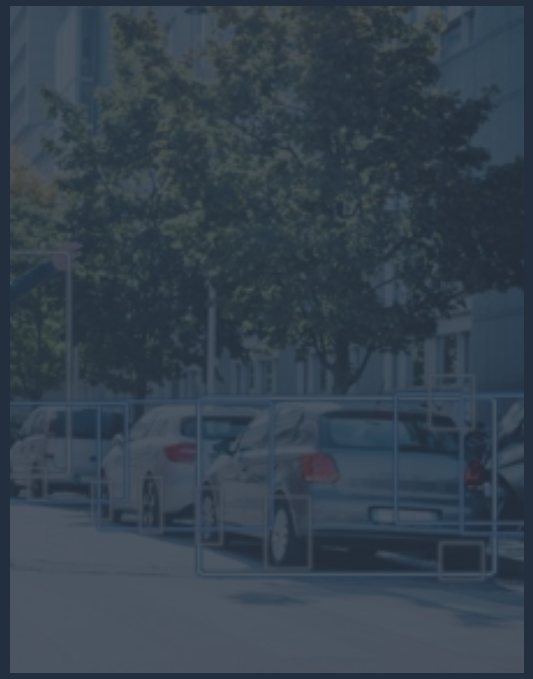
Top AI/ML use cases for state and local government



Speech and language



Intelligent document processing



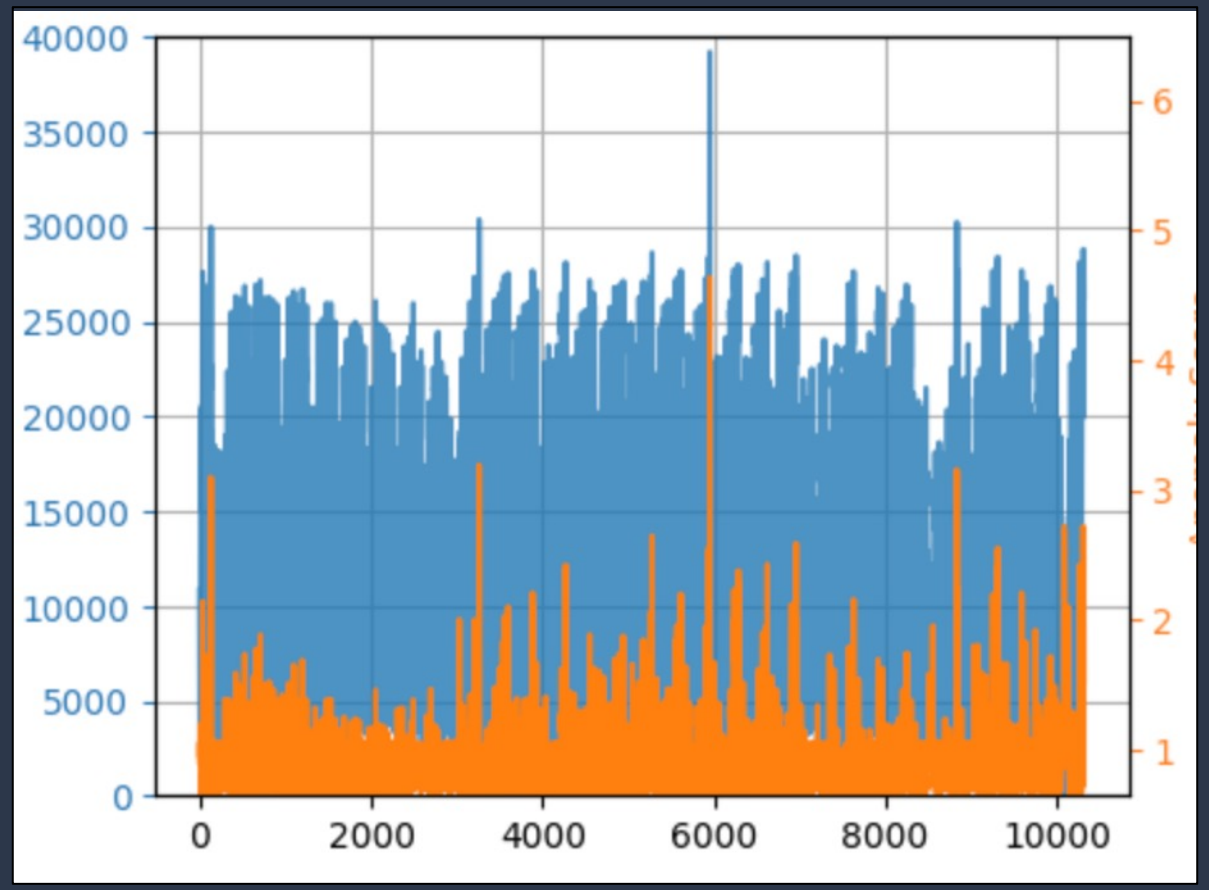
Computer vision

Predictions and insights

Fraud detection and prevention

Detect and prevent fraud, waste, and abuse

- Enhance accuracy and speed to help detect and prevent waste fraud and abuse
- Managed service approach with prebuilt ML models for fraud detection
- Supervised and unsupervised models for developing highly targeted models to utilize customer data as part of fraud prevention efforts



aws machine learning

Arizona Department of Revenue uses AI to prevent fraud

Challenge

Identifying fraudulent refund requests was a manual review process based on legacy business rules. This led to slower overall income tax return processing times and legitimate returns inaccurately marked for review.

