# AWS re:Invent

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

FSI314

# Facilitating cross-border payments with machine learning

Gloria Vargas
GAM, Global Financial Services
AWS

Sanjay Dwivedi
Principal Prototype Architect, Global Financial Services
AWS

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

# Agenda

A personal story

The business problem

Why machine learning?

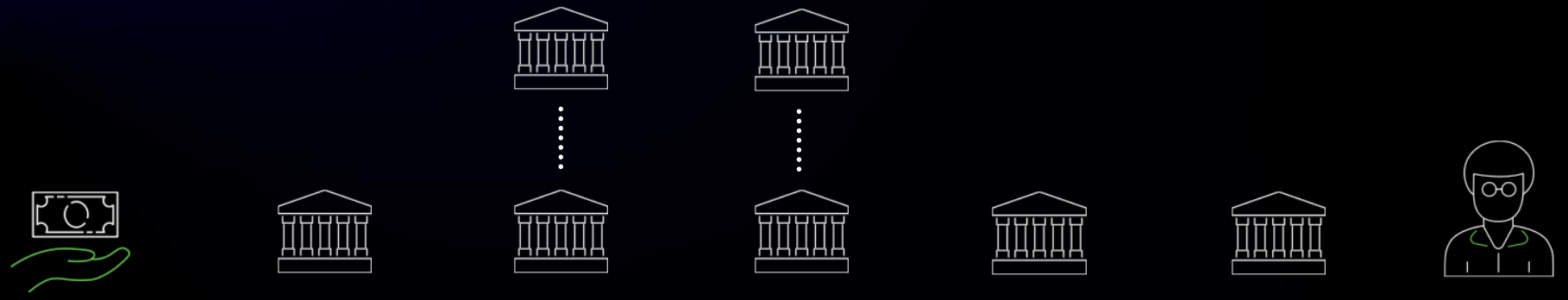
The solution

Call to action

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

# A personal story

CROSS-BORDER PAYMENTS TOUCH EVERYONE



Monthly wire back home

Schwab bank (USD)

Correspondent banks (USD): Citi, Wells Fargo, JPMC, etc.

Correspondent banks (COP): Bank of America, Wells Fargo, etc.

Domestic clearing (COP)

Banco Itau (COP)

Home: Colombia (COP)

**Payment:** Proprietary SWIFT format, ~70 data fields, unstructured beneficiary and address, varied market and bank practices

- ! Address was truncated
- ! Correspondent bank needs date of birth
- ! Recipient bank changed payments application, needs new ID field

5 out of 12 payments failed, 5 days average to solve issues

aws

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

# Payment data quality

A GLOBAL CHALLENGE MAGNIFIED BY THE ADOPTION OF NEW STANDARDS

**$6 trillion** In payment messages cross-border daily

**32 million** Cross-border payment messages/day over SWIFT

**12,000** Financial institutions and corporations globally, different adoption time lines

**200+** Countries and territories, diverse market practices

**7x** More data points in ISO20022 than existing, proprietary formats

87% of global payments on ISO20022 by 2025




Source: Deutsche Bank

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

# Change in payments is accelerating globally



**Richer data** will enable customized and predictive **customer experiences** enabled by the increase in machine-readability of payments data



**Mandatory** global moves to **ISO20022** will bring richer data to payments



**Increased volume:** digital payments are increasing about 8–10% per annum



Payments modernization



Need to keep up with the pace of **innovation** set by **new entrants**



**Time to market** in new segments and new regions



Consumers and corporations expect **instant payments** experiences

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

**“It is probably the most impactful payments industry undertaking since the introduction of SEPA more than a decade ago”**

Christian Westerhaus

Head of Cash Products, Cash Management, Global Transaction Banking
Deutsche Bank

aws

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

# Why machine learning?

Payment data quality issues prediction is not a problem solvable with rules

| 5% | 200+ | 10,000+ | 5 |
| --- | --- | --- | --- |
| Of all payments fail due to data quality issues | Possible market practices; change is not coordinated or standardized | Rules a global bank may implement to cover all possible market and bank practices of its correspondent network without being exhaustive | Years the global transition to ISO20022 will take, with each bank defining its own time line |

