



aws INNOVATE

AI/ML EDITION

24 February 2022

Innovate faster and build a smarter business with AI & ML on AWS

Dean Samuels

Chief Technologist, ASEAN
Amazon Web Services

Olivier Klein

Chief Technologist, APJ
Amazon Web Services





**Enhance
customer
experience**



**Better and faster
decision-making**

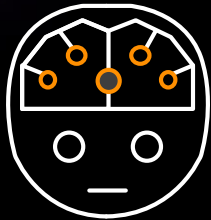


**Improve business
operations**



**New products
& services**

What is it?



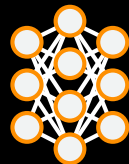
Artificial intelligence (AI)

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



Machine learning (ML)

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



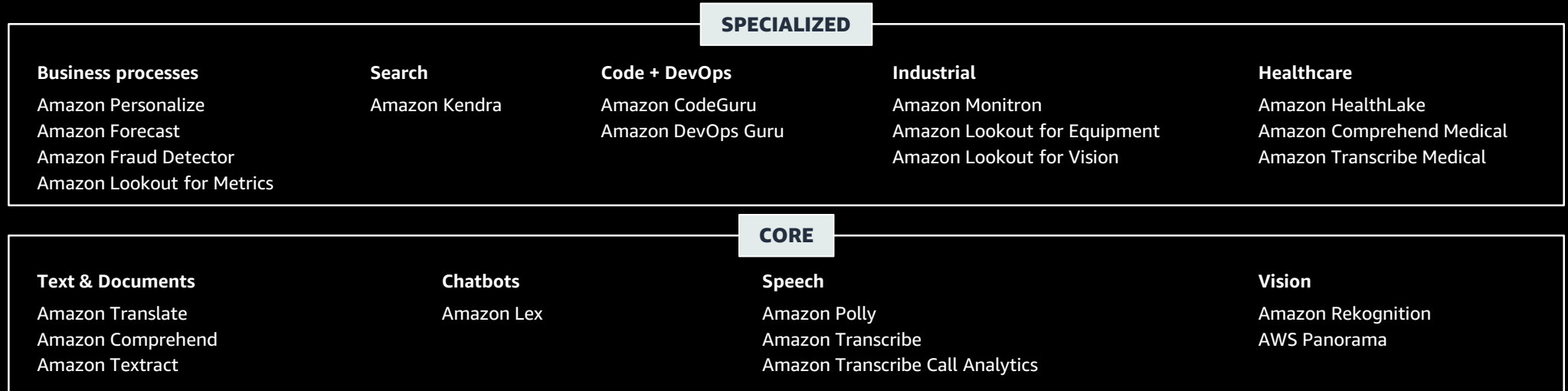
Deep learning (DL)

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

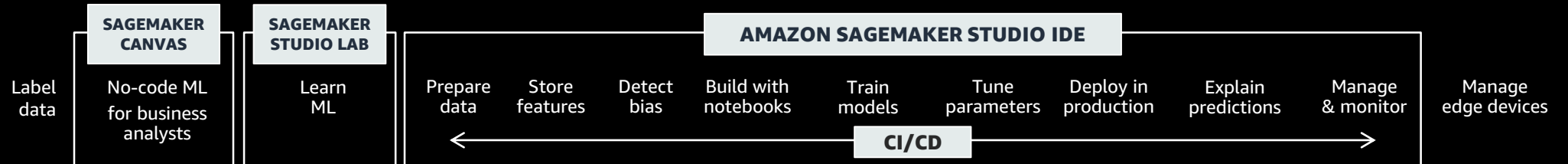
The AWS ML stack

Broadest and most complete set of machine learning capabilities

AI SERVICES



ML SERVICES



ML FRAMEWORKS & INFRASTRUCTURE

PyTorch, Apache MXNet, TensorFlow

Amazon EC2

CPUs

GPUs

AWS Inferentia

AWS Trainium

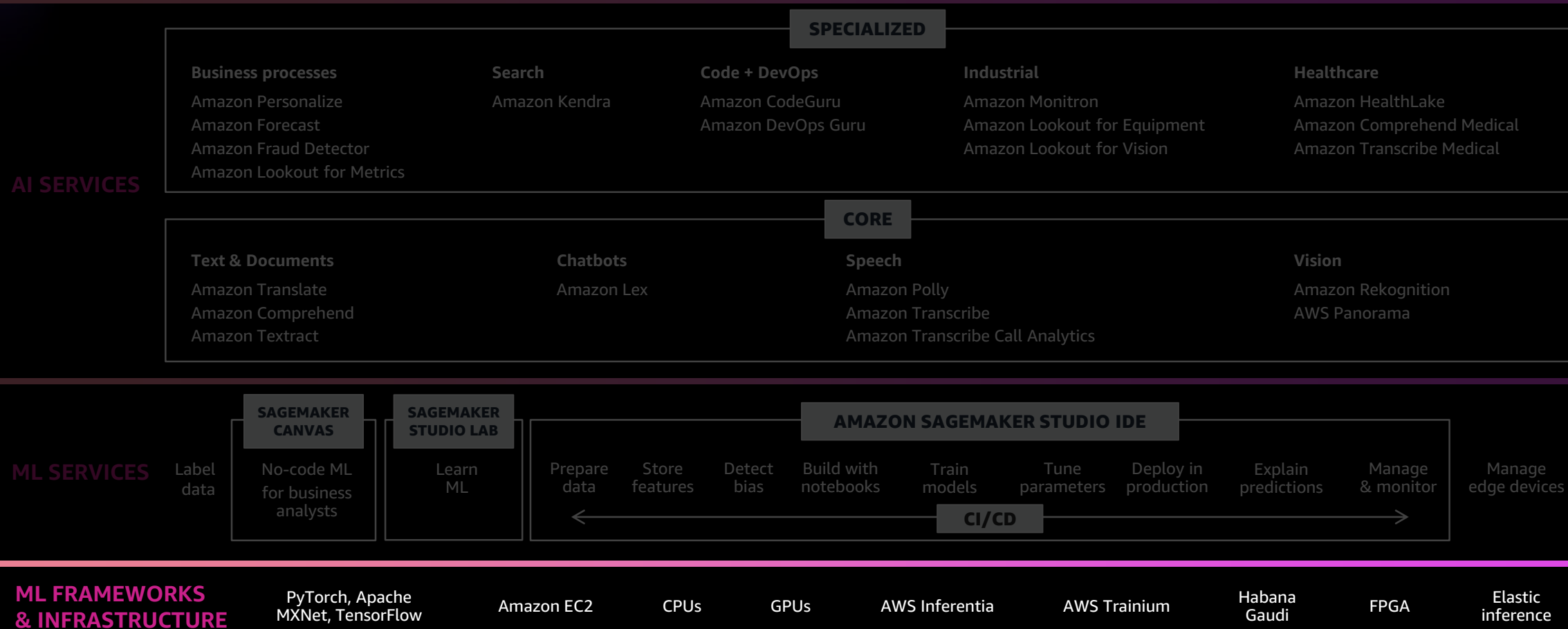
Habana Gaudi

FPGA

Elastic inference

The AWS ML stack

Broadest and most complete set of machine learning capabilities



Broadest and deepest compute infrastructure for AI/ML

Choice of CPUs, GPUs, and accelerators for your performance and budget needs

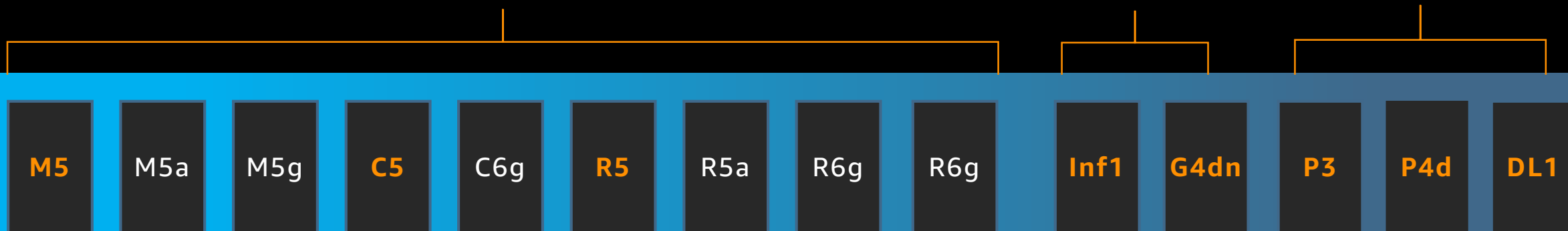
Traditional machine learning (ML)

Training and inference

Deep learning (DL)

Inference

Training



Cascade Lake CPU
Skylake CPU
Habana Gaudi accelerators



EPYC CPU



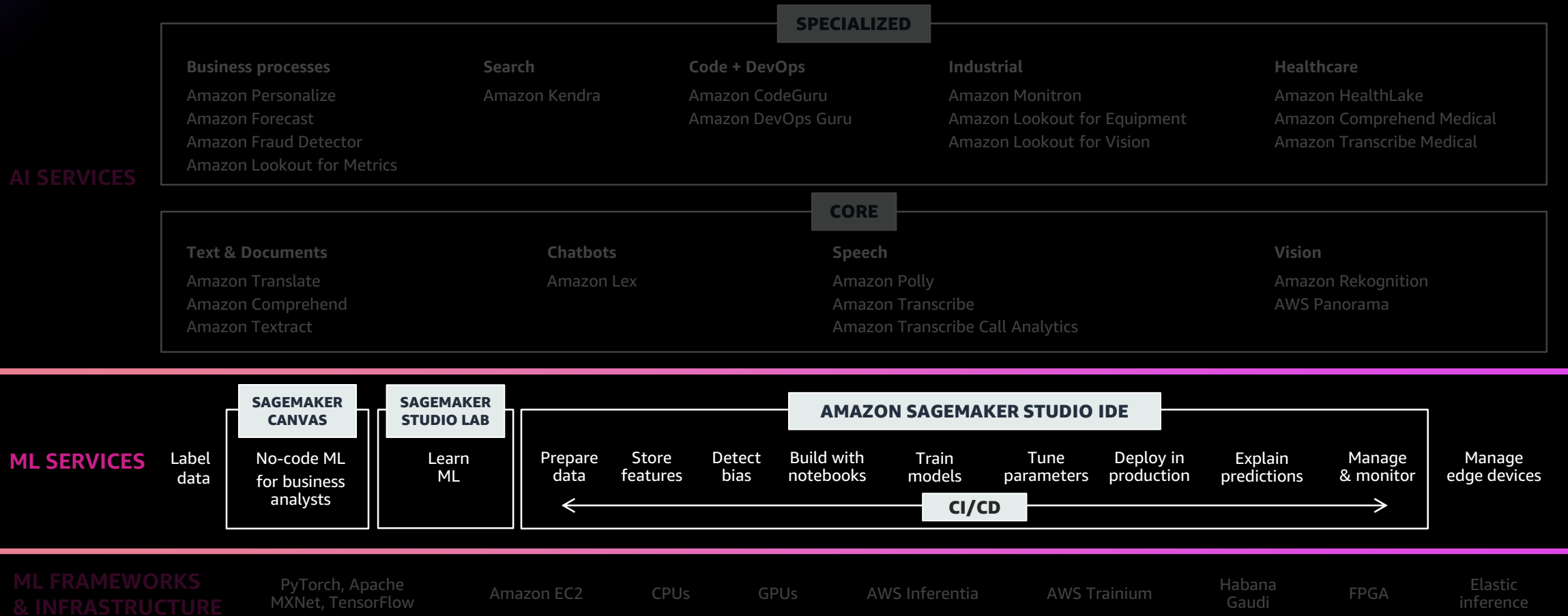
AWS Graviton
AWS Inferentia
AWS Trainium