Solution

Invested in an income tax fraud prevention solution using Machine Learning. Results from Fiscal Year 2016 to Fiscal Year 2023 were $251.3M in fraud prevented, W-2s verified, and restitution received (61 criminal convictions).

Zac Sharp
Deputy Assistant Director, Arizona Department of Revenue

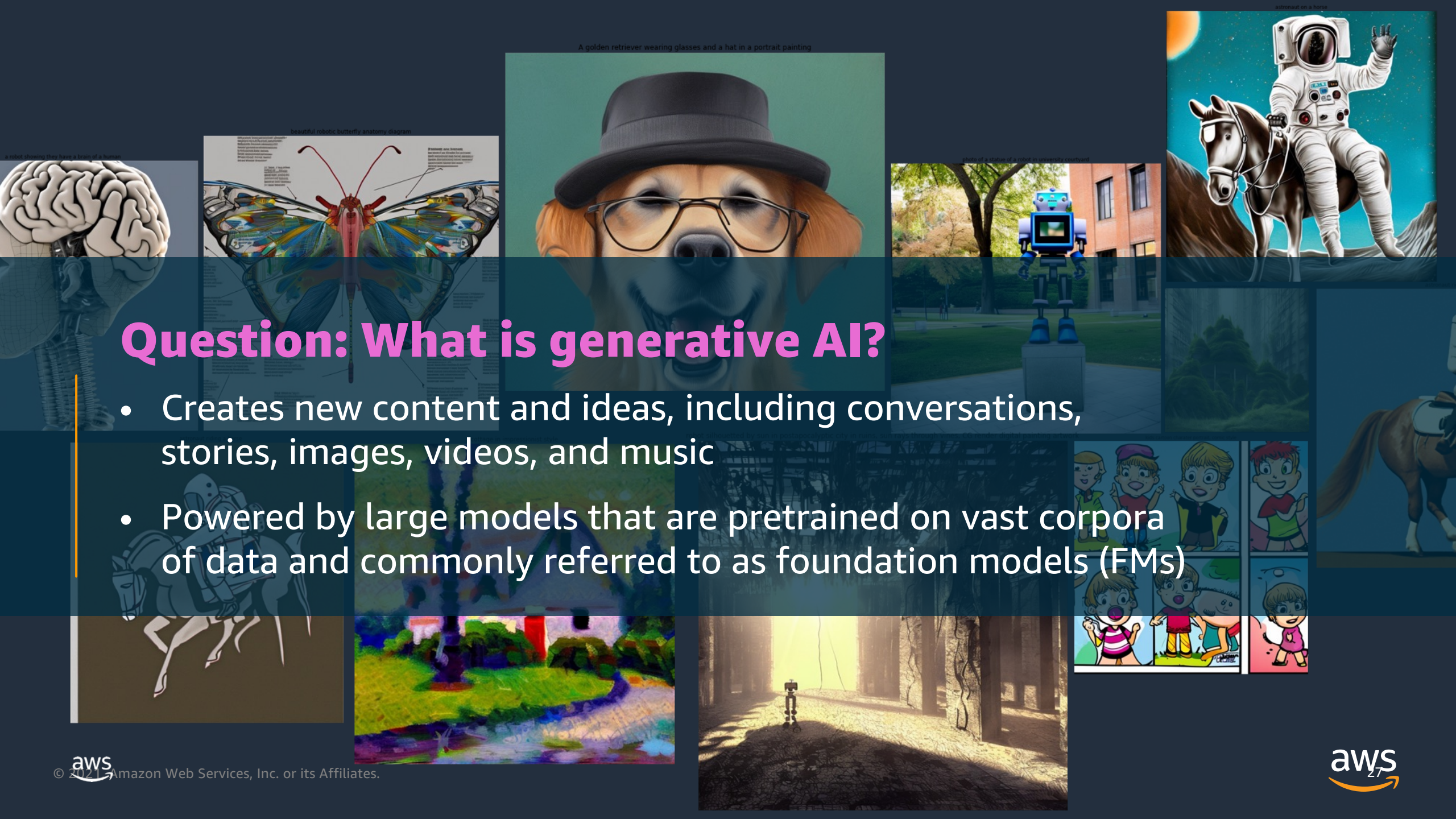
Predictions and forecasts from IoT and sensor data

Leverage data from smart cities and facilities

- **Smart cities**
- **Predictive maintenance**
- **Facility management**

Learn more about Amazon Monitron

Innovation can **transform industries**



Question: What is generative AI?

- Creates new content and ideas, including conversations, stories, images, videos, and music
- Powered by large models that are pretrained on vast corpora of data and commonly referred to as foundation models (FMs)