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

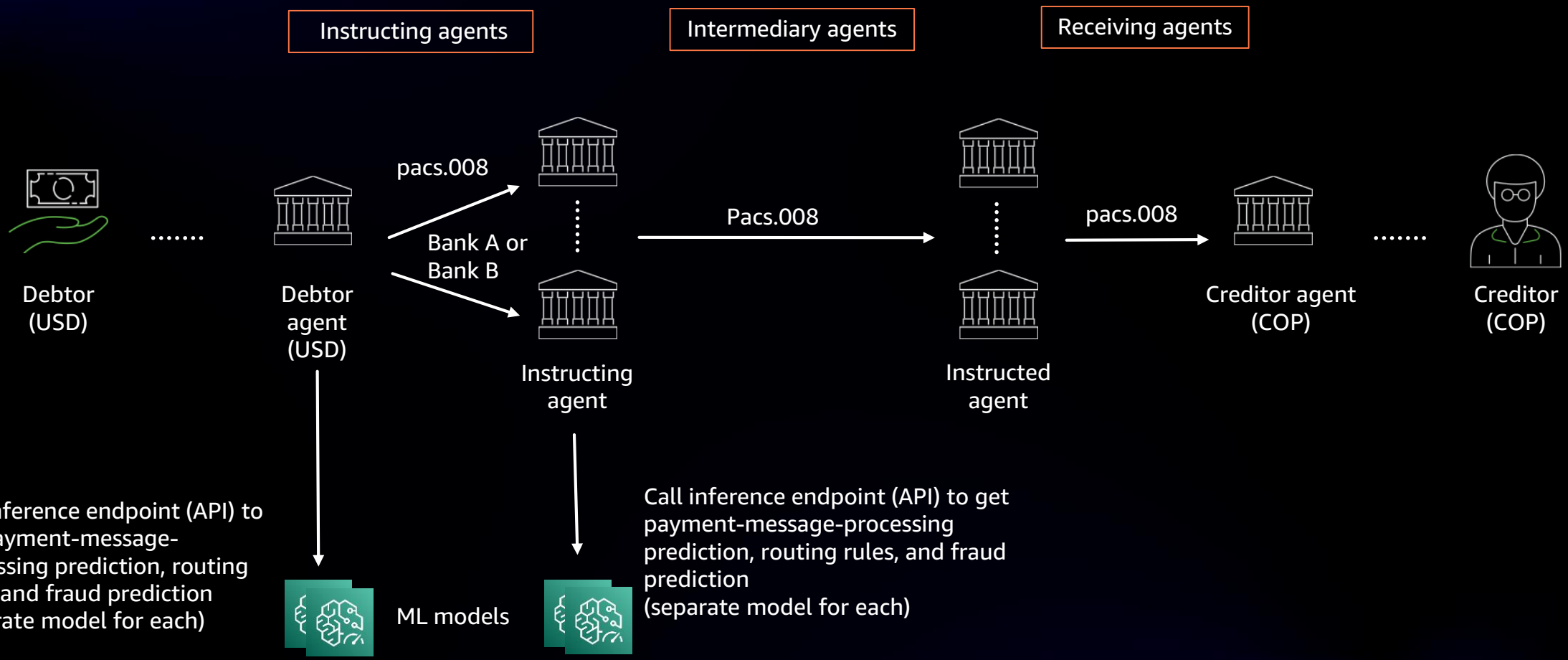
# Machine learning use cases in payments

**Machine learning can enable efficiency in cross-border payment transactions**

- Use ML models to reduce payment delays by predicting payment outcome earlier in the payment life cycle and correcting errors
- Use ML models to perform intelligent routing based on a corresponding bank's timeliness performance, cost, or some other business metric
- Use ML models to predict cross-border payment fraud and take proactive actions
- Use ML models to improve regulatory compliance and reporting

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

# ML inference in real-time payment flow




Sender-side payment systems can use their custom models in real-time message flow to gain insight on payment-processing outcome (success or failure with probability), select a correspondent agent to send payment to with best performance, or get fraud prediction. Any payment system in the flow can do the same to improve the customer experience