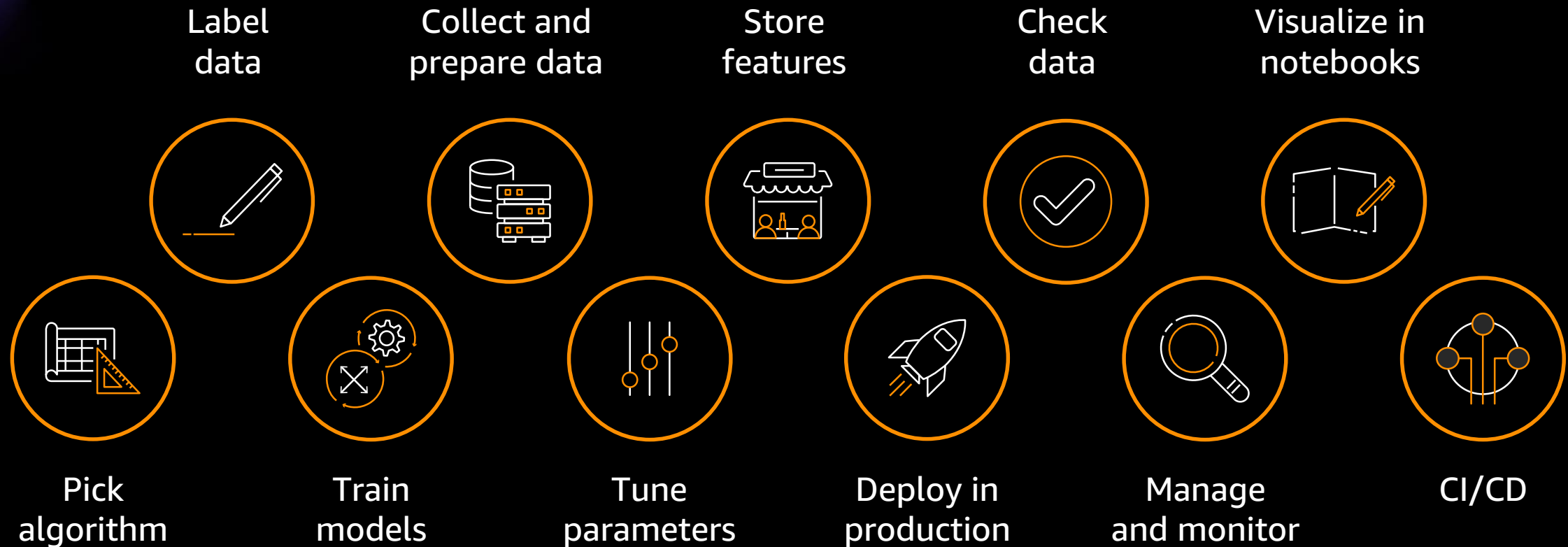
A100, V100, T4 GPUs

The AWS ML stack

Broadest and most complete set of machine learning capabilities



Amazon SageMaker: Built to make ML more accessible

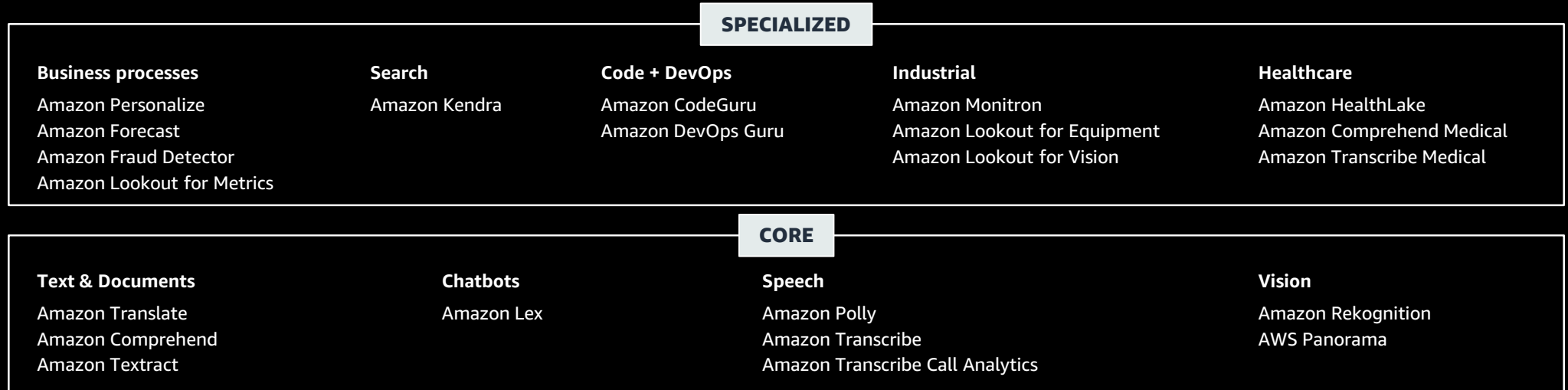


SageMaker Studio IDE

The AWS ML stack

Broadest and most complete set of machine learning capabilities

AI SERVICES



ML SERVICES



ML FRAMEWORKS & INFRASTRUCTURE

PyTorch, Apache MXNet, TensorFlow

Amazon EC2

CPUs

GPUs

AWS Inferentia

AWS Trainium

Habana Gaudi

FPGA

Elastic inference

AI Services: Easily add intelligence to applications

No machine learning skills required



Vision



Chatbots



Business tools



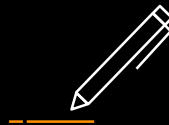
Search



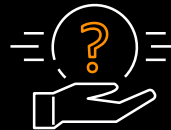
Healthcare



Speech



Text



Contact centers

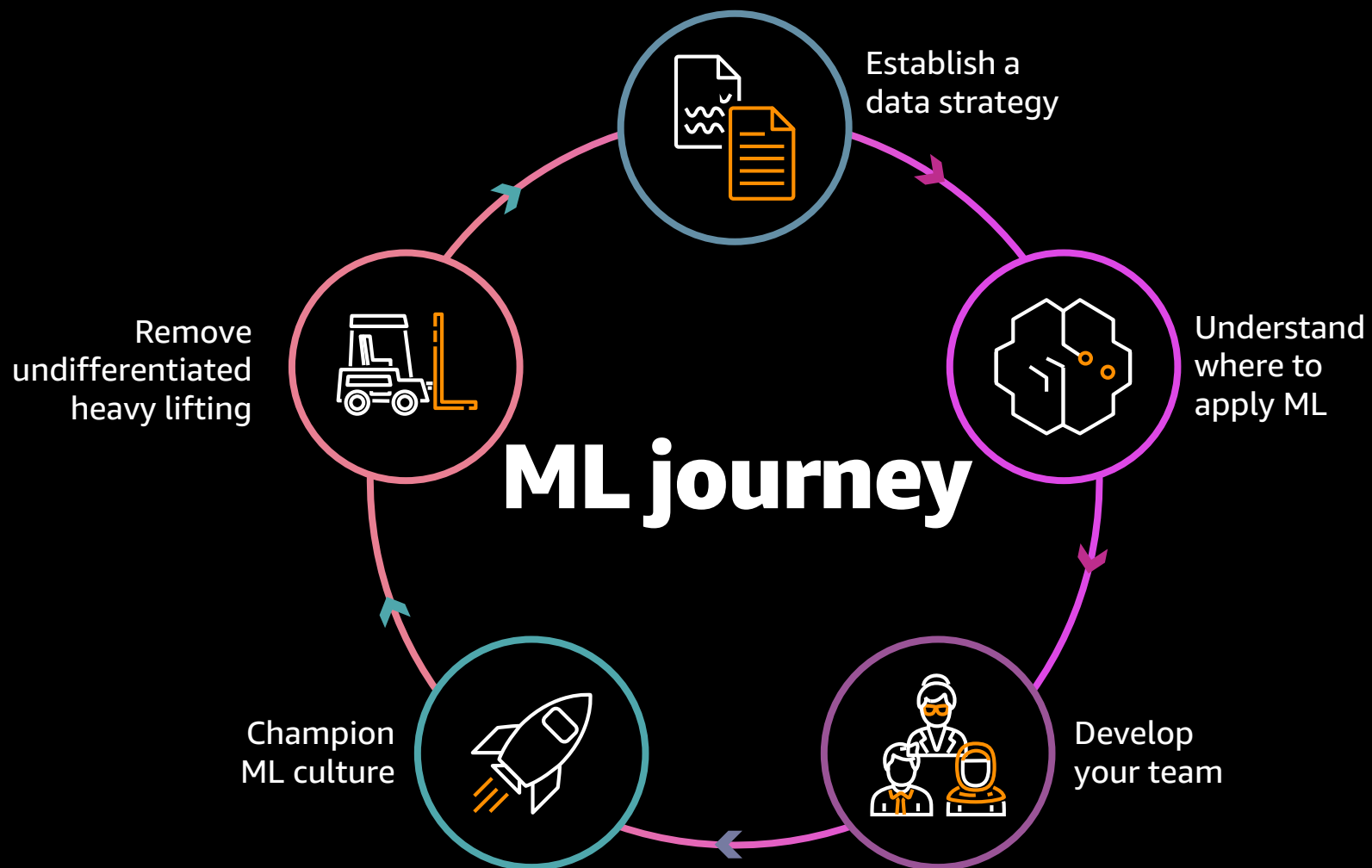


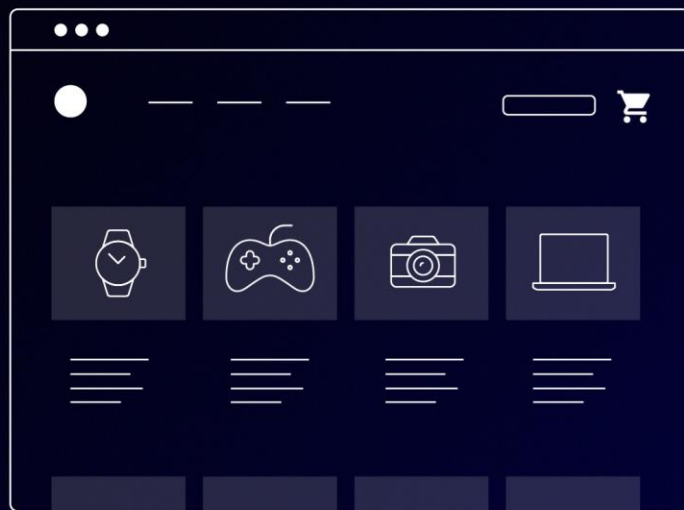
**Code and
DevOps**

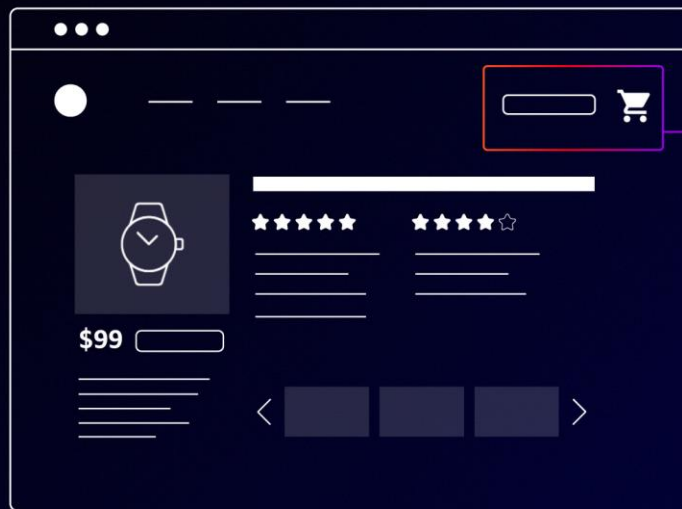


Industrial

Evolution to AI/ML

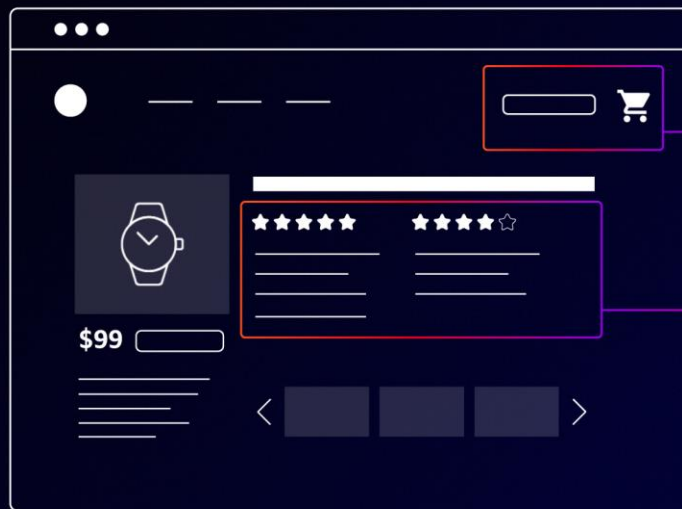






RELATIONAL



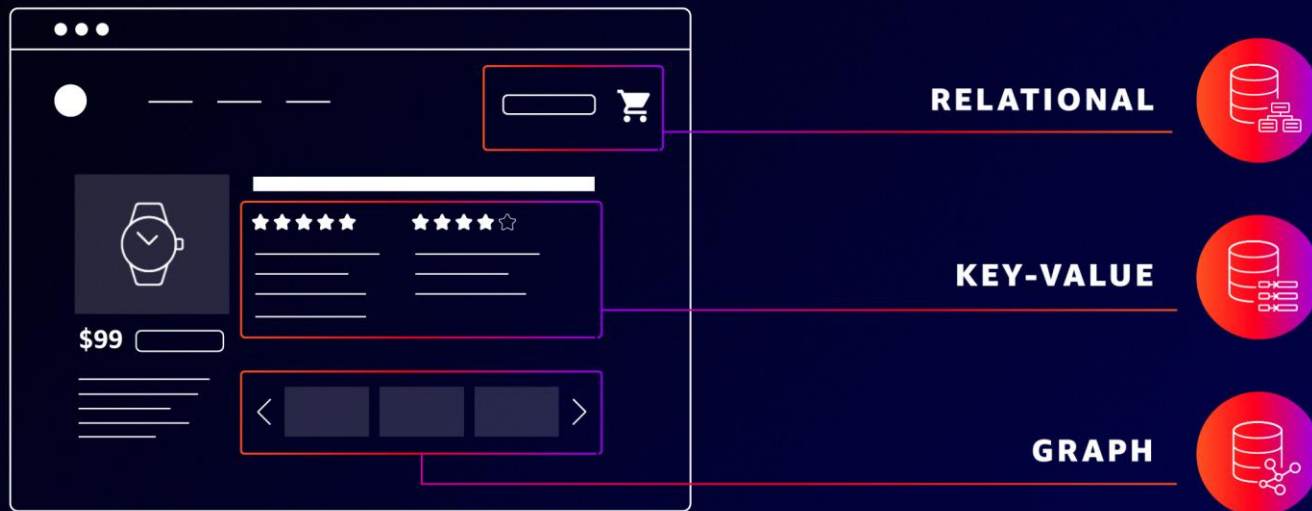


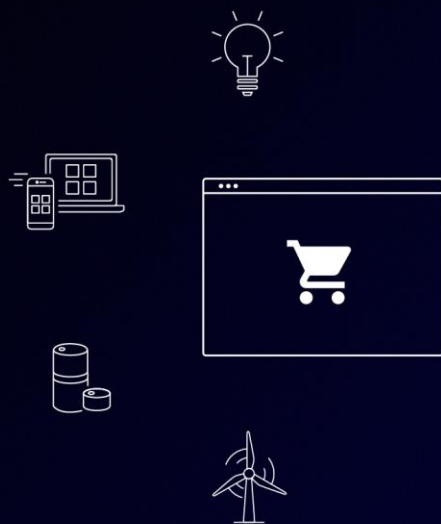
RELATIONAL



KEY-VALUE







DOCUMENT



RELATIONAL



KEY-VALUE



GRAPH



LEDGER





**"How are my customer
satisfaction scores trending
quarter over quarter?"**



**"How should I forecast
the sales** of a brand new
product to my business?"





Store + Access

Analyze + Visualize



BUSINESS INTELLIGENCE DASHBOARD



DATA STREAMING



SEARCH + ANALYZE



DATA WAREHOUSE



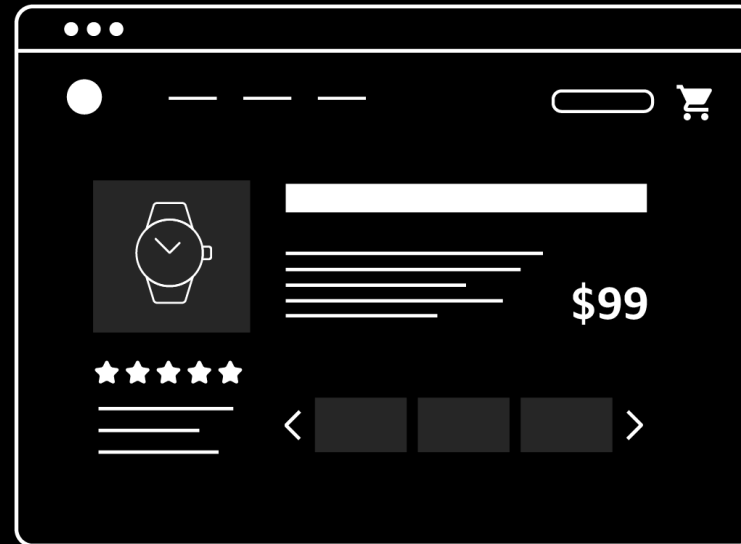
Store + Access



Analyze + Visualize



Predict





Store + Access



Analyze + Visualize



Predict



Security + Access Control

Most comprehensive set of services for **the entire end-to-end data, analytics, and ML journey** for all workloads and all types of data



