AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

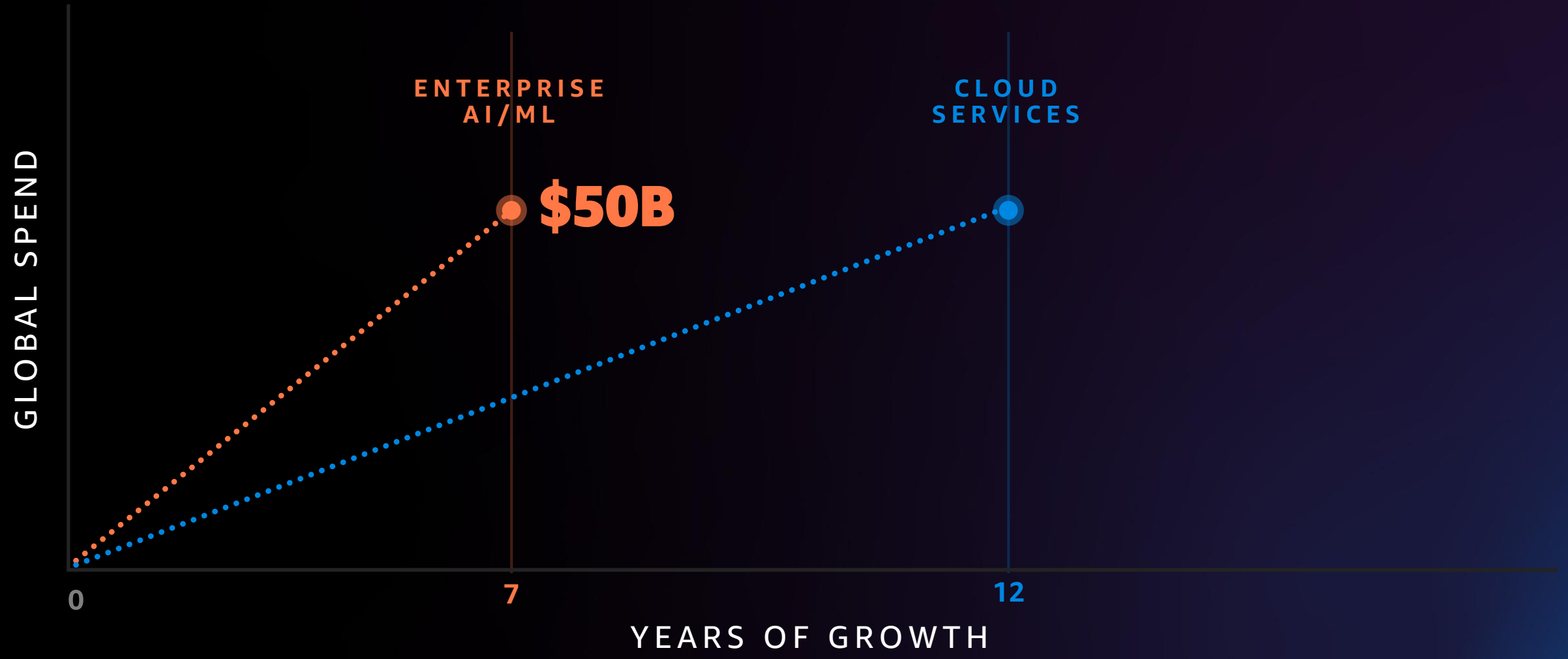
AIM212-L

Accelerate innovation with machine learning

Dr. Bratin Saha (he/him)
Vice President of Artificial Intelligence and Machine Learning Services
AWS

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Machine learning is key to innovation



Source: IDC's Worldwide Semiannual Artificial Intelligence Spending Guide, Publication August 2021; IDC Semiannual Public Cloud Services Tracker, 1H2021, November 11, 2021
Note: Enterprise AI/ML and Cloud Services (infrastructure and platform services) categories are not mutually exclusive.

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.