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

# Conceptual prototype

- AWS, in collaboration with the industry, has created and open sourced a simple ISO20022 CBPR+ message generator tool to generate synthetic data
  - Currently supports one message format (pacs.008) generation
  - Can be enhanced and customized with industry participation and contribution

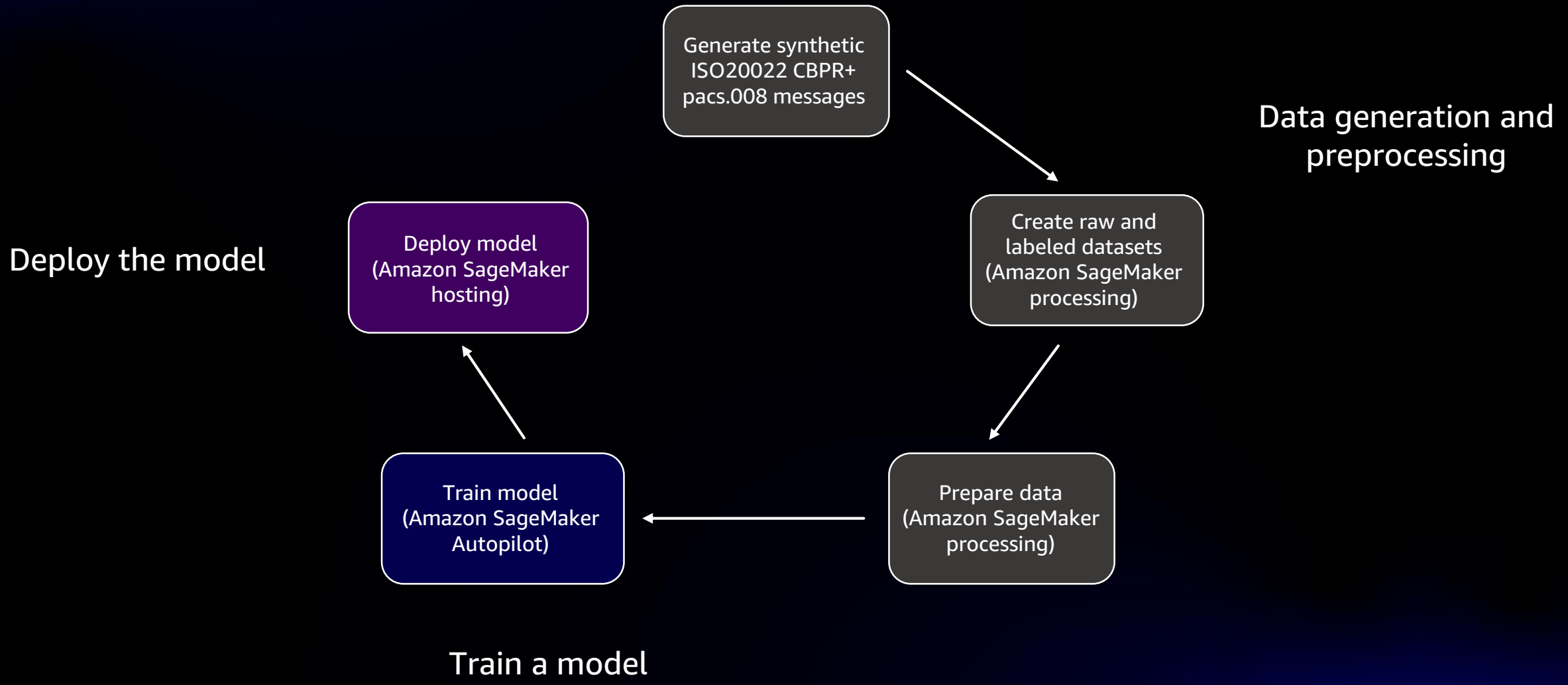
- Art of the possible: AWS used this synthetic data to train an ML model to predict whether a payment message will be accepted or rejected by a specific chain of banks
  - Data is completely synthetic and doesn't represent real-world payment transactions
  - But it is used to show the art of the possible