Store + Access

Relational and purpose-built
databases, IoT, data lake



Analyze + Visualize

Purpose-built analytics tools and
business intelligence



Predict

Machine learning and
AI services



Security + Access Control

More than 1.5 million AWS customers
use database, analytics, or machine learning AWS services

More than 1.5 million AWS customers

use database, analytics, or machine learning AWS services



Driving **business** **outcomes** with AI/ML

Business impact of machine learning



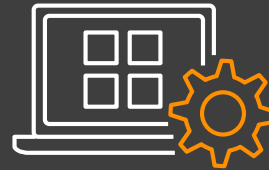
OPTIMIZING BUSINESS WITH NEW EFFICIENCIES

Create greater efficiency through sophisticated demand planning and forecasting models



MAKING SMARTER, FASTER DECISIONS

Make more informed, faster decisions to act on opportunities sooner and get better results



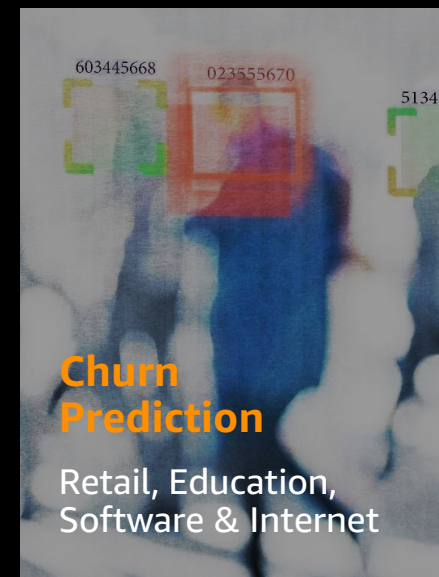
ADDING NEW CAPABILITIES TO EXISTING PRODUCTS

Enrich existing product to improve customer engagement and attract new users through deeper experiences



INVENTING NET-NEW PRODUCTS

Use data to develop innovative ideas and bring new products to market





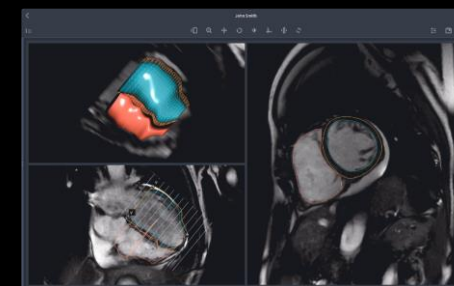
Moderna Therapeutics is using AWS **Machine Learning to automate their lab and manufacturing processes**, making it even faster for them to develop mRNA vaccines and bring them to market



Using AWS ML, GE Healthcare developed a model that can **learn from thousands of medical scans to detect anomalies** more accurately and efficiently, allowing radiologists to prioritize patients needing immediate attention.



X-Ray analysis



MRI scans analysis

Wanted to **increase customer loyalty by improving customer engagement** and driving new product purchase rates across 600,000 daily shoppers across 189 stores

Used **Amazon Personalize** to analyze sales history, product metadata, and user profiles to improve coupon recommendation service.

5x increase in response to recommended products compared to legacy solution leading and Increased **new product purchase to 40%**.



Drive down pickup and delivery times for customers, for example a store in Australia **averaged delivery times of under 5 minutes**, from order to doorstep, across an entire week

Created a data lake consisting of key order information and **predict the likelihood that an order will be placed**, so a store can begin making that order right before it is placed

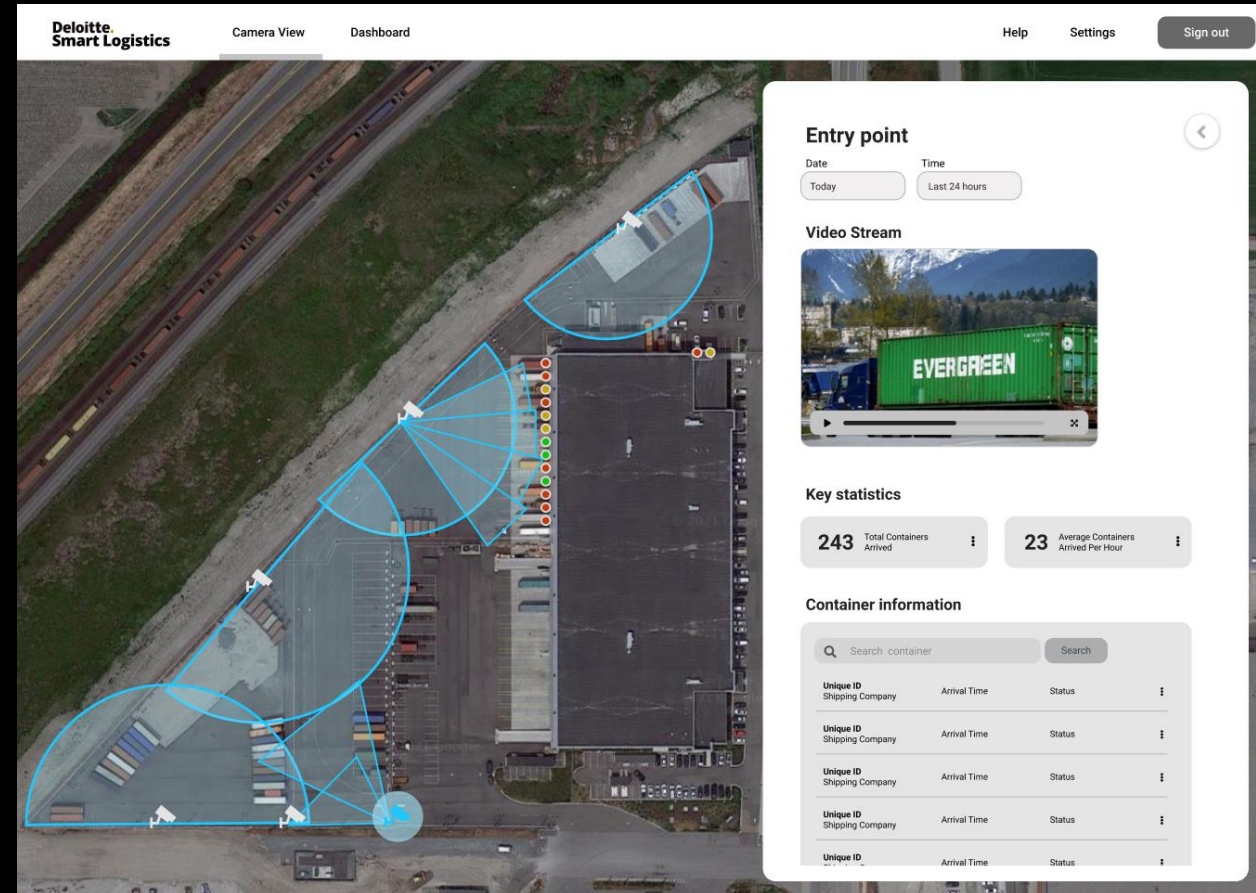
Store employees view an ordering screen **displaying specific pizzas with various color indicators** corresponding to the likelihood of those pizzas being ordered

Container tracking at the Port of Vancouver

Deloitte is working with others leveraging AWS Panorama to deploy cameras that allow for comprehensive tracking of container movement across the entire site.

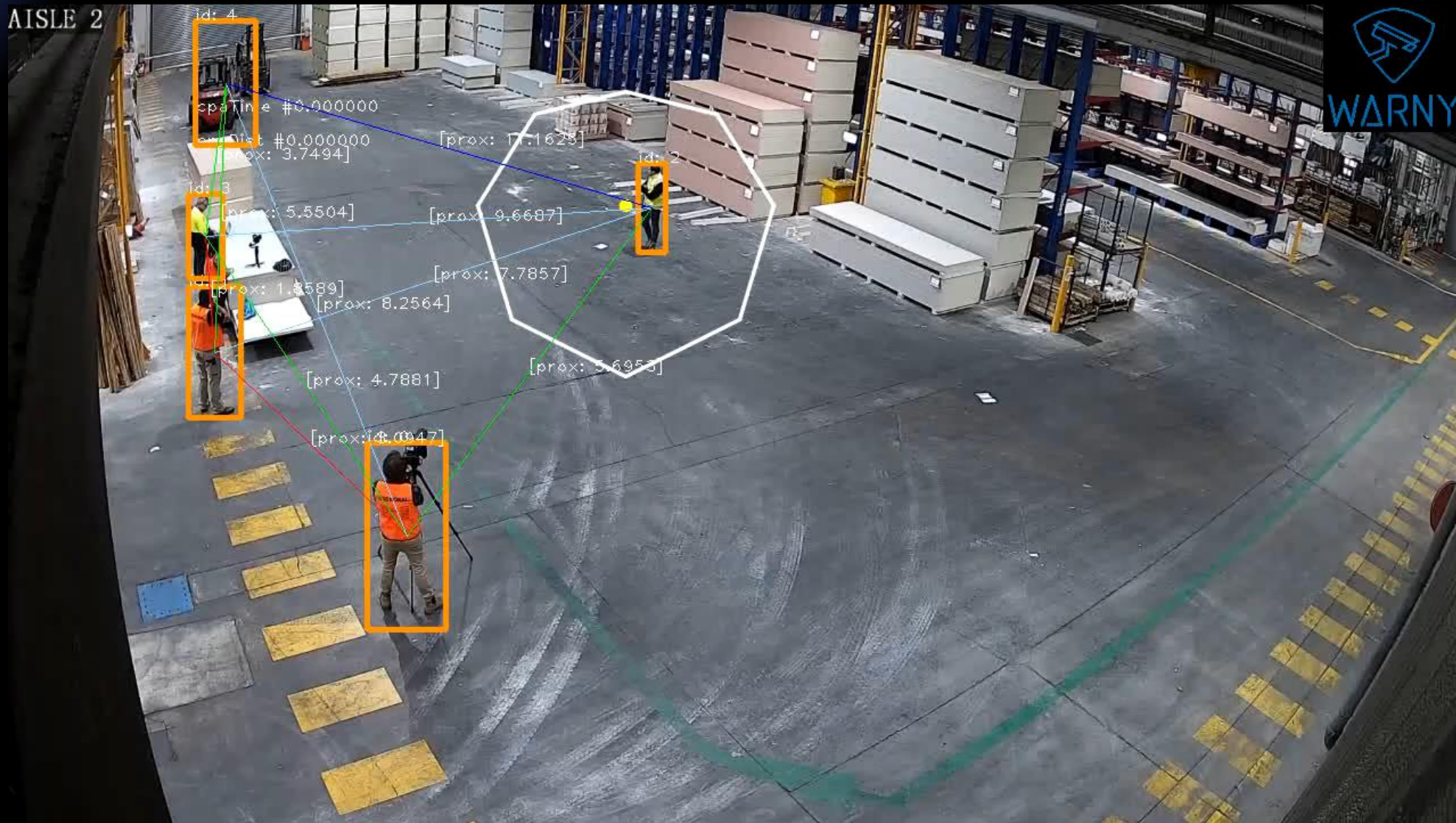


Optimized ground operations and expedited container inspection with increased data privacy and security with computer vision inference at the edge

A screenshot of the Deloitte Smart Logistics interface. The main view is an aerial camera feed of a port area with blue lines indicating camera coverage zones. On the right, there's a sidebar with controls and data. The 'Entry point' section has 'Date' set to 'Today' and 'Time' set to 'Last 24 hours'. Below it is a 'Video Stream' showing a green 'EVERGREEN' container. The 'Key statistics' section shows '243 Total Containers Arrived' and '23 Average Containers Arrived Per Hour'. The 'Container information' section has a search bar and a table with columns for 'Unique ID', 'Shipping Company', 'Arrival Time', and 'Status'.