More than 100,000 customers use AWS for machine learning

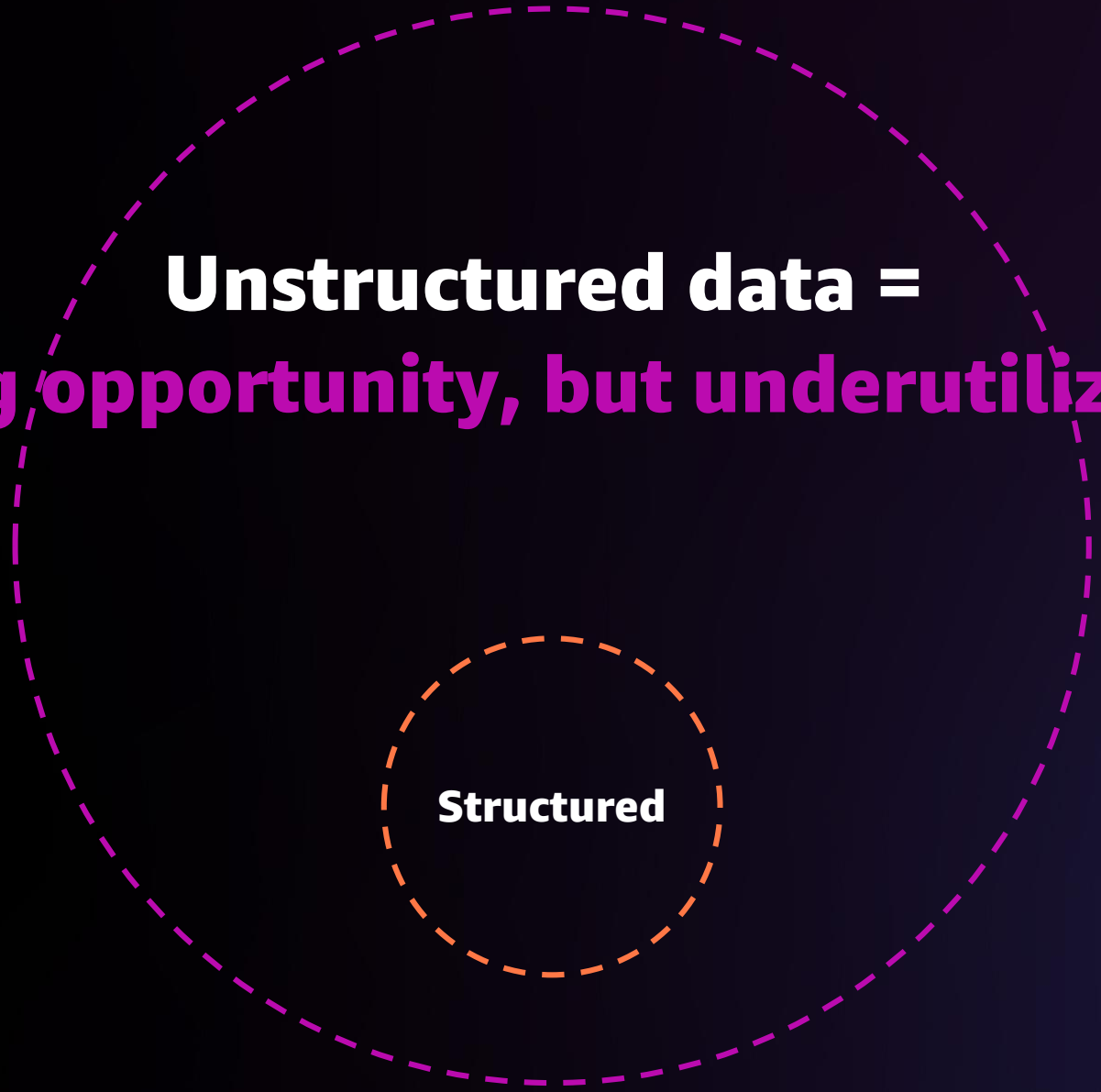
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ML INNOVATION DRIVER #1:

Use ML to push the frontier of what's possible

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Unstructured data =
big opportunity, but underutilized

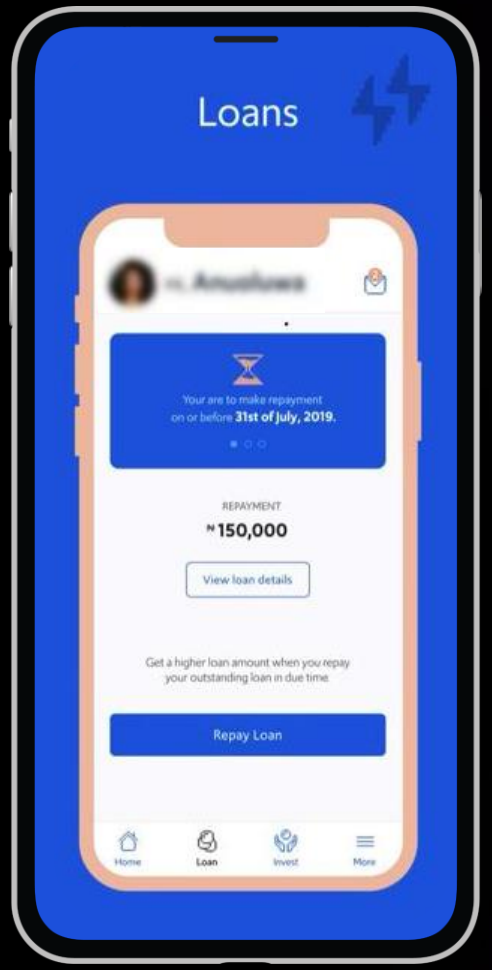


© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.


Machine learning is inspired
by the human brain

Pushing the frontier of what's possible

AWS VISION AI SERVICES



aello

Uses Amazon Rekognition to verify loan applicant's identity

Near-instant loan approvals in emerging economies

Serves over 2 million individuals and micro businesses

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Challenges of ID verification

GENERALLY AVAILABLE

Analyze ID

Extract information from different types of IDs with high accuracy in near real-time