Enhance Customer Experiences

CHATBOTS

VIRTUAL ASSISTANTS

AGENT ASSISTANCE

CONTACT CENTER

ANALYTICS

PERSONALIZATION

Boost employee productivity & creativity

CONVERSATIONAL SEARCH

SUMMARIZATION

CONTENT CREATION

CODE GENERATION

DATA TO INSIGHTS

Optimize business processes

DOCUMENT PROCESSING

DATA AUGMENTATION

CYBERSECURITY

PROCESS OPTIMIZATION

ANOMALY DETECTION

Generative AI Stack

APPLICATIONS THAT LEVERAGE LLMs AND OTHER FMs

TOOLS TO BUILD WITH LLMs AND OTHER FMs

INFRASTRUCTURE FOR FM TRAINING AND INFERENCE

Generative AI Stack

APPLICATIONS THAT LEVERAGE LLMs AND OTHER FMs

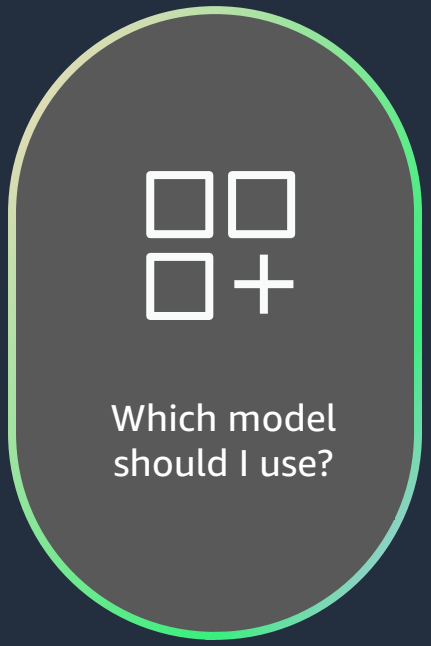
TOOLS TO BUILD WITH LLMs AND OTHER FMs

INFRASTRUCTURE FOR FM TRAINING AND INFERENCE

GPUs Trainium Inferentia SageMaker

UltraClusters EFA EC2 Capacity Blocks Nitro Neuron

- Customers have questions…



Generative AI Stack

APPLICATIONS THAT LEVERAGE LLMs AND OTHER FMs

TOOLS TO BUILD WITH LLMs AND OTHER FMs

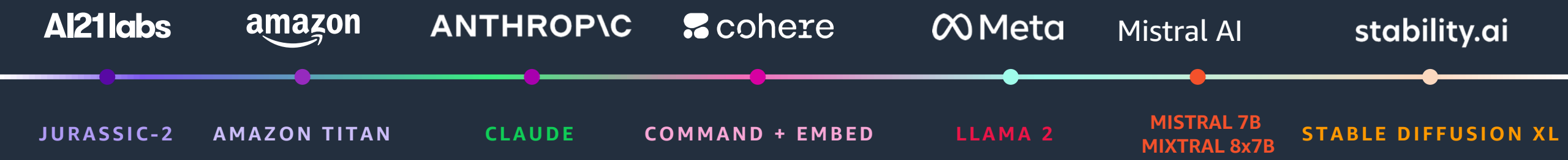
INFRASTRUCTURE FOR FM TRAINING AND INFERENCE

GPUs Trainium Inferentia SageMaker

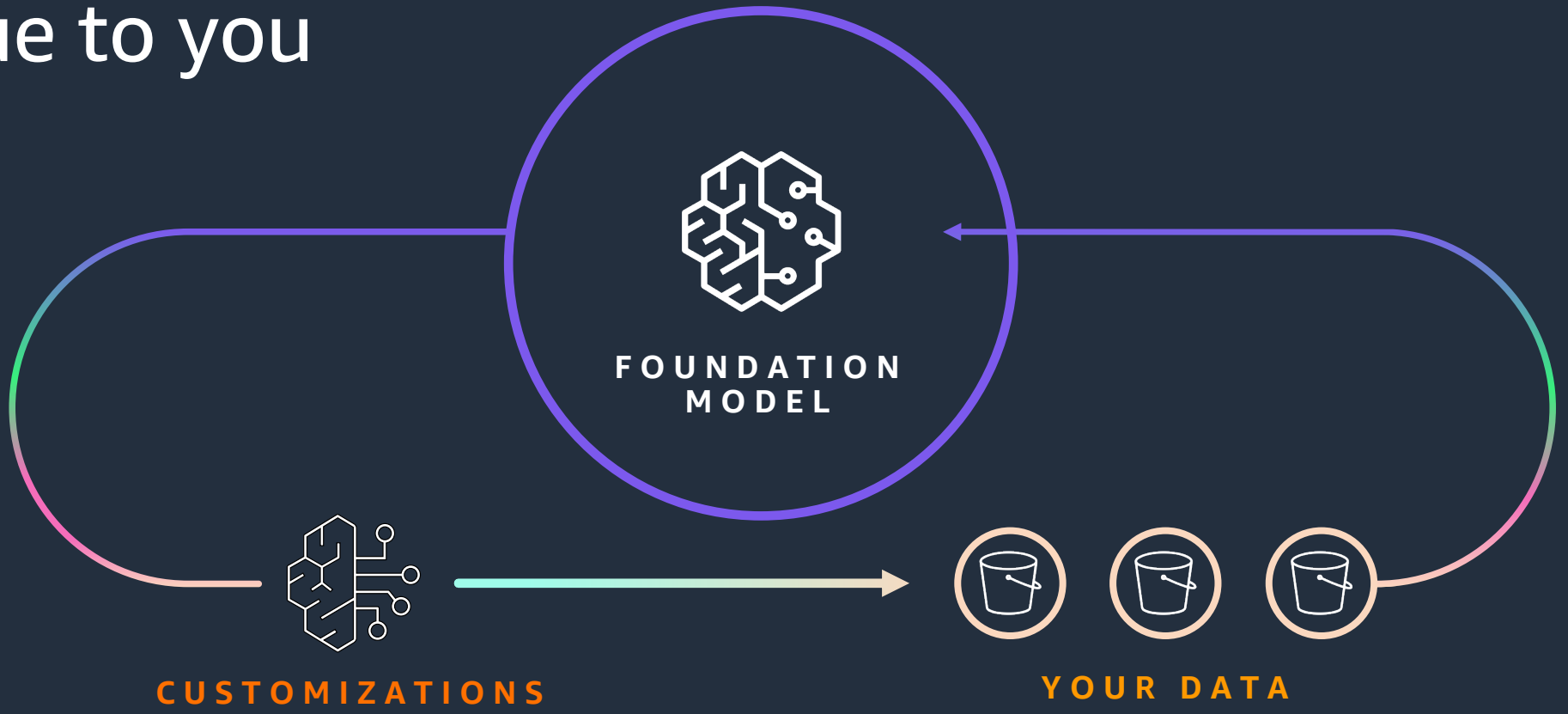
UltraClusters EFA EC2 Capacity Blocks Nitro Neuron