aws

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

# Prototype development approach




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

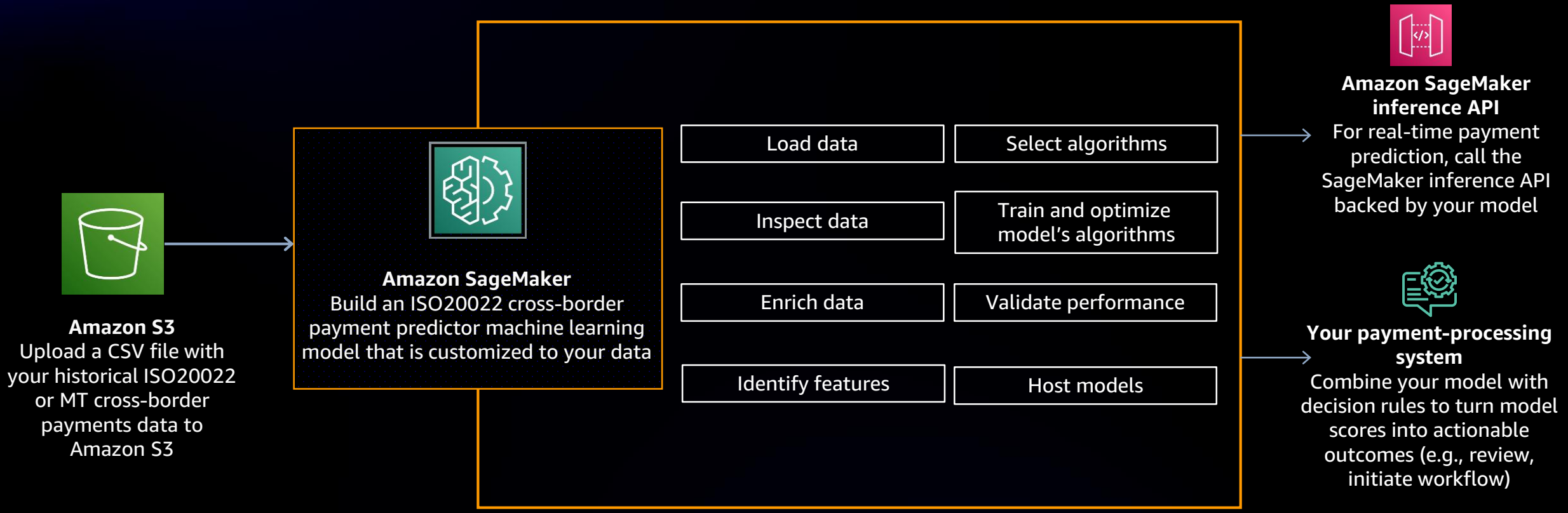
# Conceptual prototype

- The ISO20022 message generator is written in Java, Spring Boot
  - It supports the LEI database published by Global Legal Entity Foundation (GLEIF)
  - It supports BIC database
  - Only a facility for the database is provided; you bring your own data
  - The database can be hosted in memory or in a relational database such as Amazon RDS
- The ML model was developed using SageMaker
  - Python notebooks
  - Uses SageMaker processing and training
  - Uses SageMaker Autopilot to train multiple models and select best-performing model
  - SageMaker hosting services for real-time inference

aws

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

# The solution: Future




Art of the possible: use machine learning models to achieve efficiency in your payment processing flows

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

# Future: Shared datasets

Collaboration to create and share labeled real-world payments dataset, appropriately cleaned and anonymized, will be crucial for the success of ML-based solutions

- Truly effective models need to be trained with multiple entities' labeled payments data for identified business use cases
- Shared models providing predictions via an API need to be used at payment creation by the end customer
- Industry participants can use shared industry datasets, enriched with their own dataset, to create custom models for their specific use cases or market practices

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

# Future

- Use the power of AI and machine learning to make your payment-processing ecosystem efficient and intelligent
- Combine your own payments data with industry-contributed and curated real cross-border payments data to create models to fit your needs
- Use industry-contributed ML models for specific use cases as a starting point, and modify the inference layer to create your own models for specific use cases that are not addressed by the industry-contributed models

**Industry-wide synthetic data pipelines to train shared models could save billions of dollars in payment-failure investigation and repair fees**

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

# Call to action

- Leverage the synthetic data generator tool; feedback welcome
- Join industry working groups
- Leverage the prototype AI/ML model approach to develop your model with your data for payment data quality prediction; feedback welcome
- Schedule a data lab or ML lab engagement with us