NEW

Supports a wide range of id formats

Automatically maps extracted data to named fields

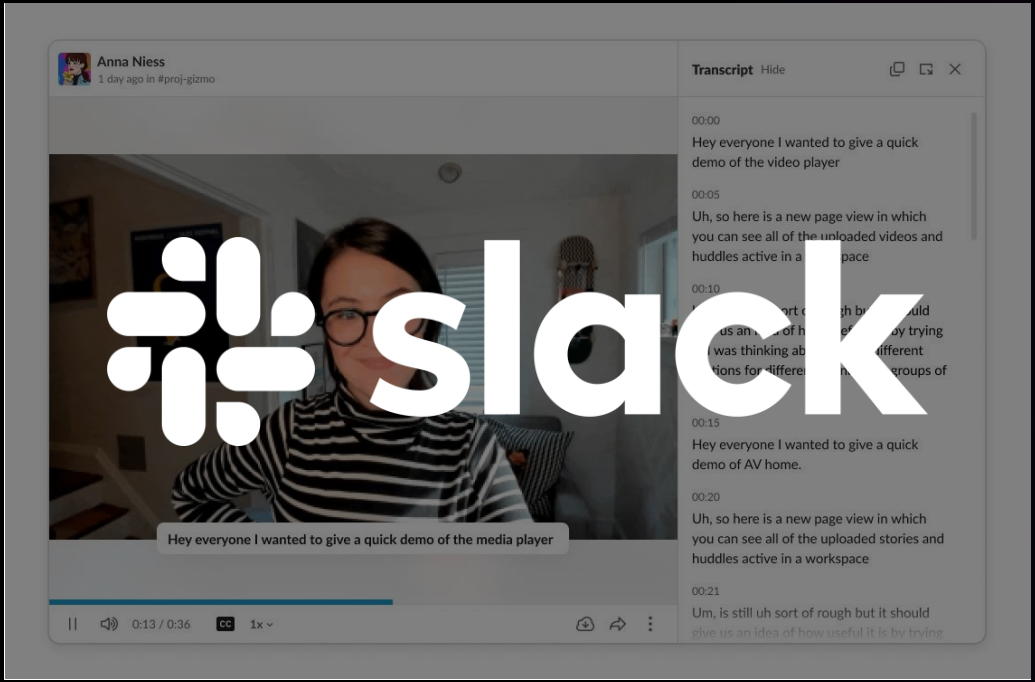
No additional post-processing required

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Pushing the frontier of what's possible

AWS LANGUAGE AI SERVICES



Uses Amazon Transcribe for voice and video messages

Captions provide accessibility

Searchable transcripts

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Our mission:

To create a seamless and connected healthcare world where everyone thrives

We serve

650,000

physician users worldwide

Clients in

35 countries

3M+

daily users

262M+

longitudinal records in Cerner systems

$800M

in R&D per year

600+ patents

More U.S. patents than all core competitors combined

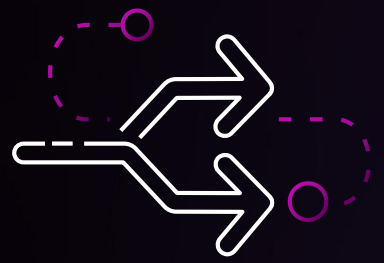
aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Machine learning at Cerner



Empower caregivers to deliver better patient outcomes



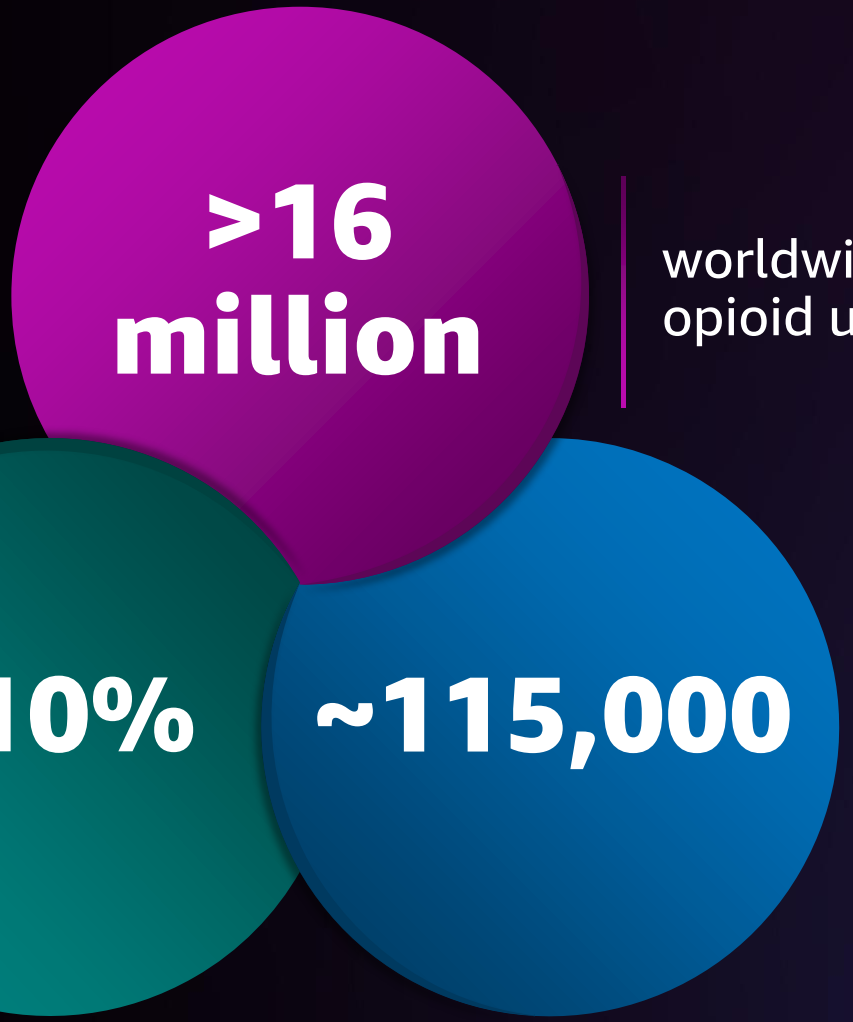
Use ML for predictive diagnostics to identify health risks early