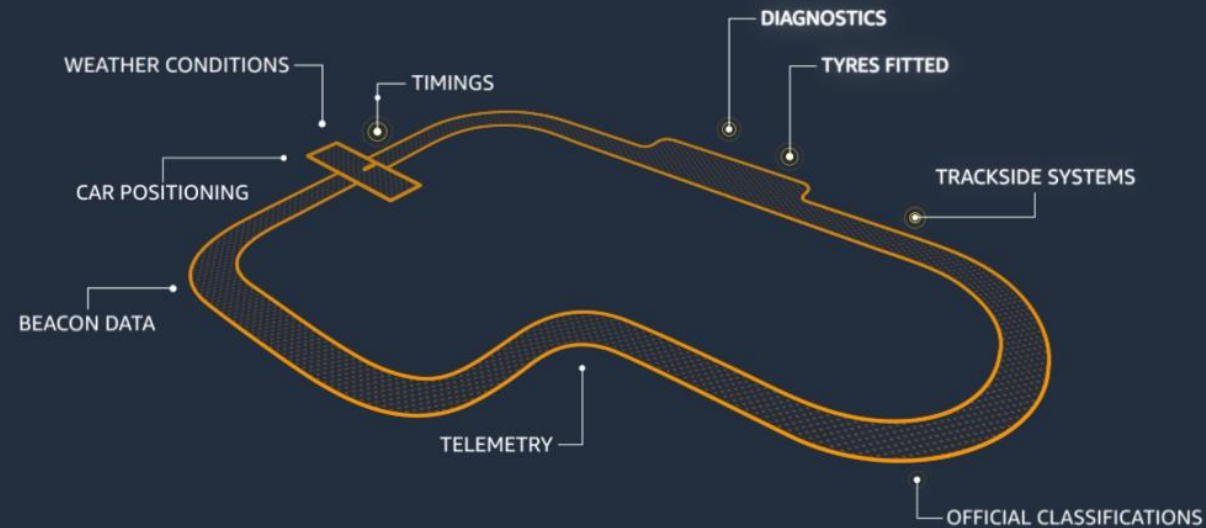
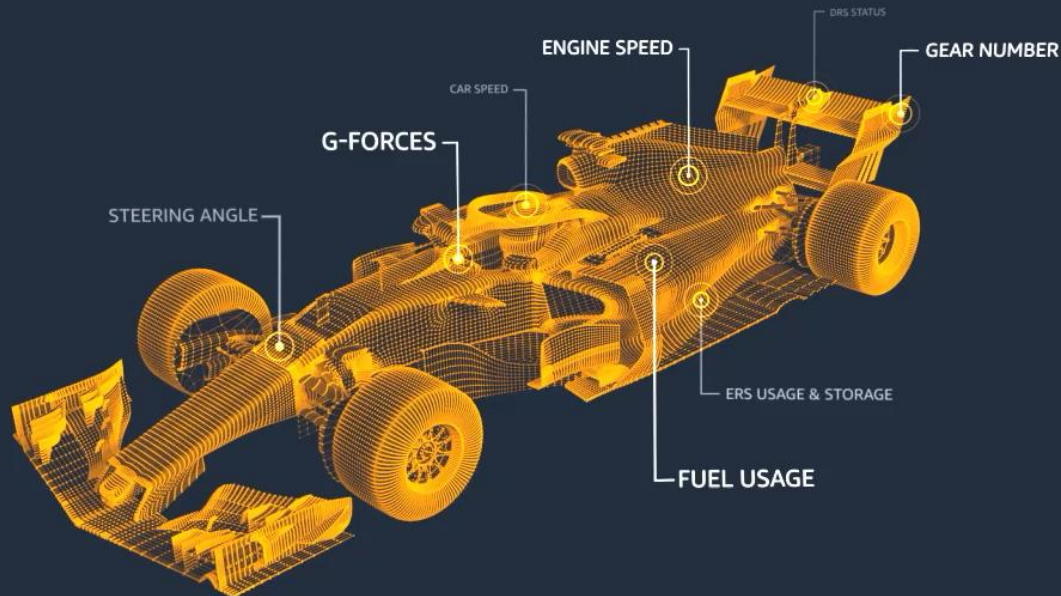
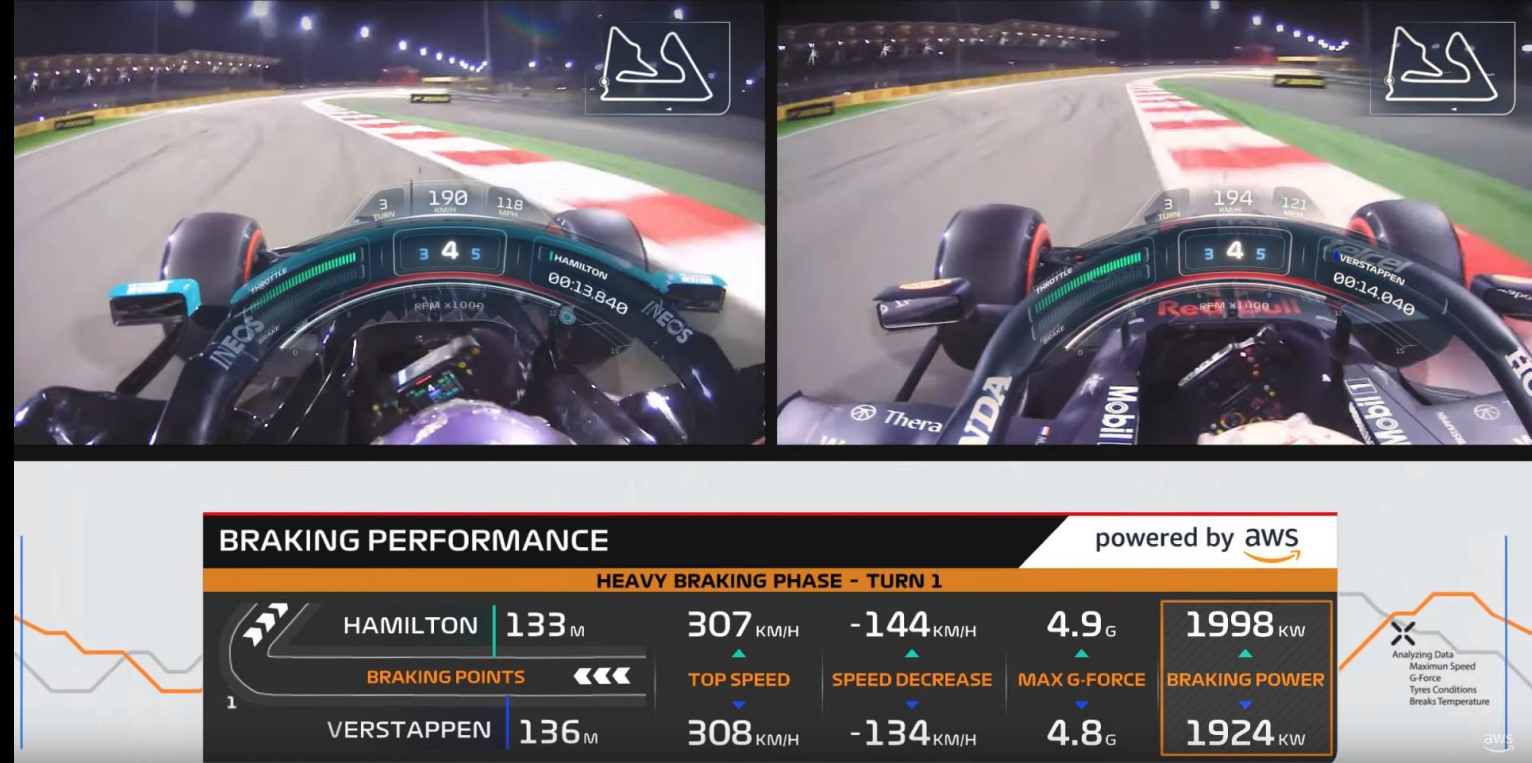
Unique ID	Shipping Company	Arrival Time	Status
Unique ID	Shipping Company	Arrival Time	Status
Unique ID	Shipping Company	Arrival Time	Status
Unique ID	Shipping Company	Arrival Time	Status
Unique ID	Shipping Company	Arrival Time	Status
Unique ID	Shipping Company	Arrival Time	Status

Bigmate Warny – Dynamic Geofencing



Formula 1 insights

Every F1 car contains **300 sensors** which **generate 1.1 million telemetry data points per second** transmitted from the cars to the pits. This real-time data is combined with over **70 years of historical race data** to extract rich insights that inform, educate, and **enrich the fan experience** and bring about more insight on the choice of race strategy that creates winning performances on the track



3D Point Clouds & Sensorfusion

Instructions Shortcuts Label View 3D Point Cloud Help

Point cloud View x

Top View x [C]

Side View x [C]

Front View x [C]

Camera_0 x [C] [F] [P] [B] [Slider]

Camera_1 x [C] [F] [P] [B] [Slider]

Camera_2 x [C] [F] [P] [B] [Slider]

Camera_3 x [C] [F] [P] [B] [Slider]

Label Category

Car

Labels (2)

Search labels

Label Id

Car:2

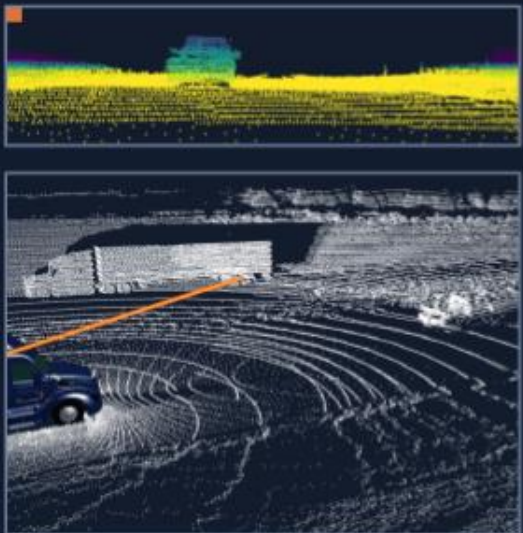
Cyclist:1

Q W E

A S D

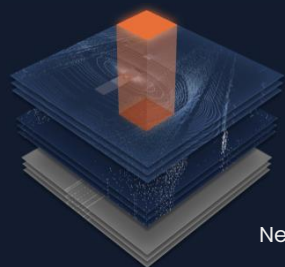
Point Size [Slider]

Save Submit

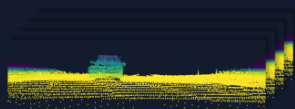


Complex object detection models for autonomous driving and needed to reduce training time to get models to market faster.

Reduced model development cycle time **from weeks to less than a day**, accelerating the pace of innovation and increased data scientist productivity by 10x.



Neural Convolution Engine



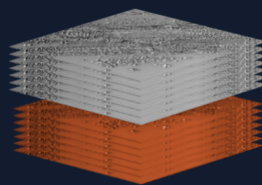
Range Conv. Engine



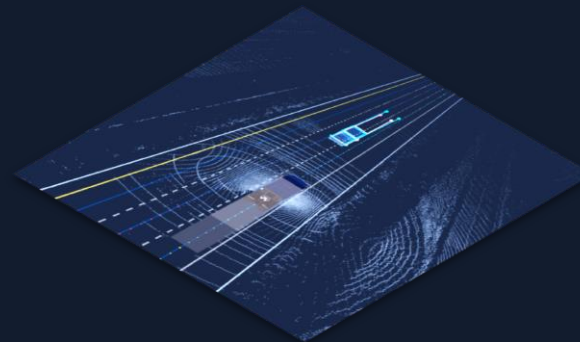
Image Fusion Engine



Euclidian Ray Scatter Engine



Neural Convolution Engine

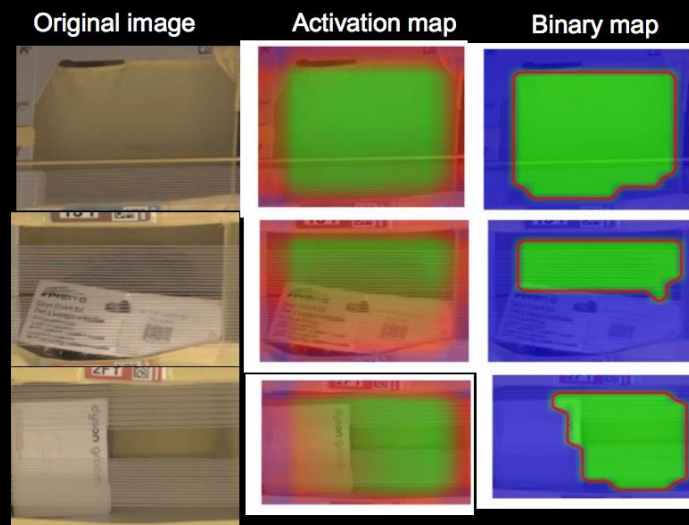
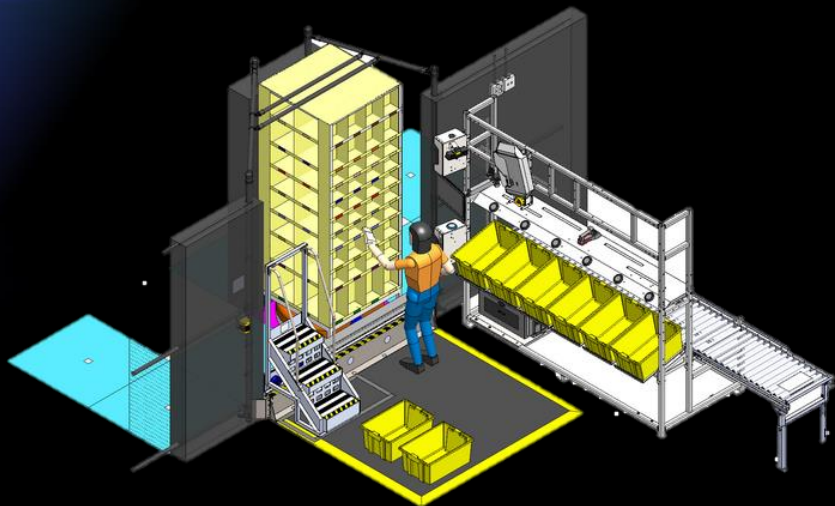




amazon

The Amazon logo, featuring the word "amazon" in a white, lowercase, sans-serif font. Below the text is a thick, orange, curved arrow that starts under the 'a' and points towards the 'n', resembling a smile.

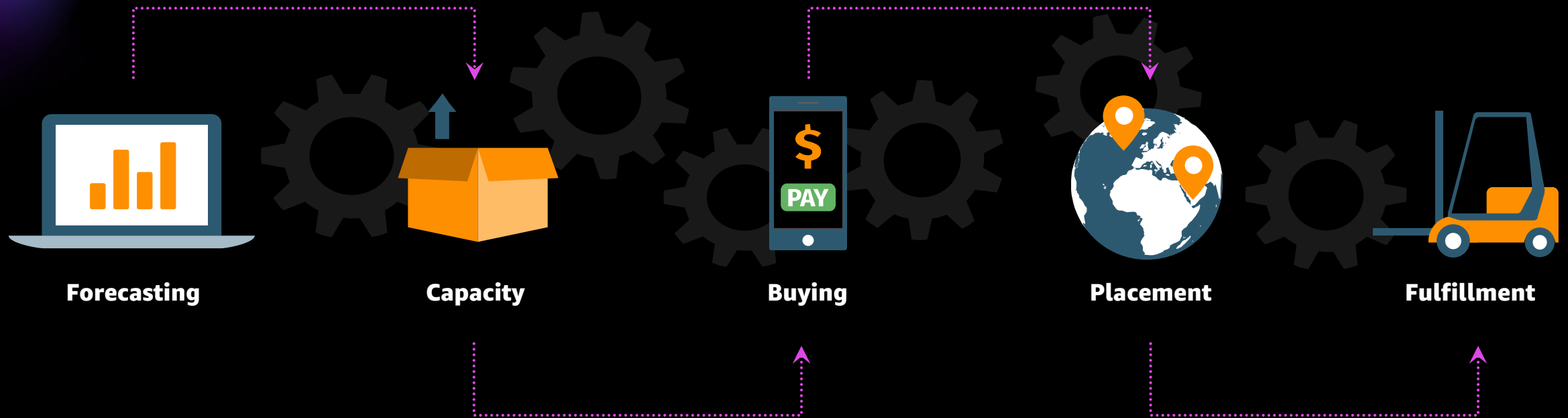




Computer Vision Implementation to identify empty space, detect object misplacements and computer assisted placement, **decreasing cognitive load**