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

# Thank you!

Gloria Vargas
vargasgl@amazon.com

Sanjay Dwivedi
sanjaydw@amazon.com



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

# Public dataset examples

- Computer vision datasets
  - MNIST – handwritten images for digits (0–9), Fashion-MNIST (clothing images, 28x28, grayscale, 60k images)
  - CIFAR-10 – a tiny image dataset (32x32 color images), 10 classes, 60 thousand images per class
  - ImageNet – 22,000 distinct categories with more than 14 million images
  - Places dataset – More than 10 million images with 400+ unique scene categories
- UCI Machine Learning Repository
  - Bank Marketing Data Set
  - Credit Card Default Data Set
- Registry of Open Data on AWS
  - The Cancer Genome Atlas
  - Amazon Sustainability Data Initiative

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

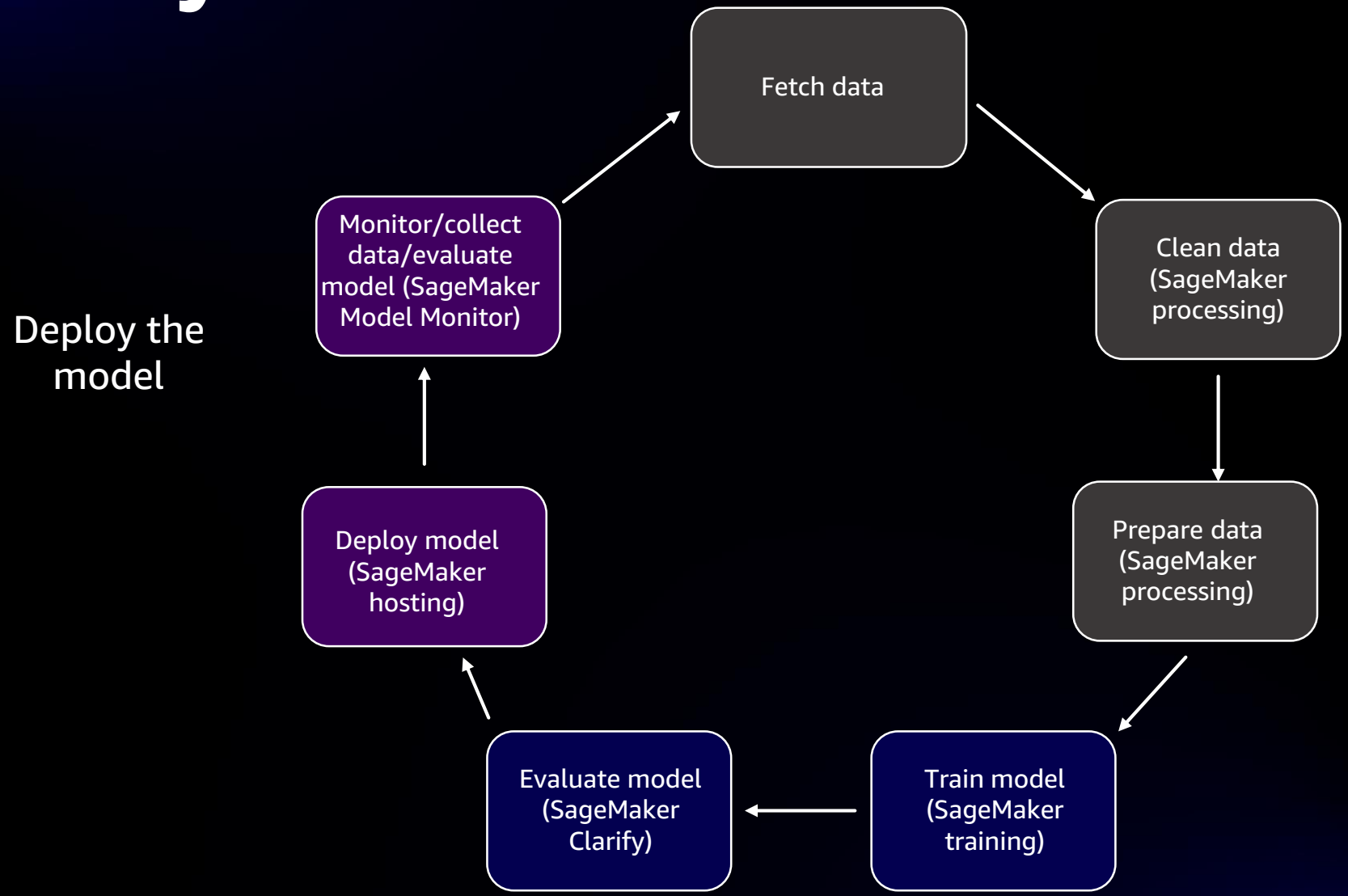
# ML model examples

- Machine learning architectures and pretrained models
  - Computer vision – CNN for image classification
    - LeNet (1998)
    - AlexNet (2012)
    - VGGNet (2014)
    - ResNet (2015)
  - Natural language processing
    - Pretrained BERT variations (this is costly; usually needs large datasets and big compute power to successfully train BERT model
      - Available in Hugging Face's Transformers library
    - 3 generations of OpenAI Generative Pre-trained Transformers (GPT) models (GPT-1, GPT-2 and GPT-3)

aws

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

# ML life cycle




© 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