Embed ML into clinical workflows for point of care assessments

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The opioid crisis



worldwide are affected by opioid use disorders (OUD)

seek or receive treatment for OUD

died of opioid overdose in 2017

Source: CDC | World Health Organization

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The opioid use disorder predictor

An ML model to identify a patient's risk for OUD

Designed for emergency departments and social workers

Over 40 input features

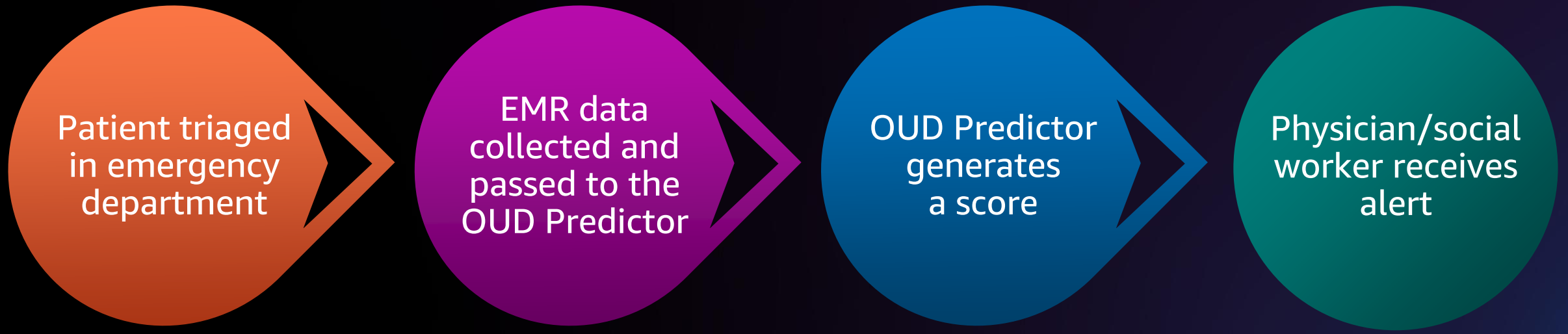
Built on Amazon SageMaker and other AWS services

Currently in validation

Planned general availability in 2022

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Identifying OUD risk in the emergency department



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Impact of machine learning

Advancing healthcare in months versus years, preventing harm, and improving the patient and clinician experience

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Pushing the frontier of what's possible

AMAZON PERSONALIZE



20M
Customers

2500+
Shows available

220
Countries & territories

50
Languages

Leveraged Amazon Personalize to deliver a better user experience, faster

3x increase in engagement with personalized recommendations

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GENERALLY AVAILABLE

Amazon Personalize Preconfigured Recommenders

Deliver ML-based recommendations regardless of ML expertise

NEW

"Top Picks," "Related Content," "Frequently Bought Together," and more

Works readily, no customization required

Optimized for media and retail use cases

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Pushing the frontier of what's possible

AMAZON DEVOPS GURU



Optimizes TV ads covering 21M US households

Uses Amazon DevOps Guru to detect anomalies in applications

Reduced application downtime

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Pushing the frontier of what's possible

AWS INDUSTRIAL AI SERVICES



Use Amazon Monitron and Amazon Lookout for Equipment for predictive maintenance

Detect equipment failures within minutes

Reduce production facility downtime

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Pushing the frontier of what's possible

AWS PANORAMA BRINGS MACHINE LEARNING TO THE EDGE



Use ML for tracking shipping containers

Deloitte built a solution with AWS Panorama

Edge computer vision using AWS Panorama identifies container movements

Expedited container inspection process

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ML INNOVATION DRIVER #2:

Process multi-modal data at scale

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Examples of multi-modal data for ML



1000's of labeled images



150 years of data in text documents



275 million minutes per year of audio and web interactions



Massive amounts of video and 3D point clouds

aws

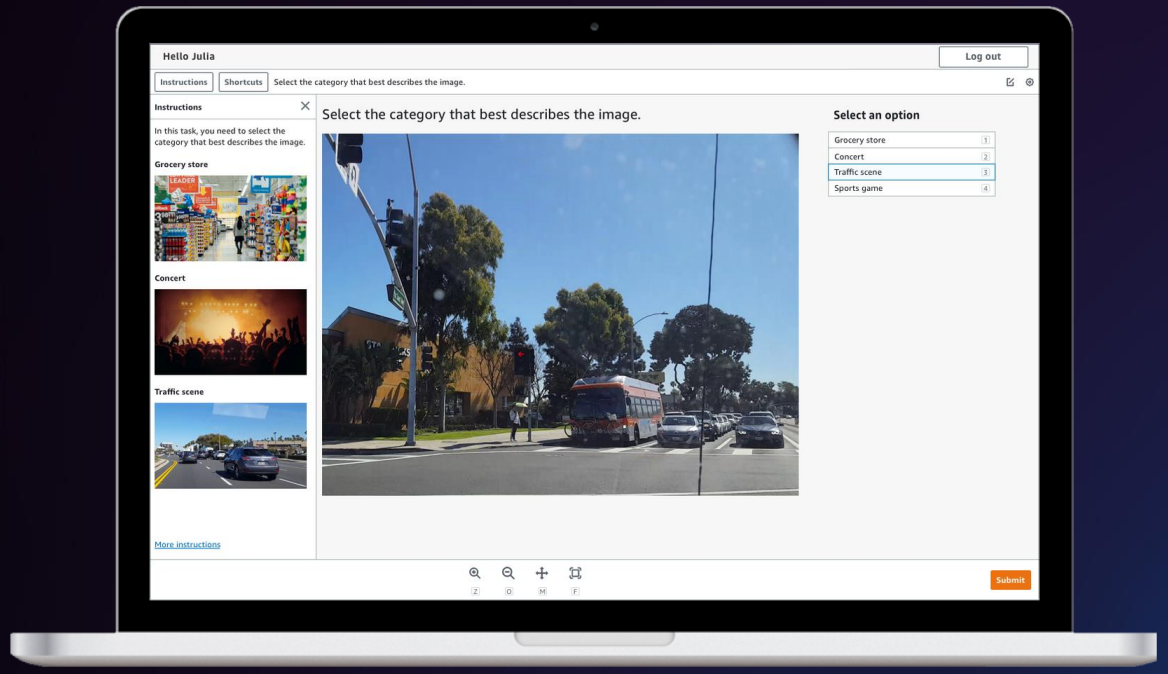
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GENERALLY AVAILABLE