Amazon Bedrock

The easiest way to build and scale
generative AI applications with LLMs and other FMs

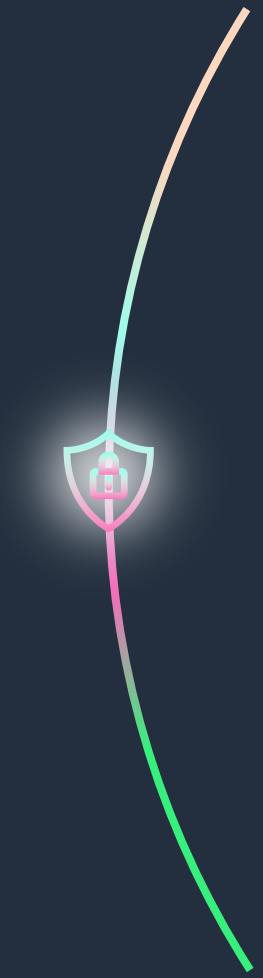


Unique to you



Fine tuning | Retrieval Augmented Generation (RAG) | Continued Pre-training

Amazon Bedrock keeps data secure & private



None of the customer's data is used to train the underlying model

All data is encrypted in transit and at rest

Data used to customize models remains within your VPC

Support for standards, (GDPR, HIPAA) and guardrails

NEW

Guardrails for Amazon Bedrock

Safeguard your generative AI applications with your responsible AI policies

Easily configure harmful content filtering based on your responsible AI policies

Apply Guardrails to any FM or agent

Redact PII information in FM responses (coming soon)