### Health AI
- Amazon HealthLake
- Amazon Transcribe Medical
- Amazon Comprehend Medical

### Industrial AI
- AWS Panorama + appliance
- Amazon Monitron
- Amazon Lookout for Equipment
- Amazon Lookout for Vision

### Anomaly detection
- Amazon Lookout for Metrics

### Code and DevOps
- Amazon DevOps Guru
- Amazon CodeGuru

### Vision
- Amazon Rekognition

### Speech
- Amazon Polly

### Text
- Amazon Transcribe Medical
- Amazon Comprehend Medical
- Amazon Translate
- Amazon Textract

### Search
- Amazon Kendra

### Chatbots
- Amazon Lex

### Personalization
- Amazon Personalize

### Forecasting
- Amazon Forecast

### Fraud
- Amazon Fraud Detector

### Contact centers
- Contact Lens for Amazon Connect
- Voice ID

## ML services

- Amazon SageMaker
- Label data

**SageMaker Studio IDE**

Aggregate & prepare data | Store & share features | Auto ML | Spark/R | Detect bias | Visualize in notebooks | Pick algorithm | Train models | Tune parameters | Debug & profile | Deploy in production | Manage & monitor | CI/CD

- Human review

SageMaker JumpStart

Model management for edge devices

## Frameworks and infrastructure

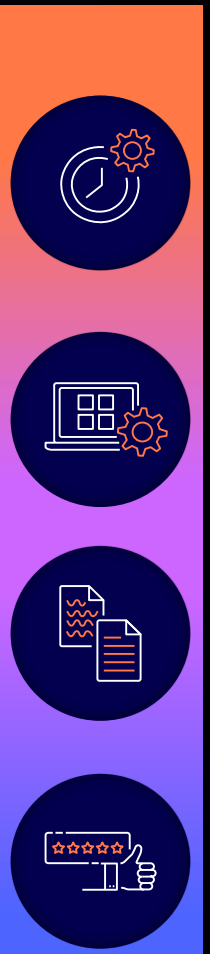
- TensorFlow
- mxnet
- PyTorch
- GLUON
- Keras
- scikit-learn
- Horovod
- DeepGraphLibrary
- AWS Deep Learning AMIs (DLAMI) and AWS Deep Learning Containers
- GPUs and CPUs
- Amazon Elastic Inference
- AWS Trainium
- AWS Inferentia
- FPGA

# Amazon SageMaker Autopilot

Automatic model creation with full visibility and control

Supports:

- Classification, binary and multiclass
- Regression



**Quick to start**
Provide your data in a tabular form, and specify target prediction

**Automatic model creation**
Get ML models with feature engineering and model tuning automatically done

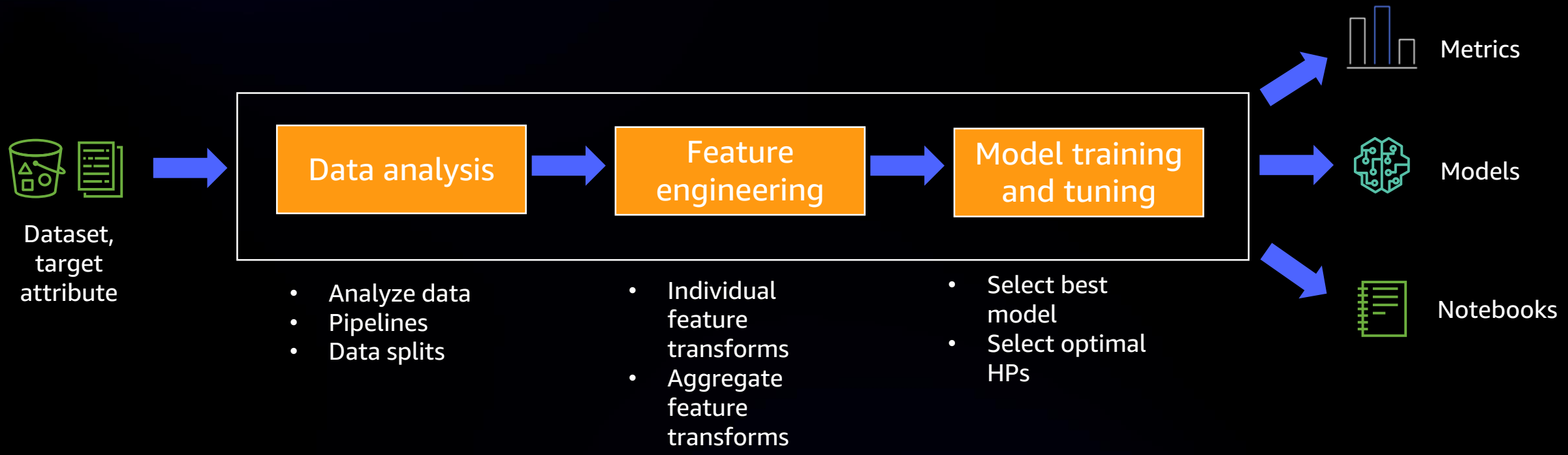
**Visibility and control**
Get notebooks for your models with source code

**Recommendations and optimization**
Get a leaderboard, and continue to improve your model

aws

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

# How Amazon SageMaker Autopilot works




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

# Amazon SageMaker Pipelines



## Accelerate machine learning development

With just a few lines of code, you can create a fully automated ML workflow, reducing months of coding to a few hours



## Automatically track model artifacts

Eliminate manual processes by automatically tracking model artifacts to keep a structured audit trail