Amazon SageMaker Ground Truth Plus

NEW

Deliver high-quality training datasets fast and reduce data labeling costs



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Benefits of Amazon SageMaker Ground Truth Plus



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ML INNOVATION DRIVER #3:

Empower more people to build ML solutions

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ML requires a lot more builders

+74%

annual growth in the past 4 years

2x

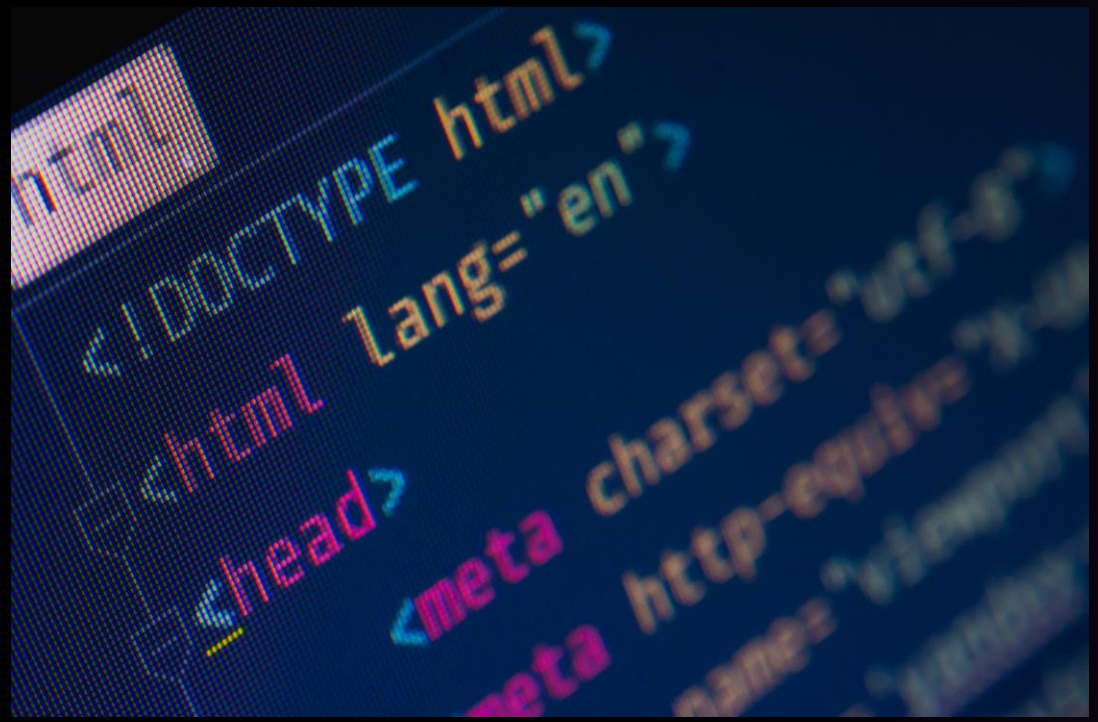
the growth of any other emerging job role

Source: LinkedIn Emerging Jobs Report 2020

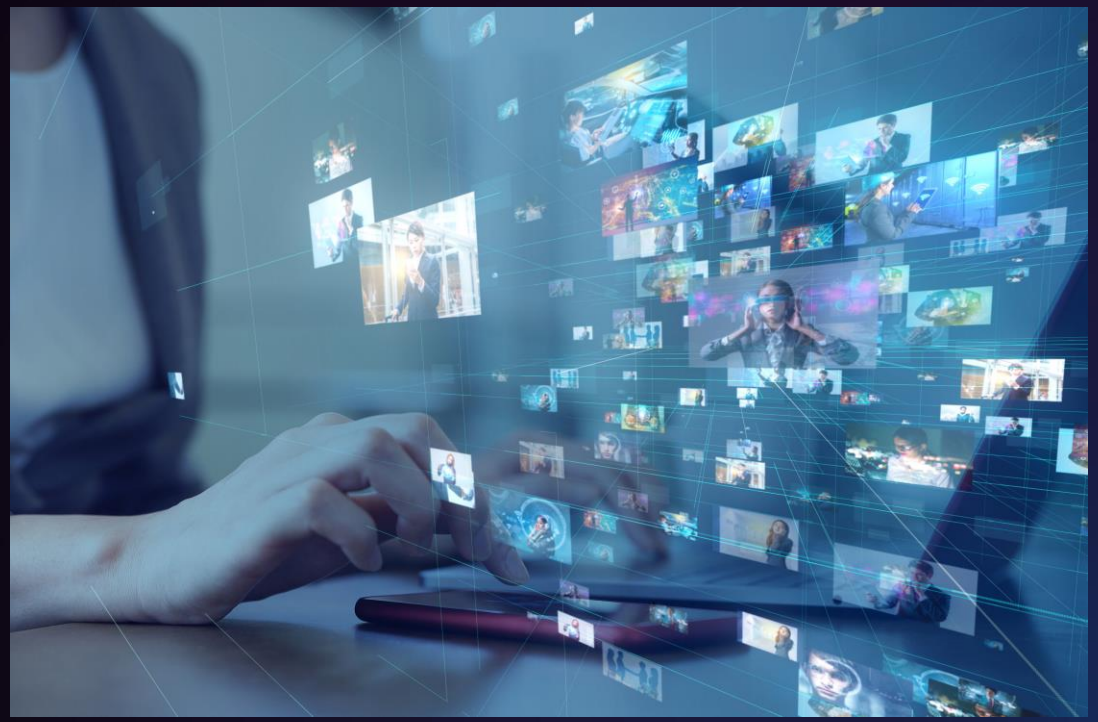
aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