Generative AI Stack

APPLICATIONS THAT LEVERAGE LLMs AND OTHER FMs

TOOLS TO BUILD WITH LLMs AND OTHER FMs

Amazon Bedrock

Guardrails | Agents | Customization Capabilities

INFRASTRUCTURE FOR FM TRAINING AND INFERENCE

GPUs Trainium Inferentia SageMaker

UltraClusters EFA EC2 Capacity Blocks Nitro Neuron

The challenges of

AI chat applications at work

Amazon Q

AMAZON Q **BUSINESS**

AMAZON Q **DEVELOPER**

EMBEDDED

Amazon Q In Connect

Amazon Q In QuickSight

Amazon Q In AWS Supply Chain

Amazon Q

in Amazon QuickSight

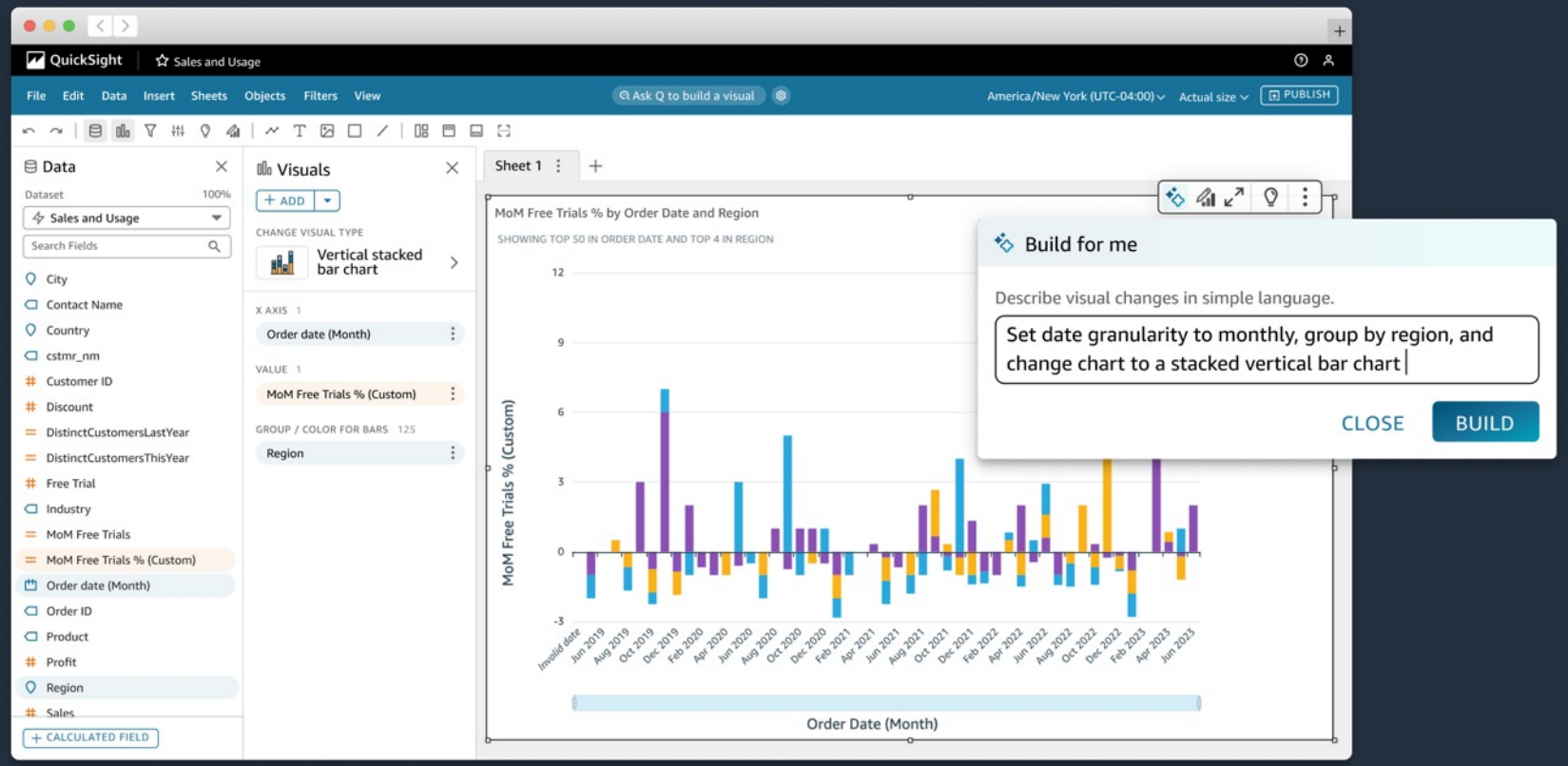
Generative dashboard authoring

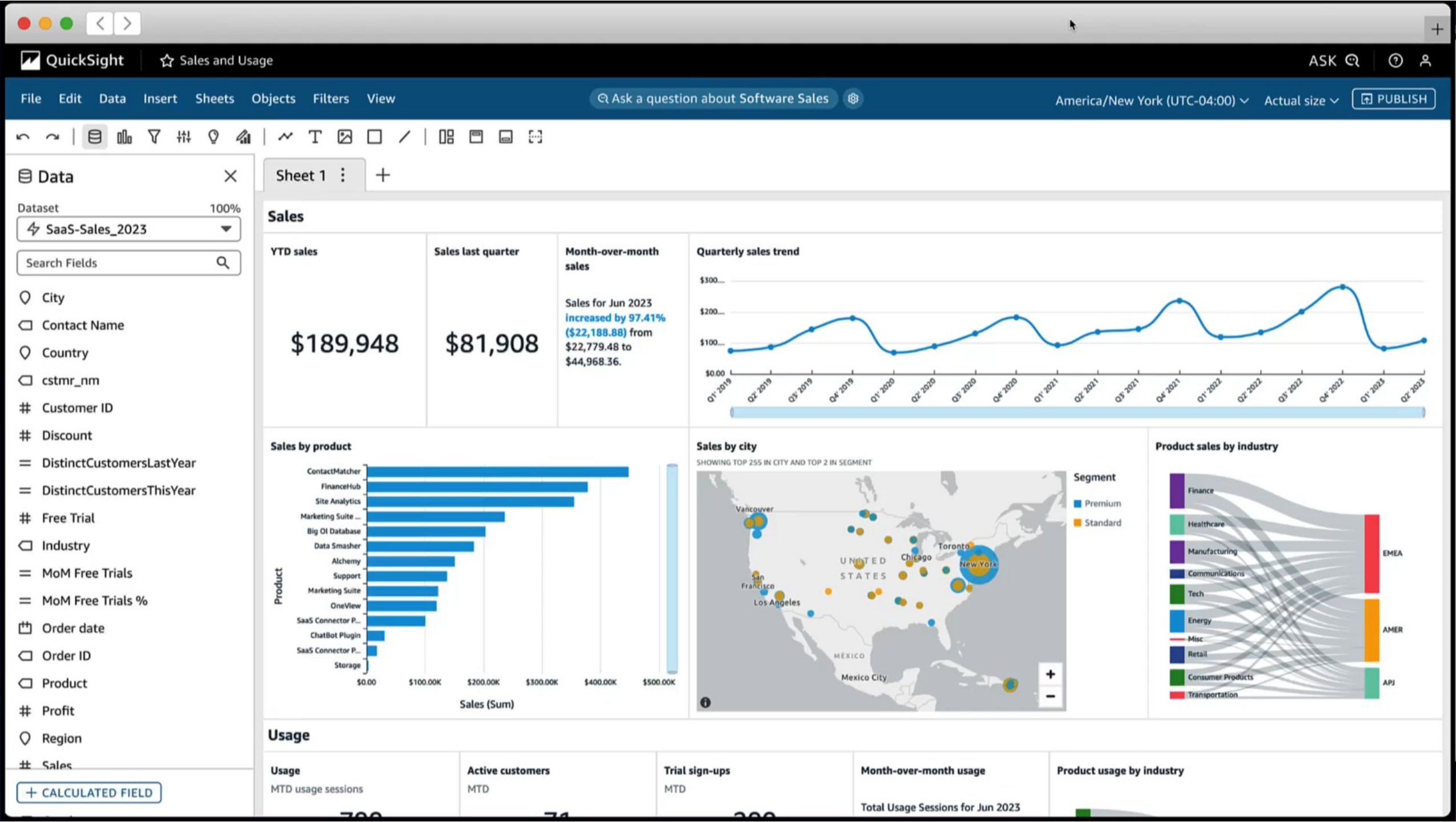
Visually compelling data stories

Reimagined Q&A experience

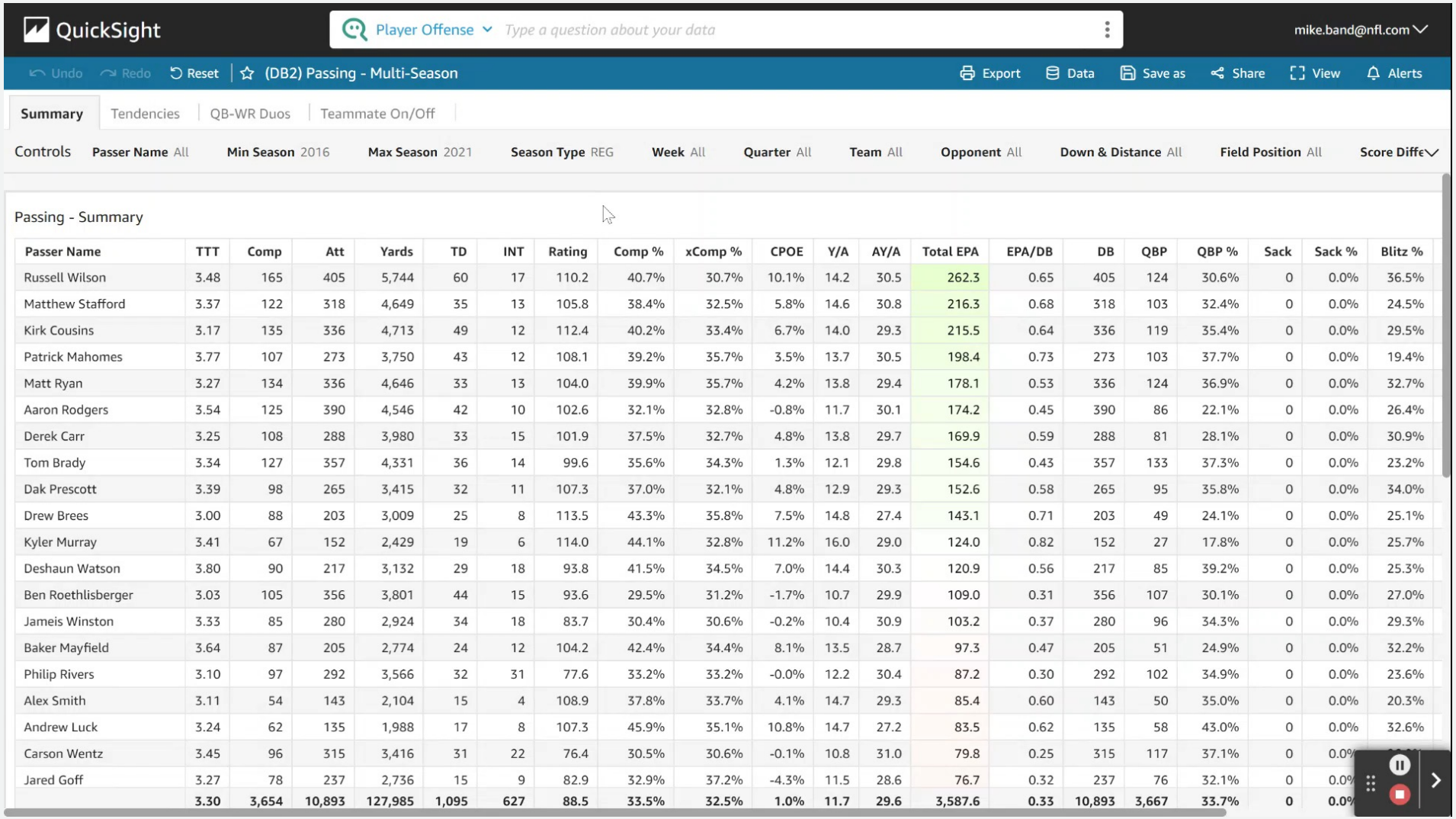
Visual authoring in **QuickSight**

Use everyday language to generate and fine-tune visuals in seconds




QuickSight
Sales and Usage
ASK
File Edit Data Insert Sheets Objects Filters View
Ask a question about Software Sales
America/New York (UTC-04:00)
Actual size
PUBLISH
Data
Dataset 100%
SaaS-Sales_2023
Search Fields
City
Contact Name
Country
cstmr_nm
Customer ID
Discount
DistinctCustomersLastYear
DistinctCustomersThisYear
Free Trial
Industry
MoM Free Trials