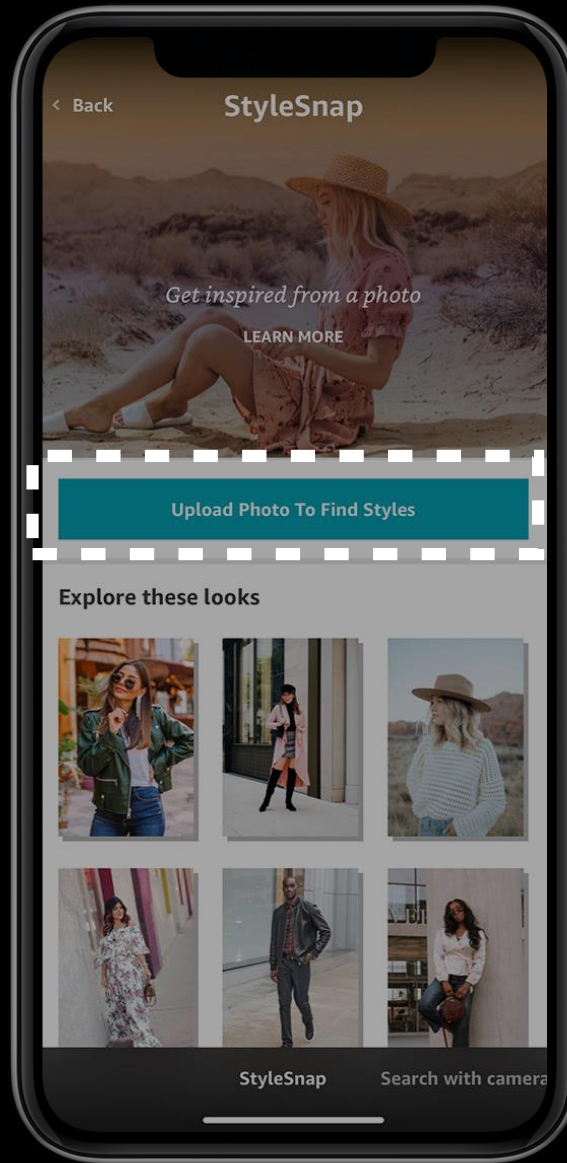
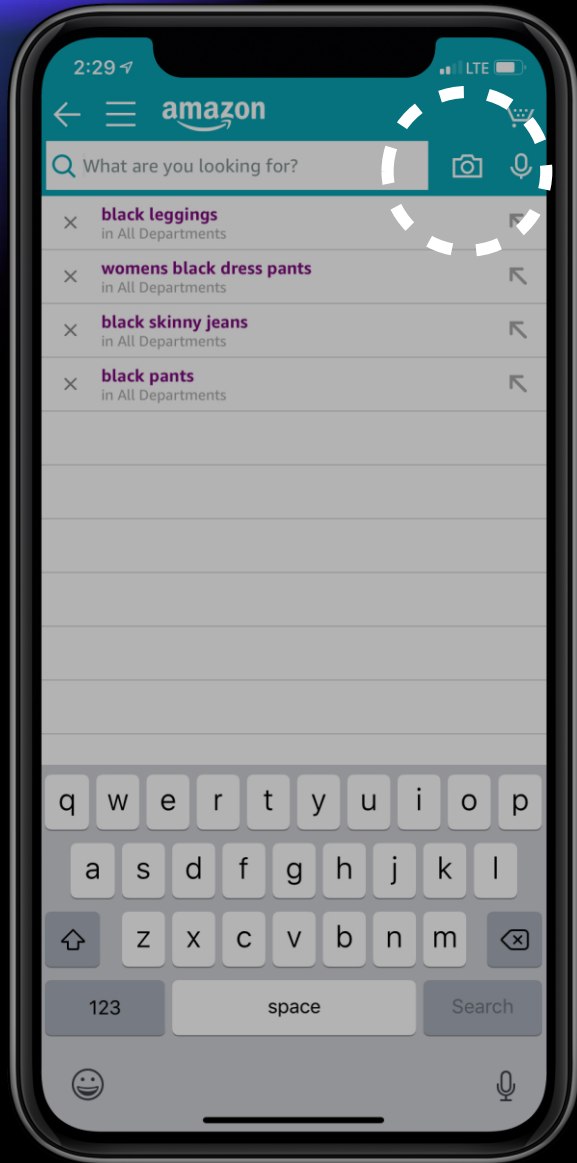
Amazon's Supply Chain – Automation End-to-End



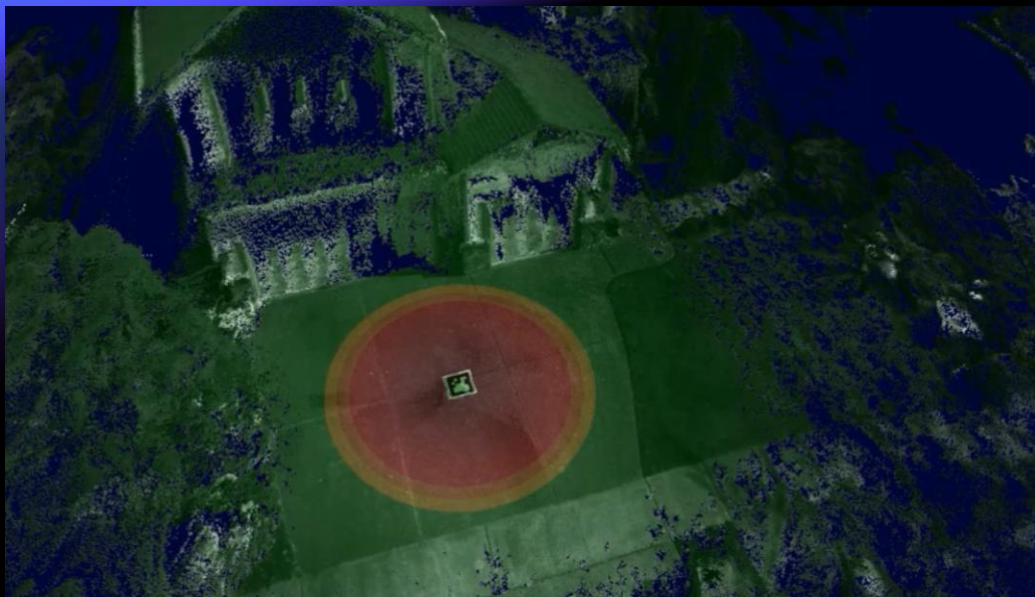
Demand forecasting
for over **400MM**
products daily

Amazon – **Over 99%**
of inventory is **ordered**
automatically

Automatic selection of the
optimal inbound carrier
and delivery location



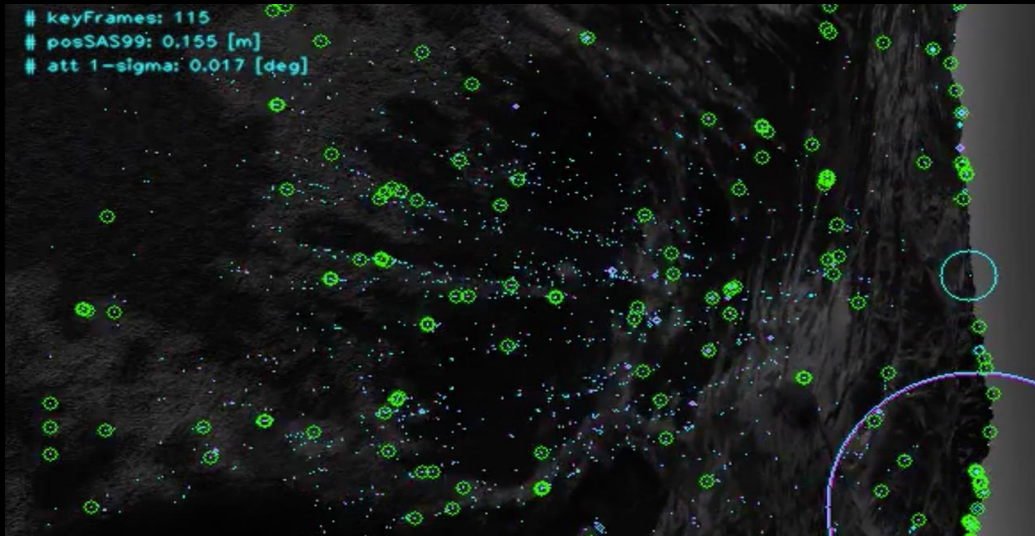




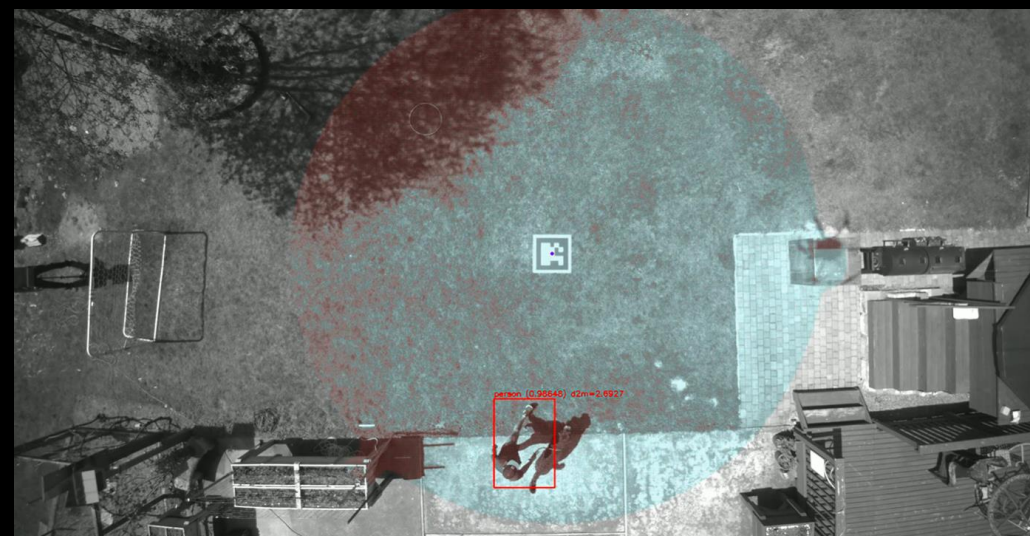
Photogrammetry



Image Segmentation

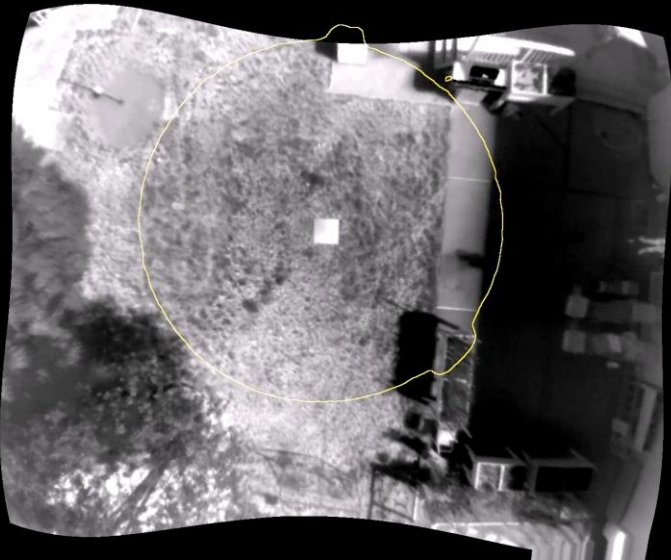


VSLAM

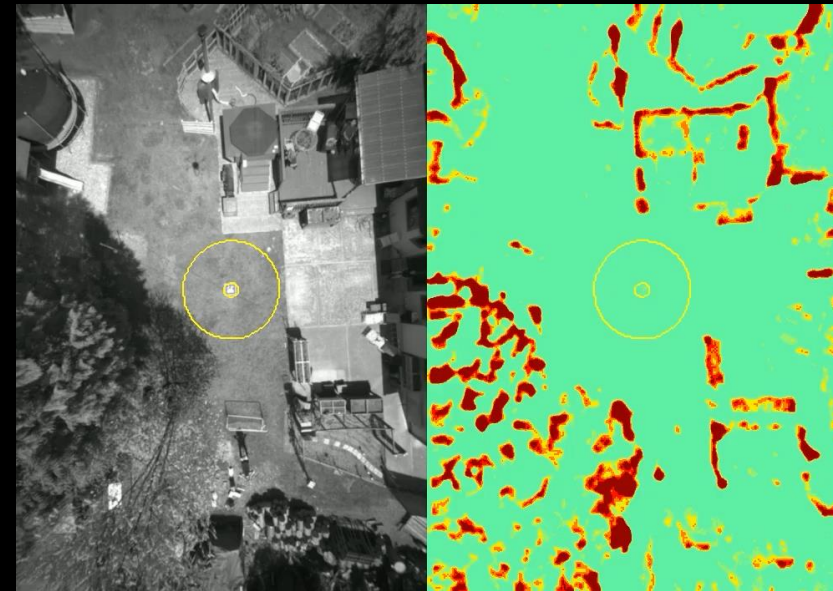
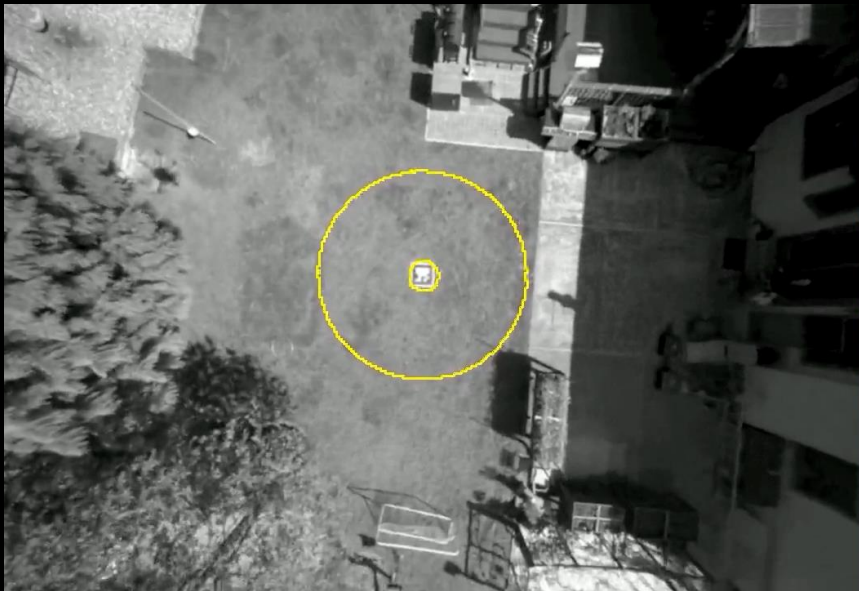