No-code interface makes products more accessible



1990: early days of the Internet



2021: no-code web authoring tools

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GENERALLY AVAILABLE

Amazon SageMaker Canvas

Build ML models and generate accurate predictions—no code required

NEW

Quickly access and prepare data sources for ML

AutoML built-in to generate accurate predictions

Share ML models with data science teams

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon SageMaker Canvas

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

PREVIEW

Amazon SageMaker Studio Lab

Learn and experiment with ML using free, no-configuration notebooks in the cloud

NEW

Instantly access SageMaker notebooks and ML frameworks

All you need is an email ID

Easily migrate projects to SageMaker Studio for production

Automatically save session data

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ML INNOVATION DRIVER #4:

Industrialize ML to scale innovation

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Moving to mass creation of ML models



Industrialization



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

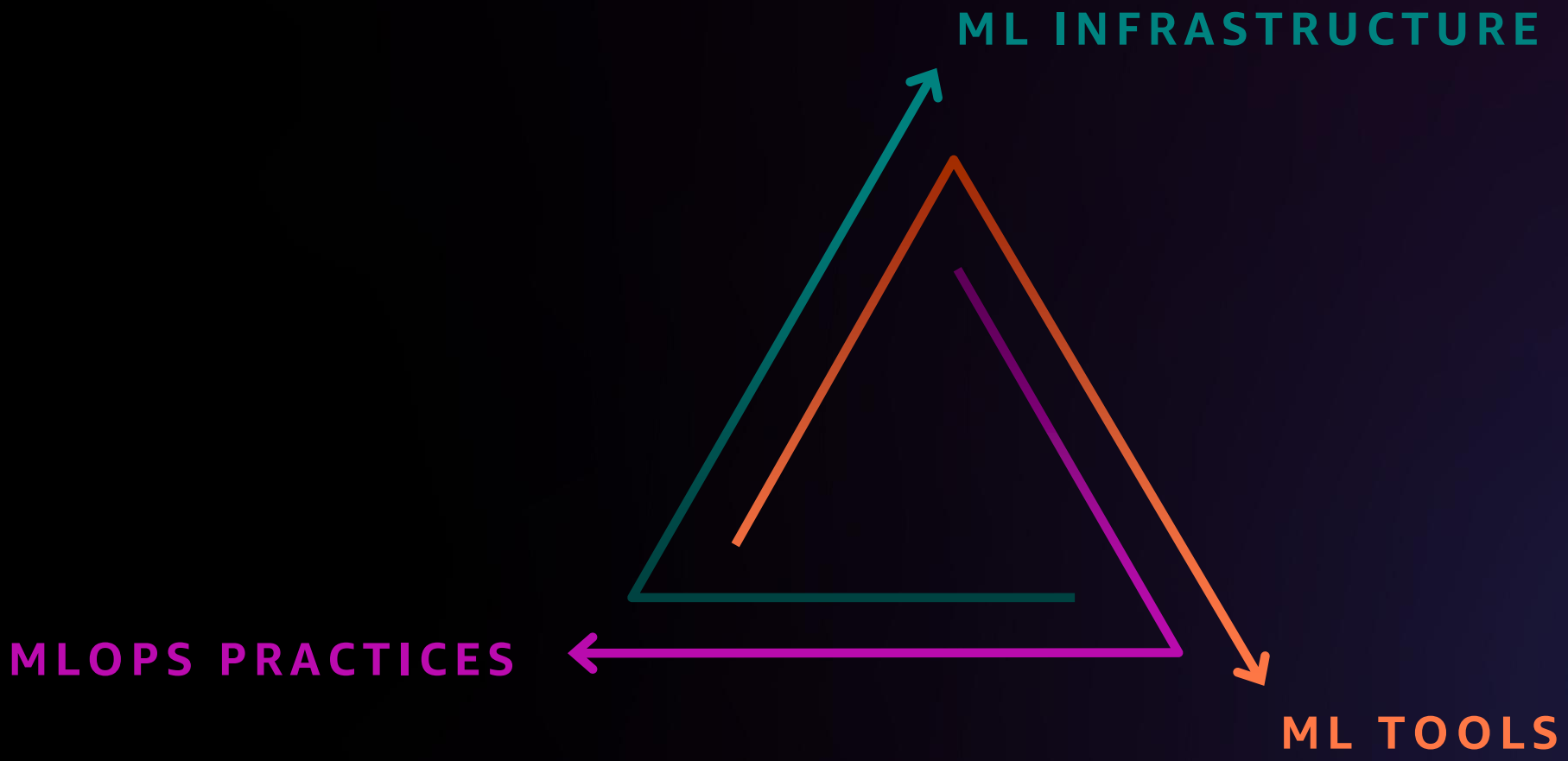
Massive scale

2021

DATA LABELING	1M+ labeling tasks/day
SINGLE CUSTOMER DEPLOYMENT	Million+ models
MODEL TRAINING	Billions of parameters
PREDICTIONS	100B+/month