MoM Free Trials %
Order date
Order ID
Product
Profit
Region
Sales
+ CALCULATED FIELD
Sheet 1
Sales
YTD sales
$189,948
Sales last quarter
$81,908
Month-over-month sales
Sales for Jun 2023 increased by 97.41% ($22,188.88) from $22,779.48 to $44,968.36.
Quarterly sales trend
$300...
$200...
$100...
$0.00
Q1' 2019
Q2' 2019
Q3' 2019
Q4' 2019
Q1' 2020
Q2' 2020
Q3' 2020
Q4' 2020
Q1' 2021
Q2' 2021
Q3' 2021
Q4' 2021
Q1' 2022
Q2' 2022
Q3' 2022
Q4' 2022
Q1' 2023
Q2' 2023
Sales by product
ContactMatcher
FinanceHub
Site Analytics
Marketing Suite ...
Big OI Database
Data Smasher
Alchemy
Support
Marketing Suite
OneView
SaaS Connector P...
ChatBot Plugin
SaaS Connector P...
Storage
Product
$0.00
$100.00K
$200.00K
$300.00K
$400.00K
$500.00K
Sales (Sum)
Sales by city
SHOWING TOP 255 IN CITY AND TOP 2 IN SEGMENT
Segment
Premium
Standard
Vancouver
Toronto
Chicago
New York
UNITED STATES
San Francisco
Los Angeles
MÉXICO
Mexico City
Product sales by industry
Finance
Healthcare
Manufacturing
Communications
Tech
Energy
Misc
Retail
Consumer Products
Transportation
EMEA
AMER
APJ
Usage
Usage
MTD usage sessions
Active customers
MTD
Trial sign-ups
MTD
Month-over-month usage
Total Usage Sessions for Jun 2023
Product usage by industry