Deep Neural Networks

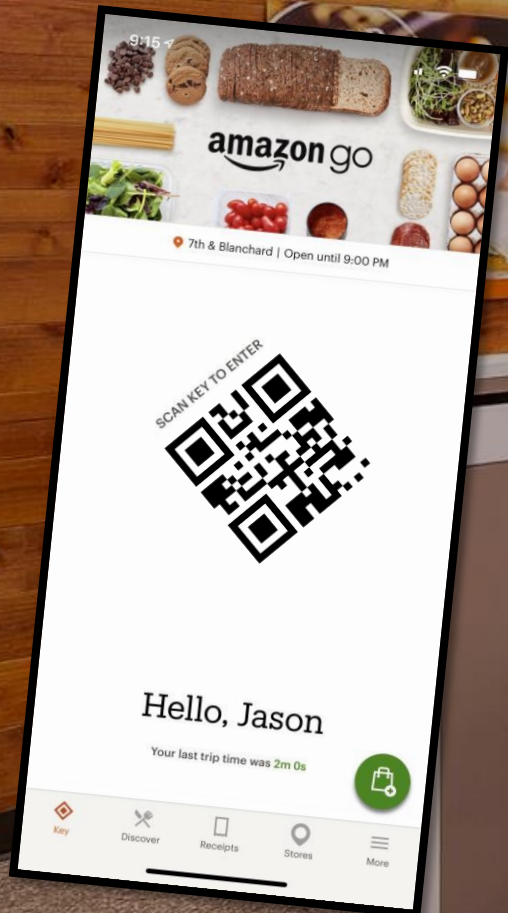
AGL = 9.24m

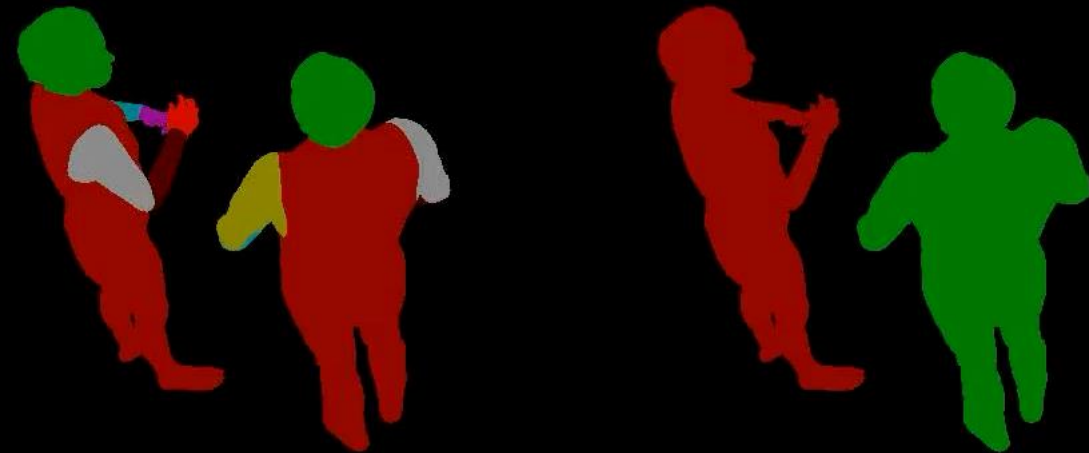


AGL = 9.24m



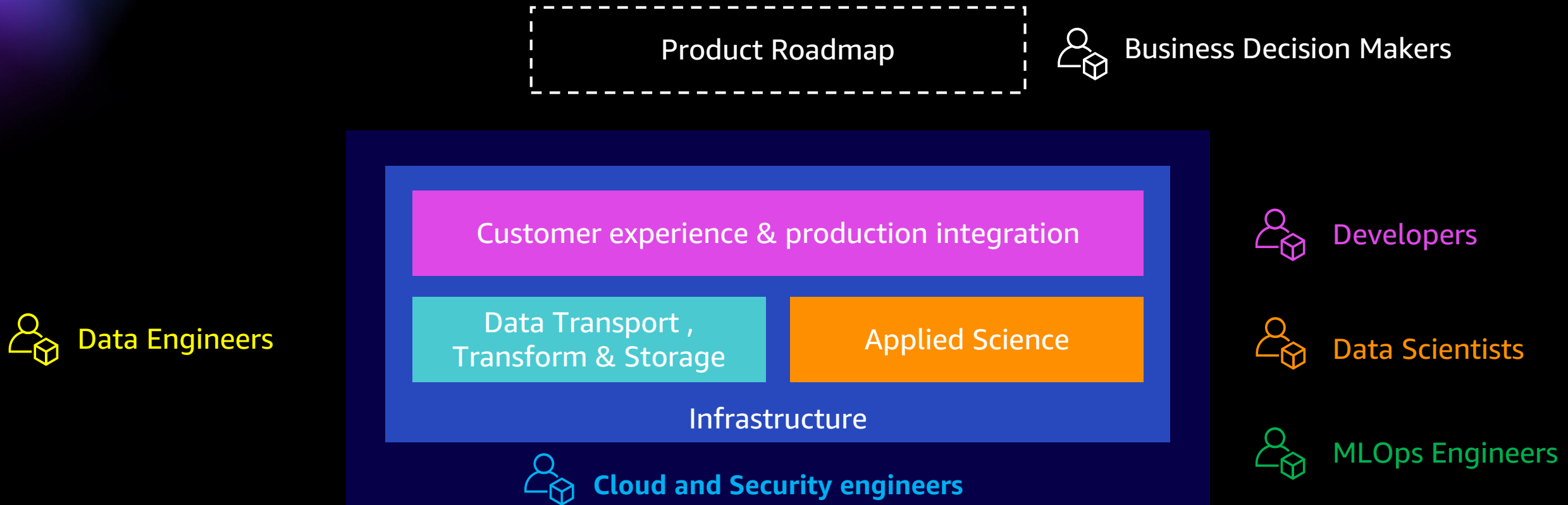
amazon go





Machine Learning **accessible to all**

It takes a diverse team to build a data product



- Other important personas in ML projects: BI analysts, Product Managers, Compliance experts
- Processes matter as well (feature prioritization, success metrics, security reviews, etc.)



Business Decision Makers

Amazon SageMaker **Canvas**

Create ML predictions without
any ML experience or writing
any code

GENERALLY AVAILABLE

Amazon **Quickstart Q**

Ask questions of your data and
receive answers with relevant
visualizations

GENERALLY AVAILABLE

Amazon QuickSight Q for Next Gen Stats



Player Offense ▾ Who has thrown the most deep TD passes since 2016?

sum of pass td ▾ from 2016 ▾ and reg ▾ by player name ▾ from dataset: Offense Player Q





Developers

Amazon **CodeGuru**

Automate code reviews and optimize application performance with ML-powered recommendations

GENERALLY AVAILABLE

Amazon **DevOps Guru**

ML-powered cloud operations service to improve application availability

GENERALLY AVAILABLE

Amazon **AI Services**

Easily add intelligence to applications. No machine learning skills required!



Data Scientists & ML Engineers

Amazon SageMaker **Studio**

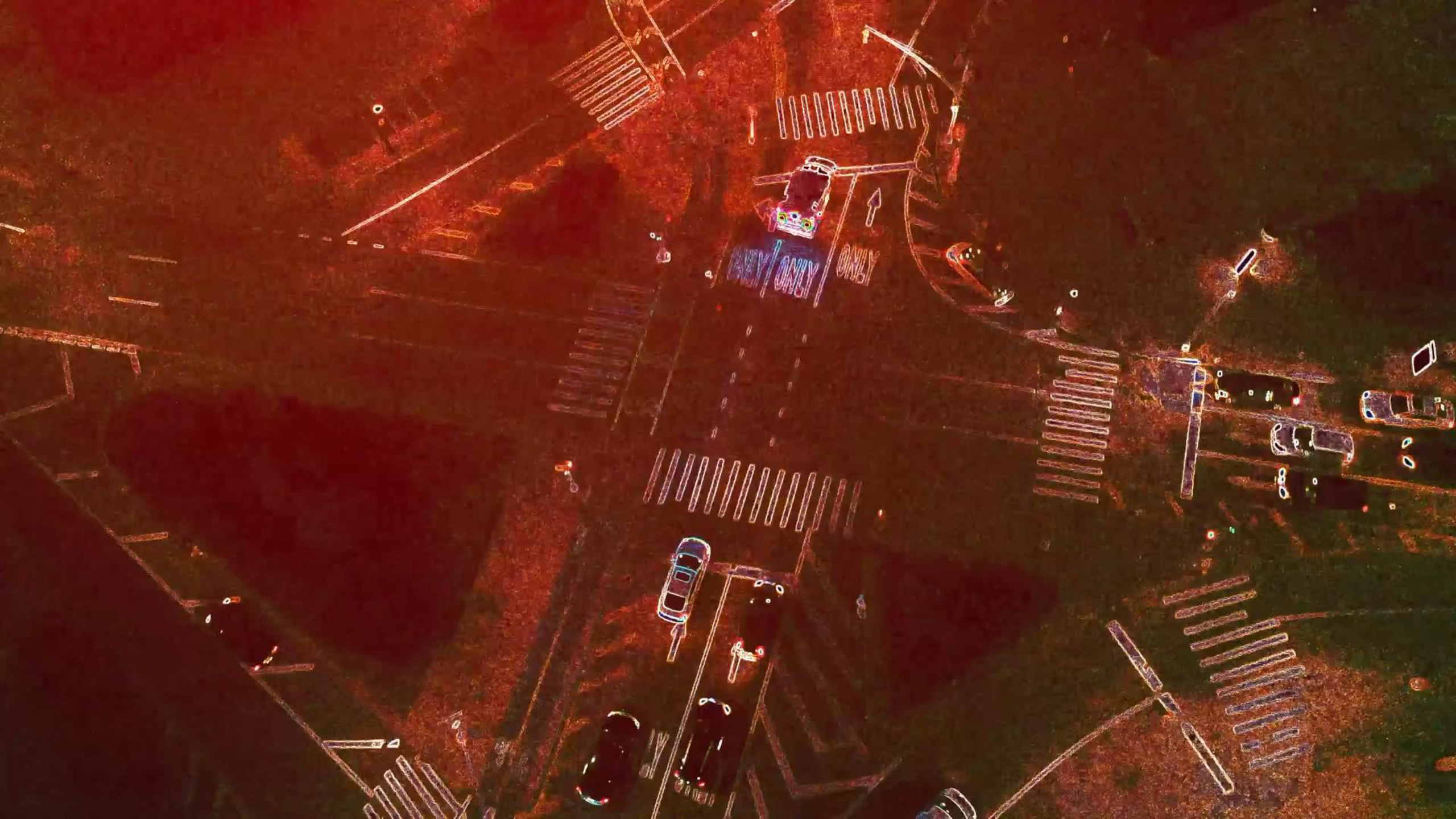
Integrated development environment (IDE) for the ML lifecycle. Prepare data, build, train, and deploy ML models

GENERALLY AVAILABLE

ML **Infrastructure** & **Frameworks**

High performance, cost-effective, scalable infrastructure for every workload. Customise algorithms with ML Frameworks

GENERALLY AVAILABLE





Data Engineers

Amazon SageMaker Ground Truth **Plus**

Deliver high-quality training datasets
fast, and reduce data labeling cost

GENERALLY AVAILABLE

ML for **database developers**

Bringing machine learning to
more builders through
databases and analytics services

GENERALLY AVAILABLE



Cloud and Security Engineers

Artificial Intelligence for IT Operations (AIOps)

Predictive Scaling for
Amazon EC2

Proactively provisions
EC2 capacity

Amazon CloudWatch
Anomaly Detection

Apply statistical & ML
algorithms analyse metrics

Amazon
GuardDuty

Protect AWS accounts
with intelligent threat
detection

Amazon
Macie

Discover and protect your
sensitive data at scale



MLOps Engineers

Amazon SageMaker Infrastructure Innovation

Amazon SageMaker **Training Compiler**

Accelerate model
training by 50%

Amazon SageMaker **Inference Recommender**

Reduce time to deploy
from weeks to hours

Amazon SageMaker **Serverless Inference**

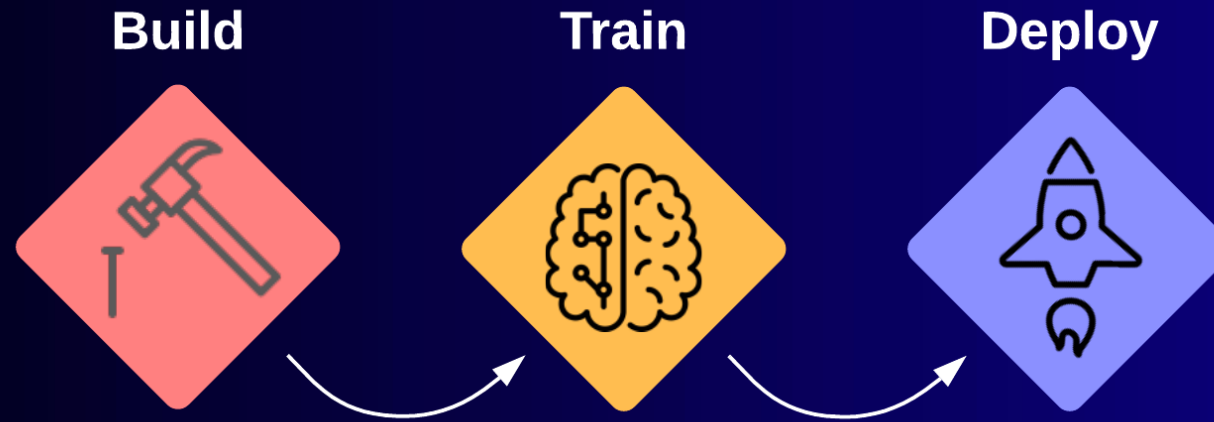
Lower cost of ownership
with pay-per-use pricing

Amazon SageMaker **Neo**

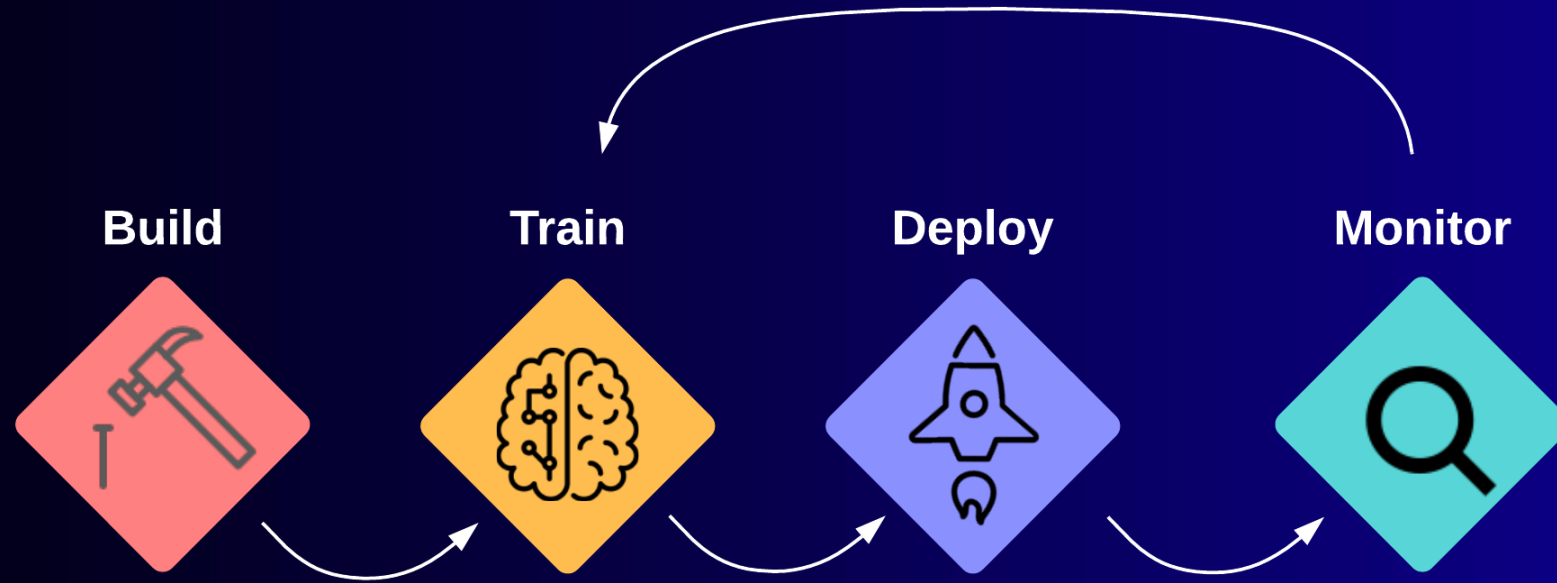
Run ML models anywhere
with better performance