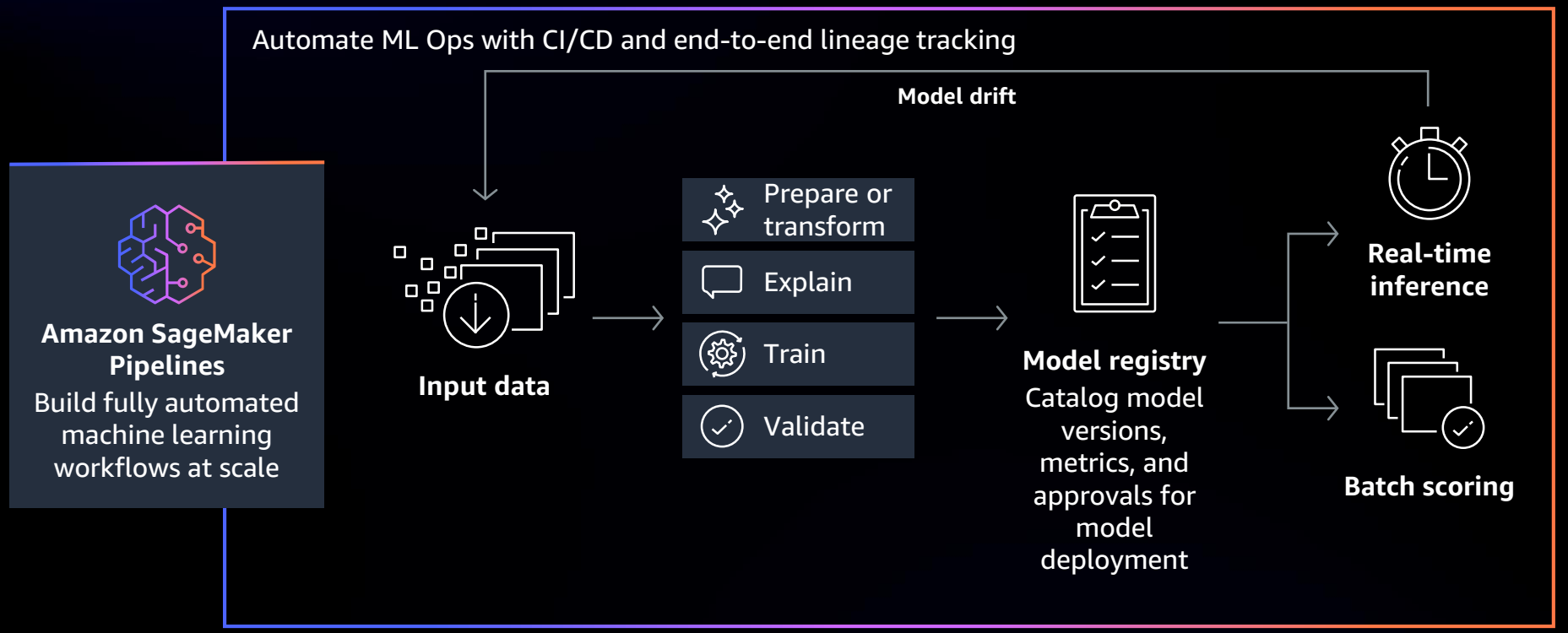
## Scale the number of ML models in production

Choose from built-in templates to set up CI/CD pipelines to automate workflows and deploy ML models at scale

aws

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

# Amazon SageMaker Pipelines

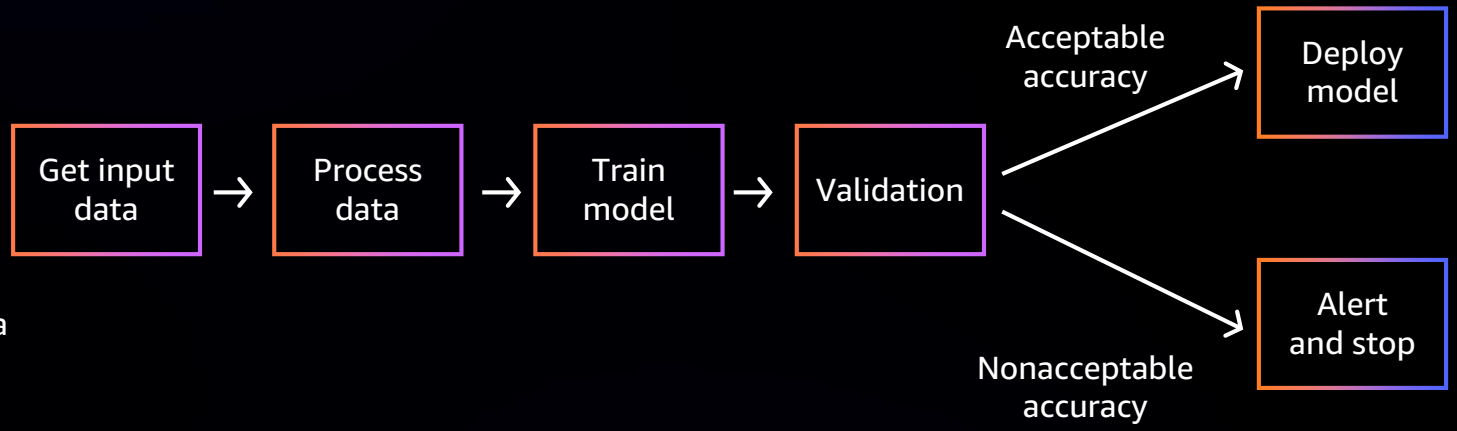



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

# How Amazon SageMaker Pipelines works

**Start pipeline execution:**

- Manually
- New data uploads via Amazon CloudWatch Events
- Upon code check-in




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