How the NFL uses Generative AI to get instant answers



Passing - Summary

Passer Name	TTT	Comp	Att	Yards	TD	INT	Rating	Comp %	xComp %	CPOE	Y/A	AY/A	Total EPA	EPA/DB	DB	QBP	QBP %	Sack	Sack %	Blitz %
Russell Wilson	3.48	165	405	5,744	60	17	110.2	40.7%	30.7%	10.1%	14.2	30.5	262.3	0.65	405	124	30.6%	0	0.0%	36.5%
Matthew Stafford	3.37	122	318	4,649	35	13	105.8	38.4%	32.5%	5.8%	14.6	30.8	216.3	0.68	318	103	32.4%	0	0.0%	24.5%
Kirk Cousins	3.17	135	336	4,713	49	12	112.4	40.2%	33.4%	6.7%	14.0	29.3	215.5	0.64	336	119	35.4%	0	0.0%	29.5%
Patrick Mahomes	3.77	107	273	3,750	43	12	108.1	39.2%	35.7%	3.5%	13.7	30.5	198.4	0.73	273	103	37.7%	0	0.0%	19.4%
Matt Ryan	3.27	134	336	4,646	33	13	104.0	39.9%	35.7%	4.2%	13.8	29.4	178.1	0.53	336	124	36.9%	0	0.0%	32.7%
Aaron Rodgers	3.54	125	390	4,546	42	10	102.6	32.1%	32.8%	-0.8%	11.7	30.1	174.2	0.45	390	86	22.1%	0	0.0%	26.4%
Derek Carr	3.25	108	288	3,980	33	15	101.9	37.5%	32.7%	4.8%	13.8	29.7	169.9	0.59	288	81	28.1%	0	0.0%	30.9%
Tom Brady	3.34	127	357	4,331	36	14	99.6	35.6%	34.3%	1.3%	12.1	29.8	154.6	0.43	357	133	37.3%	0	0.0%	23.2%
Dak Prescott	3.39	98	265	3,415	32	11	107.3	37.0%	32.1%	4.8%	12.9	29.3	152.6	0.58	265	95	35.8%	0	0.0%	34.0%
Drew Brees	3.00	88	203	3,009	25	8	113.5	43.3%	35.8%	7.5%	14.8	27.4	143.1	0.71	203	49	24.1%	0	0.0%	25.1%
Kyler Murray	3.41	67	152	2,429	19	6	114.0	44.1%	32.8%	11.2%	16.0	29.0	124.0	0.82	152	27	17.8%	0	0.0%	25.7%
Deshaun Watson	3.80	90	217	3,132	29	18	93.8	41.5%	34.5%	7.0%	14.4	30.3	120.9	0.56	217	85	39.2%	0	0.0%	25.3%
Ben Roethlisberger	3.03	105	356	3,801	44	15	93.6	29.5%	31.2%	-1.7%	10.7	29.9	109.0	0.31	356	107	30.1%	0	0.0%	27.0%
Jameis Winston	3.33	85	280	2,924	34	18	83.7	30.4%	30.6%	-0.2%	10.4	30.9	103.2	0.37	280	96	34.3%	0	0.0%	29.3%
Baker Mayfield	3.64	87	205	2,774	24	12	104.2	42.4%	34.4%	8.1%	13.5	28.7	97.3	0.47	205	51	24.9%	0	0.0%	32.2%
Philip Rivers	3.10	97	292	3,566	32	31	77.6	33.2%	33.2%	-0.0%	12.2	30.4	87.2	0.30	292	102	34.9%	0	0.0%	23.6%
Alex Smith	3.11	54	143	2,104	15	4	108.9	37.8%	33.7%	4.1%	14.7	29.3	85.4	0.60	143	50	35.0%	0	0.0%	20.3%
Andrew Luck	3.24	62	135	1,988	17	8	107.3	45.9%	35.1%	10.8%	14.7	27.2	83.5	0.62	135	58	43.0%	0	0.0%	32.6%
Carson Wentz	3.45	96	315	3,416	31	22	76.4	30.5%	30.6%	-0.1%	10.8	31.0	79.8	0.25	315	117	37.1%	0	0.0%	
Jared Goff	3.27	78	237	2,736	15	9	82.9	32.9%	37.2%	-4.3%	11.5	28.6	76.7	0.32	237	76	32.1%	0	0.0%	
	3.30	**3,654**	**10,893**	**127,985**	**1,095**	**627**	**88.5**	**33.5%**	**32.5%**	**1.0%**	**11.7**	**29.6**	**3,587.6**	**0.33**	**10,893**	**3,667**	**33.7%**	**0**	**0.0%**	

(Play video)

Generative AI stack

APPLICATIONS THAT LEVERAGE LLMs AND OTHER FMs

Amazon Q | Amazon Q in Amazon QuickSight | Amazon Q in Amazon Connect | Amazon CodeWhisperer