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ML industrialization has three components



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Best price performance with ML infrastructure

TRAINING	Trainium	Best price performance for deep learning training in the cloud
	P4D	60% lower cost and 2.5x faster training compared to P3 and P3dn
	DL1	40% better price performance than GPU instances
INFERENCE	Inferentia	2.3x higher throughput, 70% lower cost than GPU instances
	G5	3x the performance of G4dn for inference

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

PREVIEW

Amazon SageMaker Serverless Inference

Deploy ML models for inference without worrying about servers

NEW

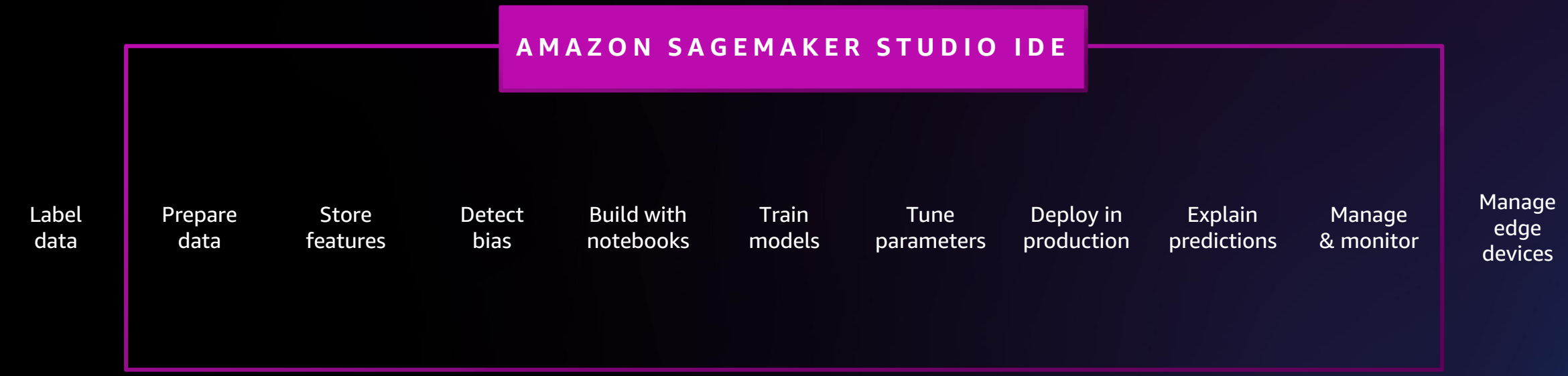
- Fully managed serverless environment
- Automatic scaling
- Pay per use pricing—no traffic means no charge
- Lower inference cost for intermittent usage pattern

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Integrated ML tools

AMAZON SAGEMAKER



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Improved productivity with Amazon SageMaker