Streamline the Machine Learning **lifecycle**

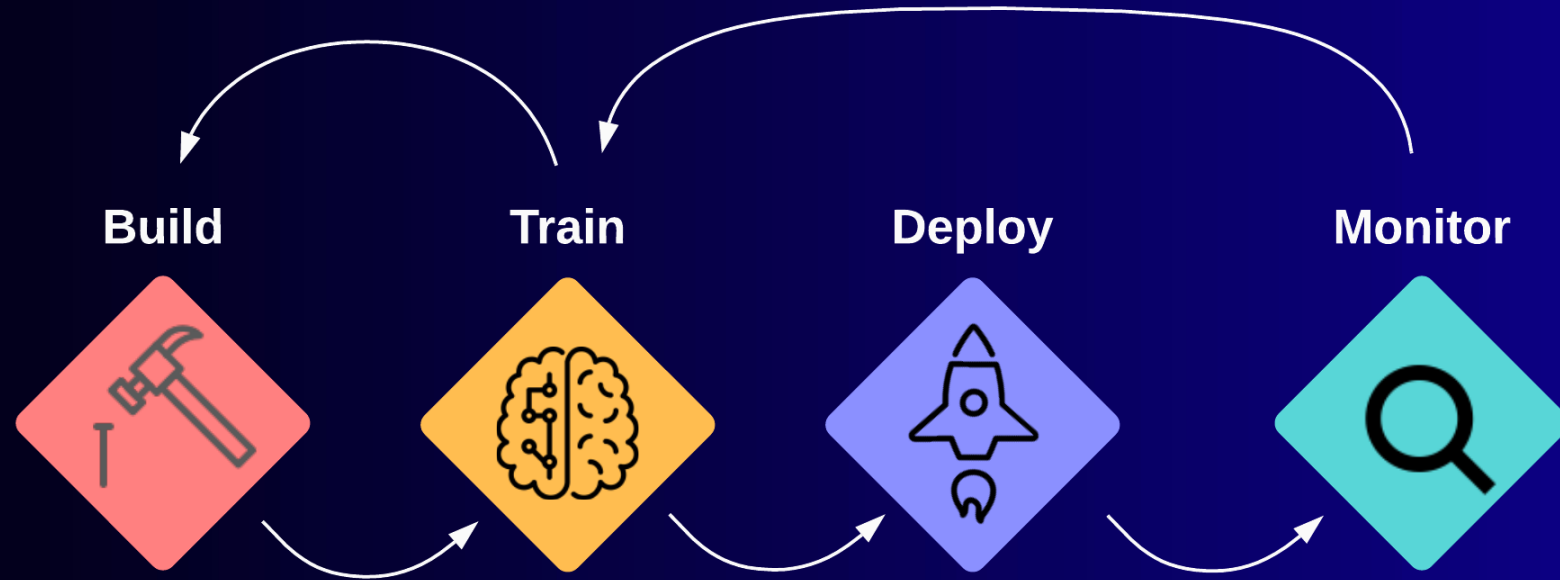
The development cycle of ML



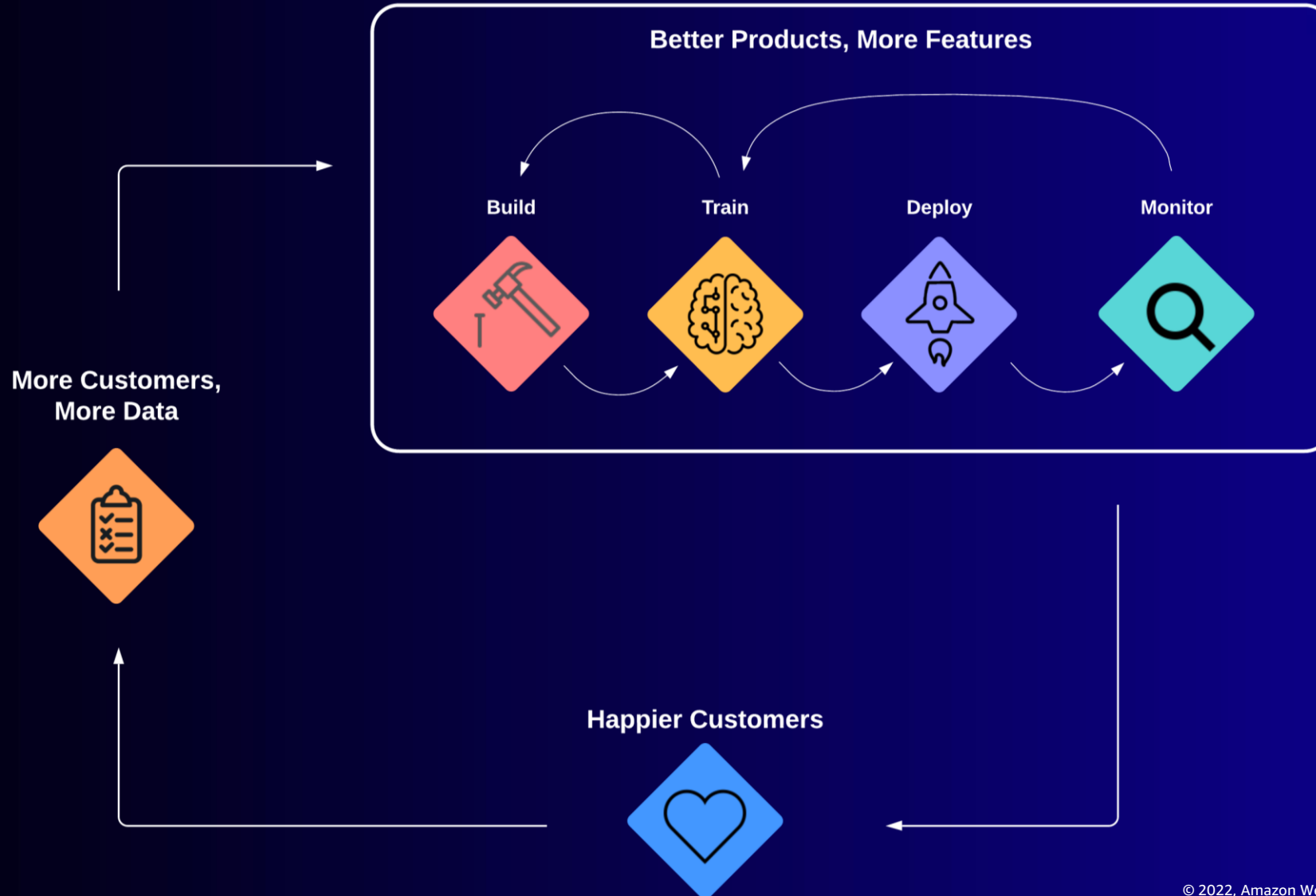
The development cycle of ML



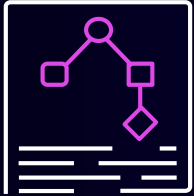
The development cycle of ML



The development cycle of ML



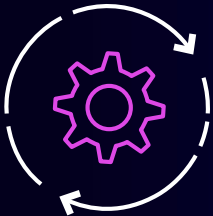
Key tenets to optimize your MLOps lifecycle



Create automated and reproducible ML workflows



Manage models with a model registry



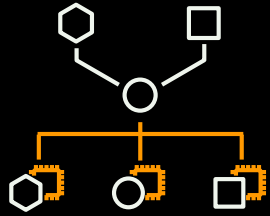
Enable continuous delivery with Infrastructure-as-Code and CI/CD pipelines



Monitor performance and feedback information to your models

Amazon SageMaker Pipelines

Managed machine learning CI/CD service



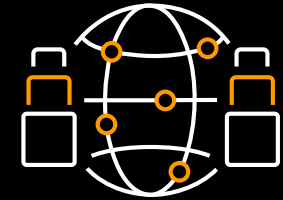
Centrally **manage**
each step of
the workflow



Share and
re-run
workflows

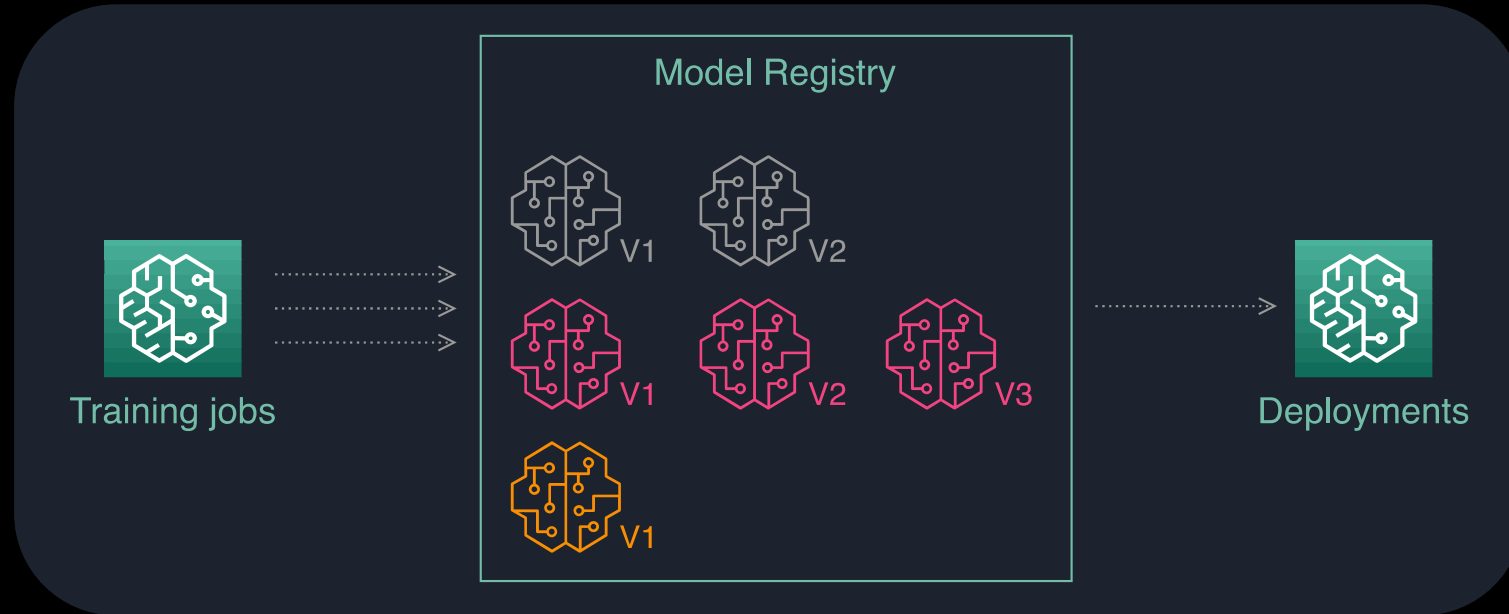


Choose from
built-in
templates



Compare
workflows
visually

Manage your models with a model registry



A **centralized tracking system** that stores lineage, versioning, and related metadata for published machine learning models. A registry may **capture governance data** required for auditing purposes, such as who trained and published a model, which datasets were used for training, the values of metrics measuring predictive performance, and when the model was deployed to production.

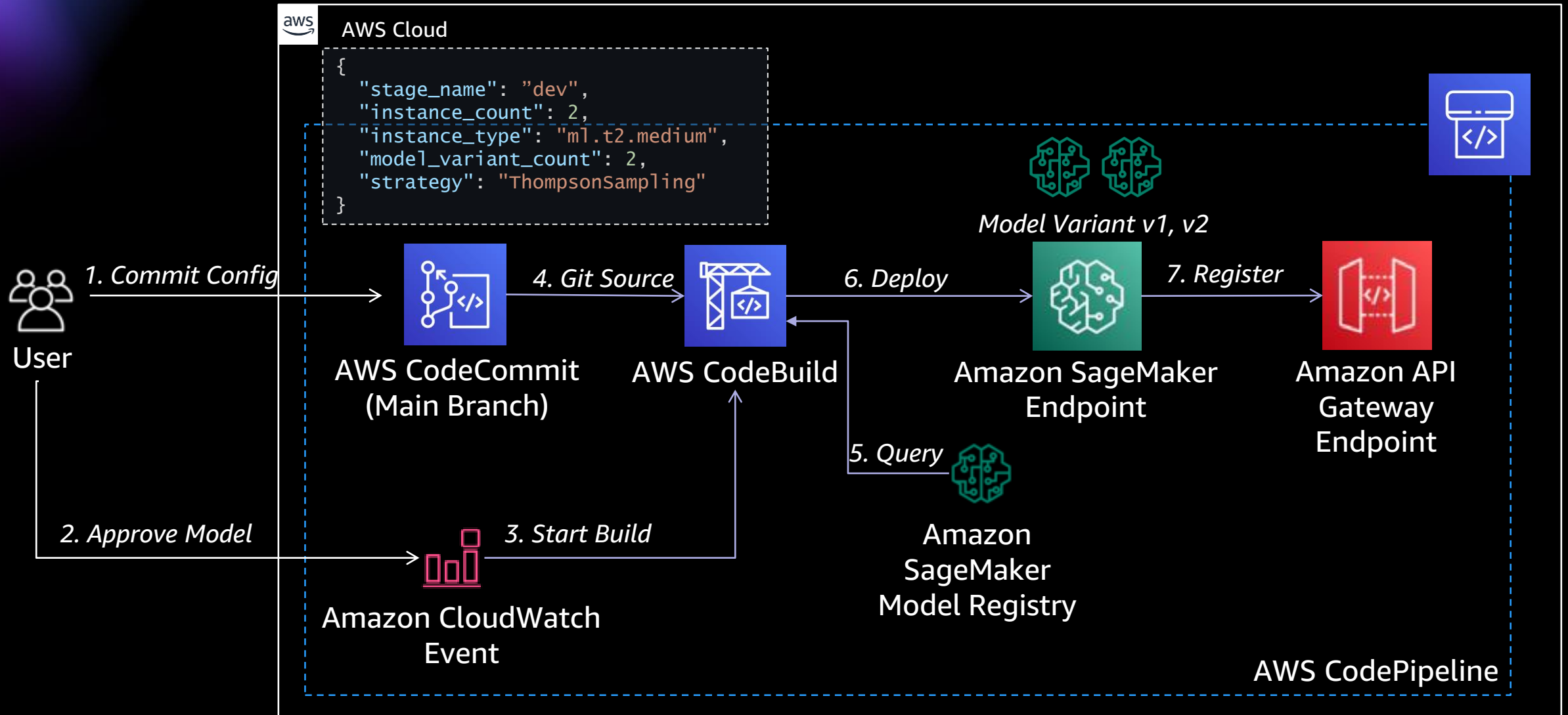
Approve model for deployment

The screenshot displays the Amazon SageMaker Studio interface. On the left, the 'Components and registries' sidebar shows a 'Model registry' section with a search bar and a 'Create model group' button. Below this is a table of model groups. The main panel is titled 'Recommendations Model - Latin America' and shows a list of model versions. A context menu is open for version 5, displaying options to 'Open model version' and 'Update model version status...'.