TOOLS TO BUILD WITH LLMs AND OTHER FMs

Amazon Bedrock

Guardrails | Agents | Customization Capabilities

INFRASTRUCTURE FOR FM TRAINING AND INFERENCE

GPUs | Trainium | Inferentia | SageMaker

UltraClusters | EFA | EC2 Capacity Blocks | Nitro | Neuron

Bedrock Studio (In Preview)

BUILD GENAI APPLICATIONS FASTER AND MORE SECURELY

Easy to use playground

Projects based collaboration

Easy access with corporate SSO

Purpose of Gen/AI Sandbox

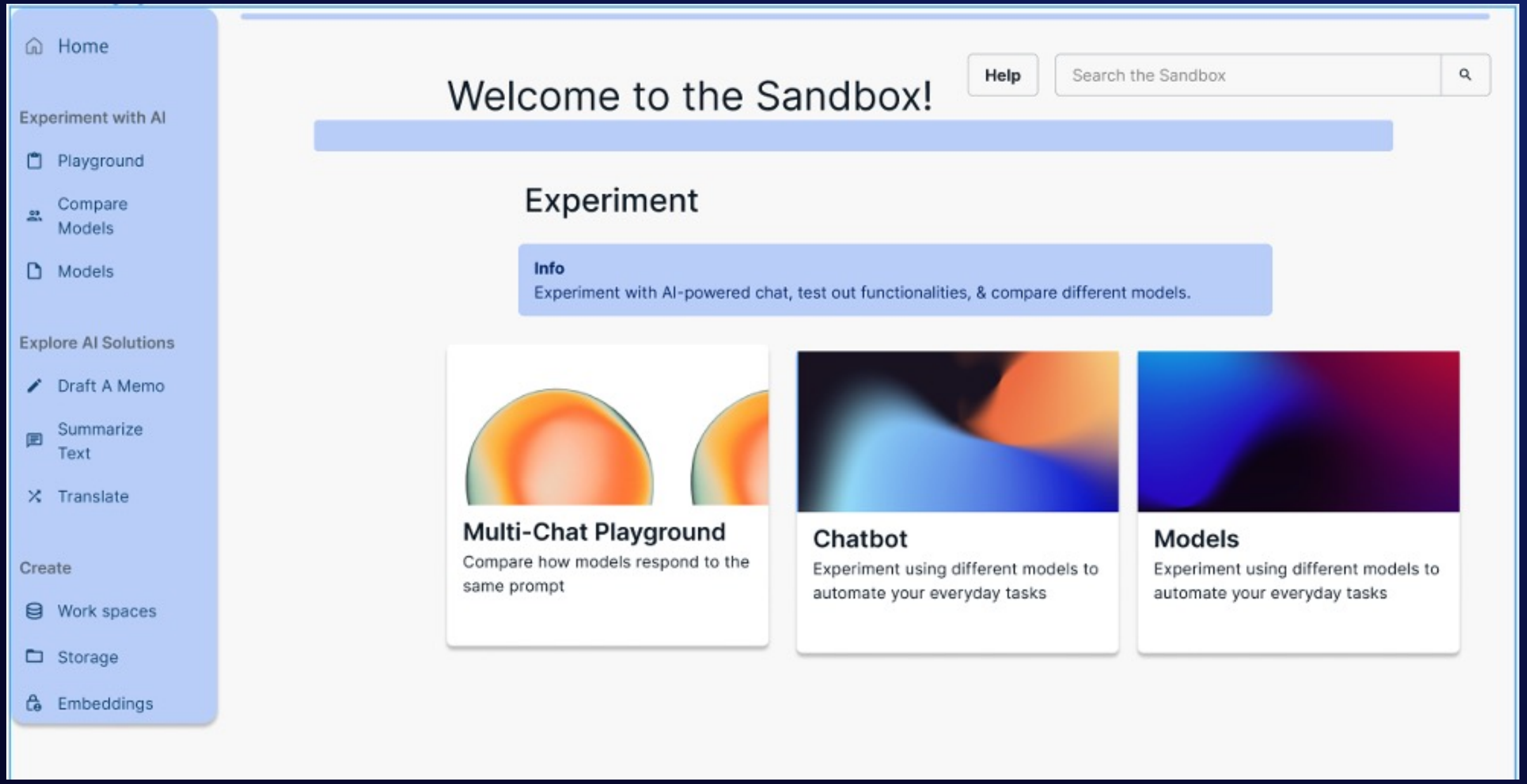
Security - An environment to allow teams to experiment the Generative AI securely and safely

Scale - AI Center Of Excellence team (build the AI environment, data integration and training/support)

Empower - Non-Technical and Technical State staffs for Innovation

Accelerate - high impact use cases opportunity to production.

Preview : If scaling, Gen/AI Sandbox Application – Landing Page



Generative AI Sandbox Requirements

Technical

- ❑ Use AWS Multi-Accounts Cloud Foundation with guardrails and budgets
- ❑ Isolate environment with network boundary
- ❑ Use your AD for user access management
- ❑ Leverage “Canned” Gen/AI applications to AI Sandbox where is possible
- ❑ Apply guardrails to large language models
- ❑ Usage reports
- ❑ Multi-Languages support

Business

- ❑ Self-service
- ❑ Single Sign-On(SSO)
- ❑ Gen/AI playground
- ❑ User Data Store
- ❑ Historical Prompts

Call to action:

- Close the loop between business and technology
- Select your use case and champion
- Connect with us at the Ask the Experts Table
- Learn more at https://aws.amazon.com/ai

Thank you!

Hiren Deliwala
hdaws@amazon.com

Please complete the session survey!

Data and analytics track

AI/ML for Data and Analytics