Vanguard

Models deployed **20x faster**

AstraZeneca

Environments deployed in **5 minutes** vs. one month before

nerdwallet

Reduced training costs by **around 75%**

zendesk

Reduced inference costs by **90%**

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GENERALLY AVAILABLE

Amazon SageMaker Training Compiler

Accelerate training by up to 50% through more efficient use of GPUs

NEW

Integrated with versions of TensorFlow and PyTorch in SageMaker

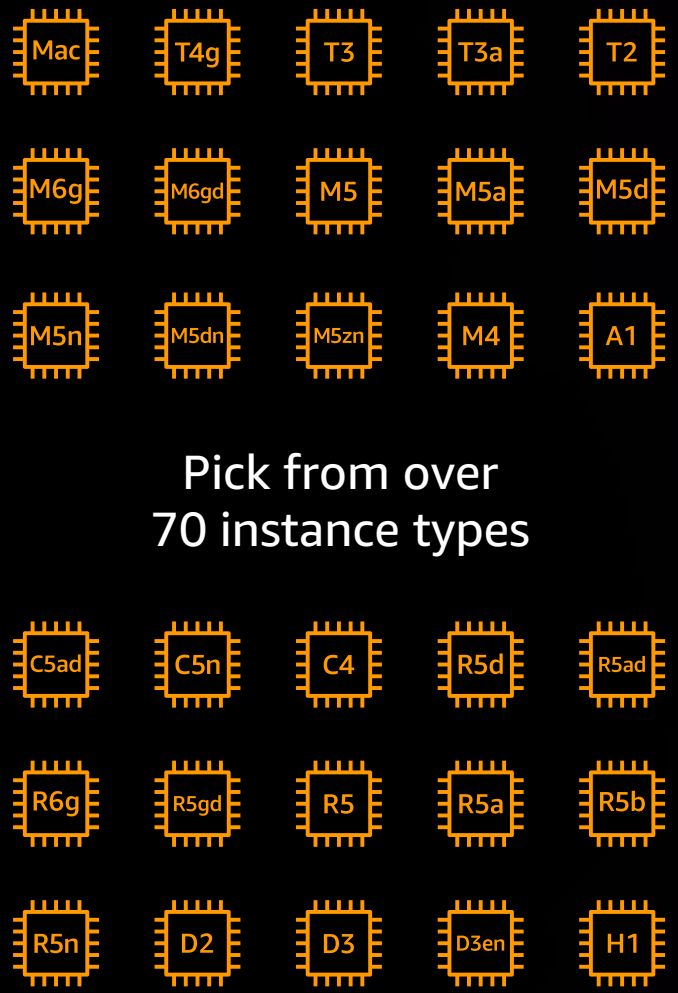
No change to existing workflows needed

Designed for deep learning model training

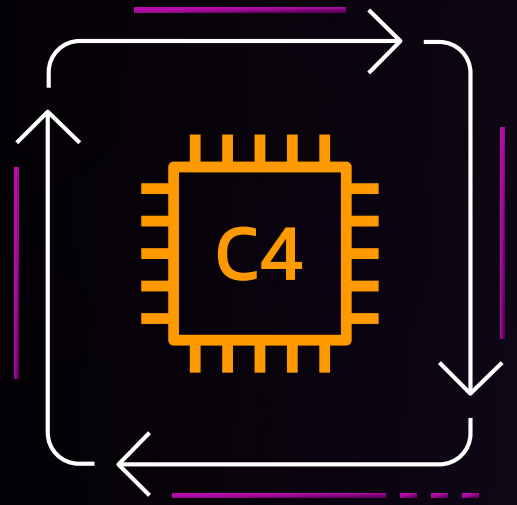
aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Model deployment takes a lot of work



Pick from over
70 instance types



Optimize model for the
selected instance type



Manage custom load
testing infrastructure

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GENERALLY AVAILABLE

Amazon SageMaker Inference Recommender

Reduce the time required to deploy ML models from weeks to hours

NEW

Automatically selects the best instance type, configuration, and model optimizations

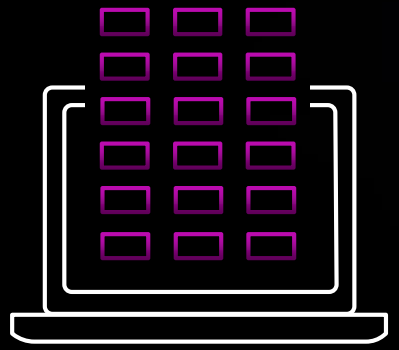
Fully managed load testing

Evaluate tradeoffs between latency, throughput, and cost for optimal deployment

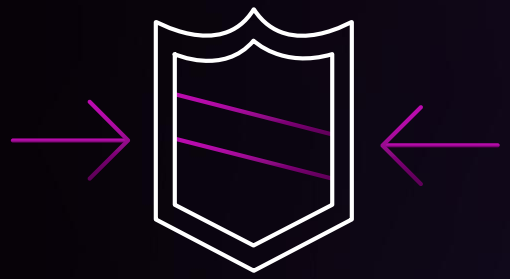
aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Streamline ML lifecycle with MLOps practices



Get more models to production faster



Enhance governance and security posture



Foster responsible use of ML

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MLOps practices

Streamline the ML lifecycle

Amazon SageMaker Pipelines

Automate ML workflows with CI/CD

Track lineage end-to-end

NEW Add bias detection to workflows

Explainable, high-quality predictions

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

One of the fastest growing AWS services

TENS OF THOUSANDS OF CUSTOMERS USE AMAZON SAGEMAKER



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The AWS ML stack

BROADEST AND MOST COMPLETE SET OF MACHINE LEARNING CAPABILITIES

AI SERVICES

SPECIALIZED

Business processes
Amazon Personalize
Amazon Forecast
Amazon Fraud Detector
Amazon Lookout for Metrics

Search
Amazon Kendra

Code + DevOps
Amazon CodeGuru
Amazon DevOps Guru

Industrial
Amazon Monitron
Amazon Lookout for Equipment
Amazon Lookout for Vision

Healthcare
Amazon HealthLake
Amazon Comprehend Medical
Amazon Transcribe Medical

CORE

Text/Documents
Amazon Textract
Amazon Comprehend
Amazon Translate

Chatbots
Amazon Lex

Speech
Amazon Polly
Amazon Transcribe
Amazon Transcribe Call Analytics

Vision
Amazon Rekognition
AWS Panorama

ML SERVICES

Label data

SAGEMAKER CANVAS
ML for business analysts

SAGEMAKER STUDIO LAB
Learn ML

AMAZON SAGEMAKER STUDIO IDE
Prepare data | Store features | Detect bias | Build with notebooks | Train models | Tune parameters | Deploy in production | Explain predictions | Manage & monitor

Manage edge devices

ML FRAMEWORKS & INFRASTRUCTURE

TensorFlow, PyTorch, Apache MXNet, Hugging Face | Deep learning AMIs & containers | Amazon EC2 | CPUs | GPUs | AWS Inferentia | AWS Trainium | Habana Gaudi | FPGA | Elastic Inference

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SUMMARY

Key drivers of ML innovation

Use ML to push the frontier of what's possible

Process multi-modal data at scale

Empower all your employees to build ML solutions

Industrialize ML to scale innovation

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Visit https://aws.amazon.com/machine-learning/

Check out AI/ML breakout sessions, workshops, and chalk talks

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.