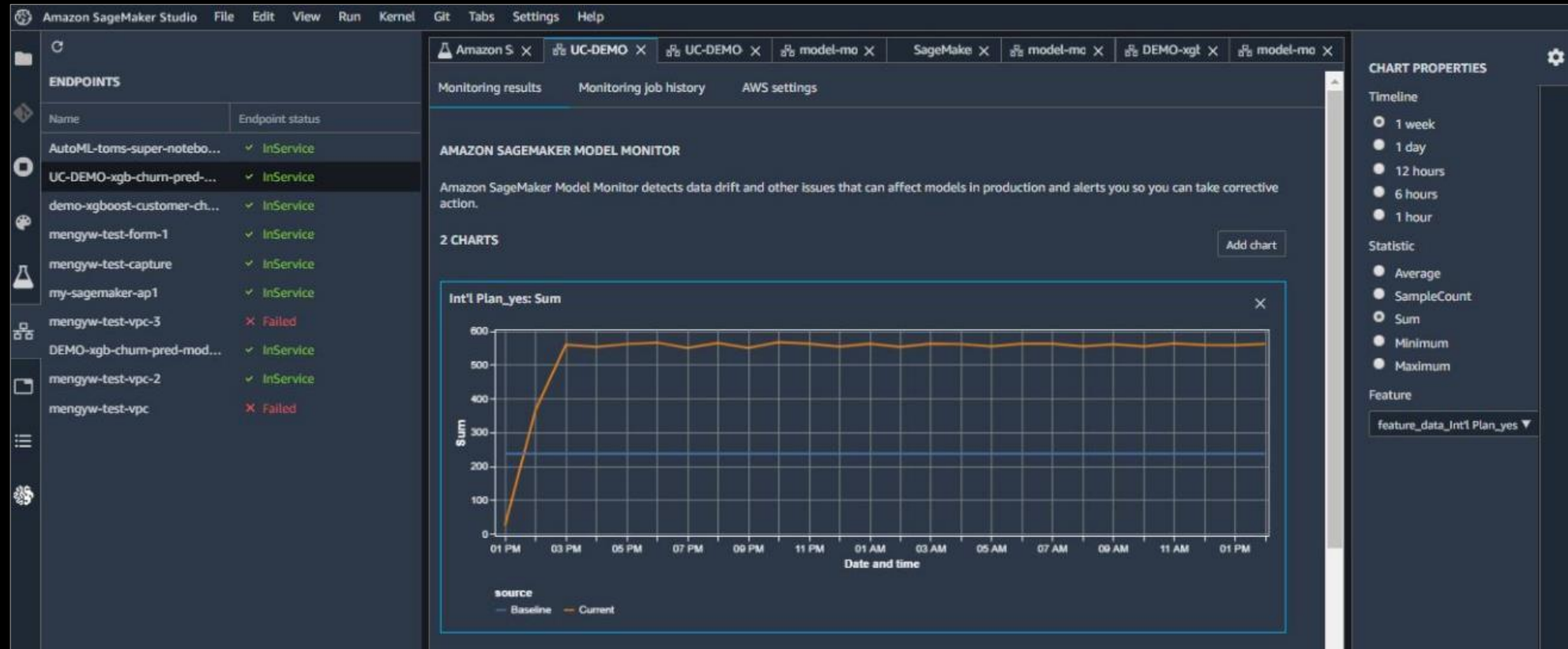
Name	Modified on	Actions
Recommendatio Model...	Today	
Recommendatio Model...	Yesterday	
Recommendatio Model...	10/22/20	
Recommendatio Model...	09/12/20	

Name	Status	Step	Description	Status updated by	Modified on	Actions
version 6	Pending	Staging	New model with SKLe...	Pending	10/10/20	
version 5	Approved	Production	Model updated on 8/1...	Jen Cabro		Open model version ↗ Update model version status...
version 4	Approved	Archived	Model updated on 7/15...	Jen Cabro		
version 3	Approved	Archived	Model updated on 6/15...	Jen Cabro		
version 2	Approved	Archived	Model built on 5/15/20...	Jen Cabro	10/10/20	
version 1	Approved	Archived	Model built on 4/15/20...	Jen Cabro	10/10/20	

Deploy a multi-variant endpoint

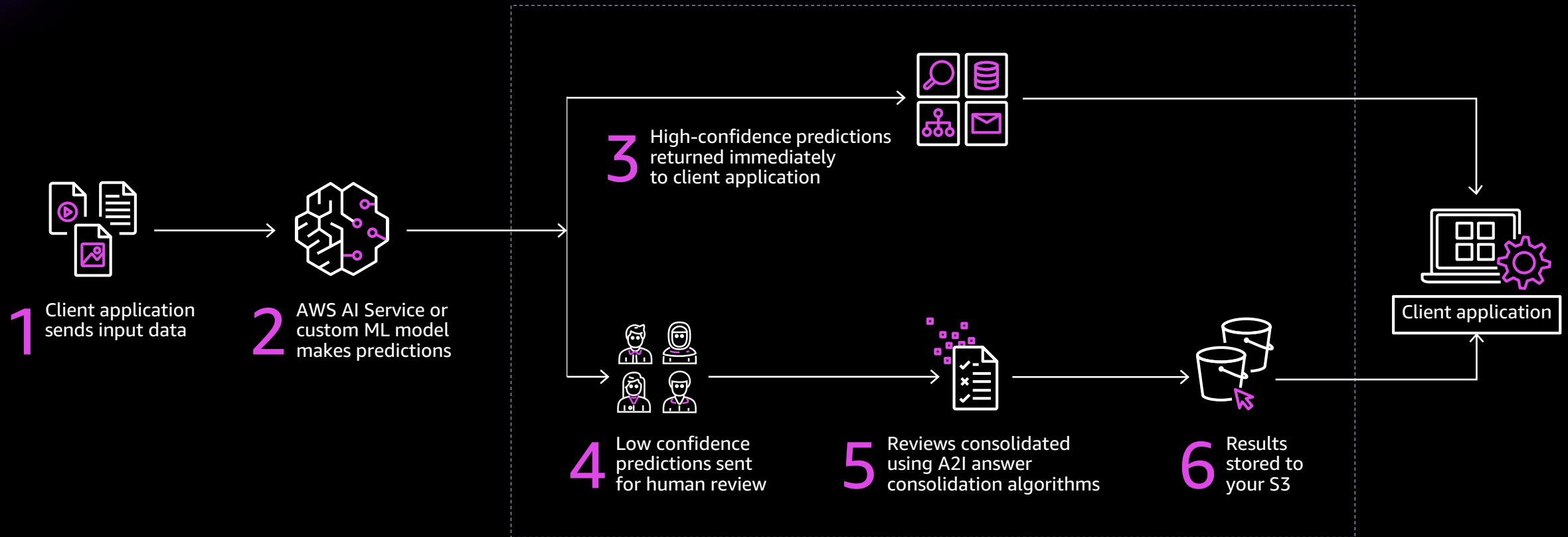


Identify model drift with Amazon SageMaker Model Monitor and take action



Amazon SageMaker Model Monitor automatically monitors ML models in production, and alerts you when data quality issues appear

Build human review workflows required for your ML predictions with Amazon Augmented AI



Balancing needs of ML builders and Cloud IT

ML Builders
Stay agile



Innovate with the speed and agility of AWS

Self-service access
Experiment fast
Respond quickly to change

Cloud IT
Establish governance



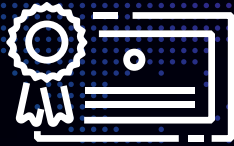
Govern at scale with central controls

Security & Compliance
Operations
Spend management

29 Million trained by 2025



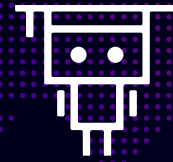
DeepRacer
League



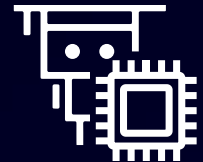
Training &
Certification



Amazon SageMaker
Studio Lab



Machine Learning
University



Machine Learning
Nanodegree

Visit the AI & Machine Learning resource hub for more resources

Dive deeper into these resources, get inspired and learn how you can use AI and machine learning to accelerate your business outcomes.

- The machine learning journey e-book
- 7 leading machine learning use cases e-book
- A strategic playbook for data, analytics, and machine learning e-book
- Accelerate machine learning innovation with the right cloud services & infrastructure e-book
- Choosing the right compute infrastructure for machine learning e-book
- Improving service and reducing costs in contact centers e-book
- Why ML is essential in your fight against online fraud e-book
- ... and more!



<https://bit.ly/3mwi59V>

Visit resource hub

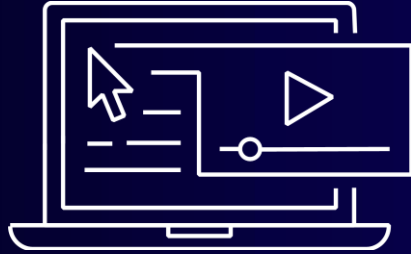
AWS Machine Learning (ML) Training and Certification



AWS is how you build machine learning skills

Courses built on the curriculum leveraged by Amazon's own teams. Learn from the experts at AWS.

aws.training/machinelearning



Flexibility to learn your way

Learn online with on-demand digital courses or live with virtual instructor-led training, plus hands-on labs and opportunities for practical application.

explore.skillbuilder.aws/learn



Validate your expertise

Demonstrate expertise in building, training, tuning, and deploying machine learning models with an industry-recognized credential.

aws.amazon.com/certification

Thank you for attending AWS Innovate – AI/ML Edition

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



twitter.com/AWSCloud



facebook.com/AmazonWebServices



youtube.com/user/AmazonWebServices



slideshare.net/AmazonWebServices



twitch.tv/aws

Overview



Live Keynote and Closing

FOR BUSINESS

2 RETHINK POSSIBLE TRACKS

Rethink possible:
Innovation stories

Rethink possible:
Accelerate AI & ML innovations

FOR BUILDERS

6 CORE CONCEPTS TRACKS

AI/ML use
case solutions

Build, train &
deploy ML models

ML
infrastructure

Data infrastructure
for ML workloads

Scaling
with ML

Uplevel your
ML skills

3 LOCALIZED

LANGUAGE
AGENDA

Japanese

Korean

Bahasa Indonesia

ACTIVITIES



50+ breakout
sessions



Builders
zone



Hands-on
demos & labs



Customer
stories



AWS e-books
& guides



Ask the
experts




AGENDA AT A GLANCE

60 mins	Keynote: Innovate faster and build a smarter business with AI & ML on AWS					
	For Business: Rethink Possible	For Builders: Create Tomorrow				
	Rethink Possible: Innovation Stories	AI/ML use case solutions	Build, train & deploy ML models	ML infrastructure	Scaling with ML	AI & ML showcase
30 mins	Rethink Possible: Innovation stories	Modernize your CX Innovation with ML-powered Amazon Connect	Get started with Amazon SageMaker in minutes	Rapidly launch ML solutions at scale on AWS infrastructure	Setting up secure and well-governed ML environments on AWS	<div>• Ask the Experts</div> <div>• Customer stories</div> <div>• Builders Zone</div> <div>• Hands-on labs</div>
		LEVEL 200	LEVEL 100	LEVEL 200	LEVEL 200	
30 mins		Build an intelligent document processing and search solution at scale	Data preparation: Using Amazon SageMaker and AWS Glue DataBrew	Build an end-to-end machine learning platform using Kubeflow on Amazon EKS	A no-code alternative to quickly predict churn using Amazon SageMaker Canvas	
		LEVEL 200	LEVEL 300	LEVEL 300	LEVEL 200	
30 mins		Fraud detection on AWS	Build machine learning models with Amazon SageMaker efficiently	Machine learning Inference with AWS Lambda and Amazon EFS	Solving challenges with AI/ML powered by Intel	
		LEVEL 100	LEVEL 200	LEVEL 200	LEVEL 200	
Break	Rethink Possible: Accelerate AI & ML Innovations	AI/ML use case solutions	Build, train & deploy ML models	Data infrastructure for ML workloads	Uplevel your ML skills	
30 mins	Rethink machine learning for regulated industries	Easily build custom computer vision models using Amazon Rekognition	Train ML models quickly and cost-effectively with Amazon SageMaker	Sentiment analysis using Amazon Aurora machine learning integration	Getting started to learn and experiment ML with Amazon SageMaker Studio Lab	
	LEVEL 100	LEVEL 200	LEVEL 300	LEVEL 200	LEVEL 100	
30 mins	Rethink machine learning for supply chain	Using AI/ML to automate content moderation and compliance	Bias detection and explainability in ML	Integrate easy-to-use machine learning into your analytics workloads	Applying AWS machine learning to next-gen DevOps	
	LEVEL 100	LEVEL 200	LEVEL 200	LEVEL 200	LEVEL 300	
30 mins	Reinventing Industrial operations with AI and ML	Using Hugging Face models on Amazon SageMaker	Rightsizing and optimizing cost of Amazon SageMaker endpoints using serverless deployments and automated instance type recommendations	Democratize machine learning with Amazon Redshift ML to drive employee attraction, not attrition	Machine learning for developers	
	LEVEL 100	LEVEL 200	LEVEL 200	LEVEL 300	LEVEL 200	
30 mins	AI and ML product development: How you can leverage AI and create productive teams to deliver transformative experiences	Running Intelligent applications at edge using AWS AI/ML	Implementing MLOps with Amazon SageMaker	Accelerate your Amazon SageMaker model training with Amazon FSx for Lustre and Amazon S3	Start your engines with AWS DeepRacer	
	LEVEL 100	LEVEL 200	LEVEL 300	LEVEL 300	LEVEL 300	
45 mins	Closing: Accelerate rapid innovation with AI & ML					

Thank you!

Dean Samuels

Chief Technologist, ASEAN
Amazon Web Services

 @the_cloud_ninja

Olivier Klein

Chief Technologist, APJ
Amazon Web Services

 @captainklein