AWS re:Invent

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

AIM320

Implementing MLOps practices with Amazon SageMaker

Shelbee Eigenbrode
Principal AI/ML Specialist Solutions Architect
AWS

Ritesh Shah
Head of Architecture for Data, Analytics, and Machine Learning
Vanguard

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

Agenda

300 Level



- MLOps: Setting the stage
- SageMaker MLOps-ready capabilities
- Vanguard MLOps journey
- Wrap-up

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

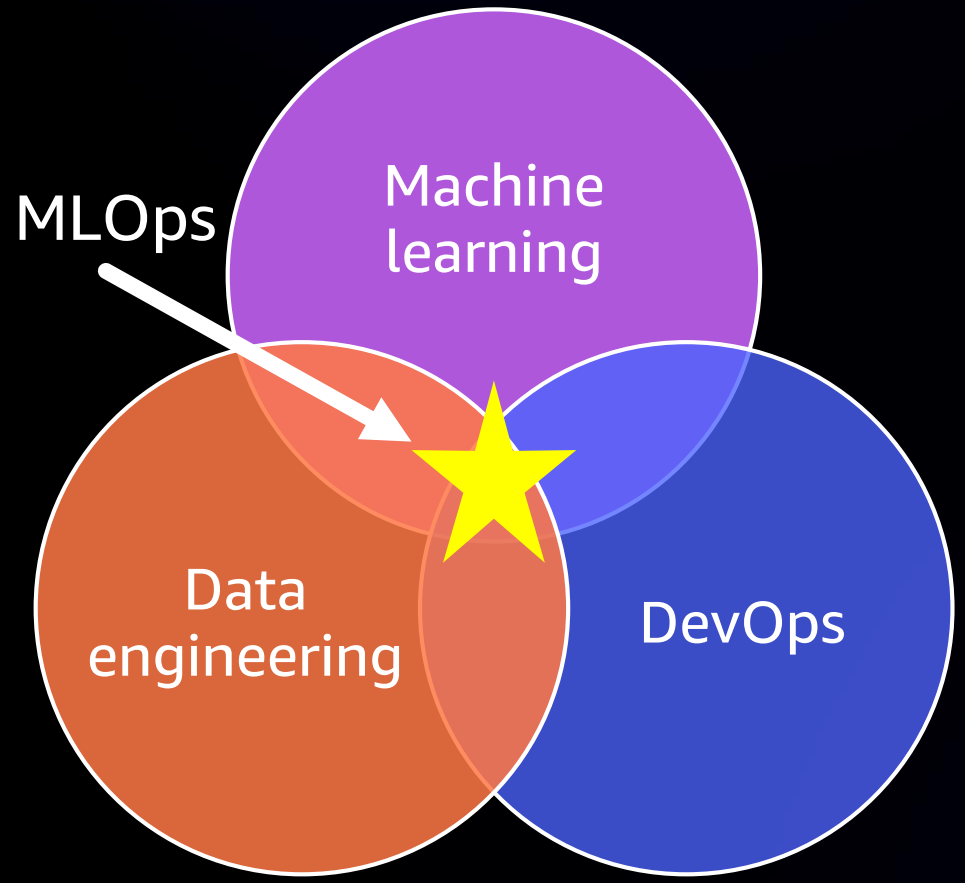
MLOps: Setting the stage

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

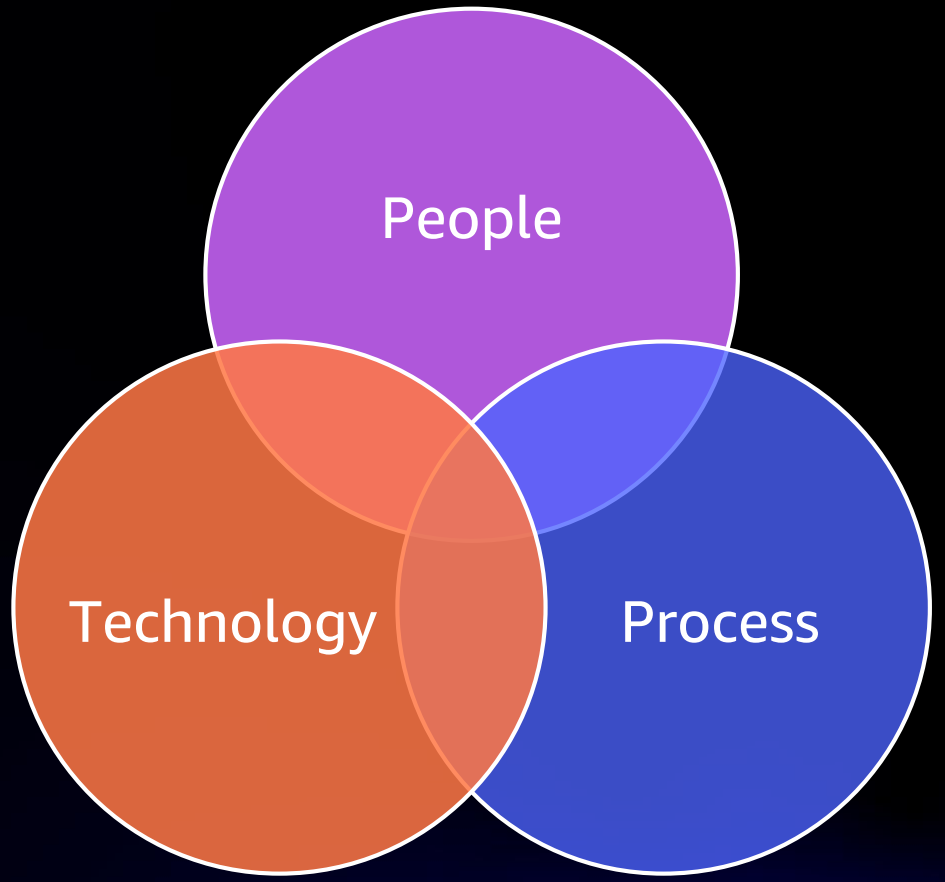
MLOps: High-level view

INTERSECTION OF DOMAINS WITH ENABLING PRACTICES

Domains



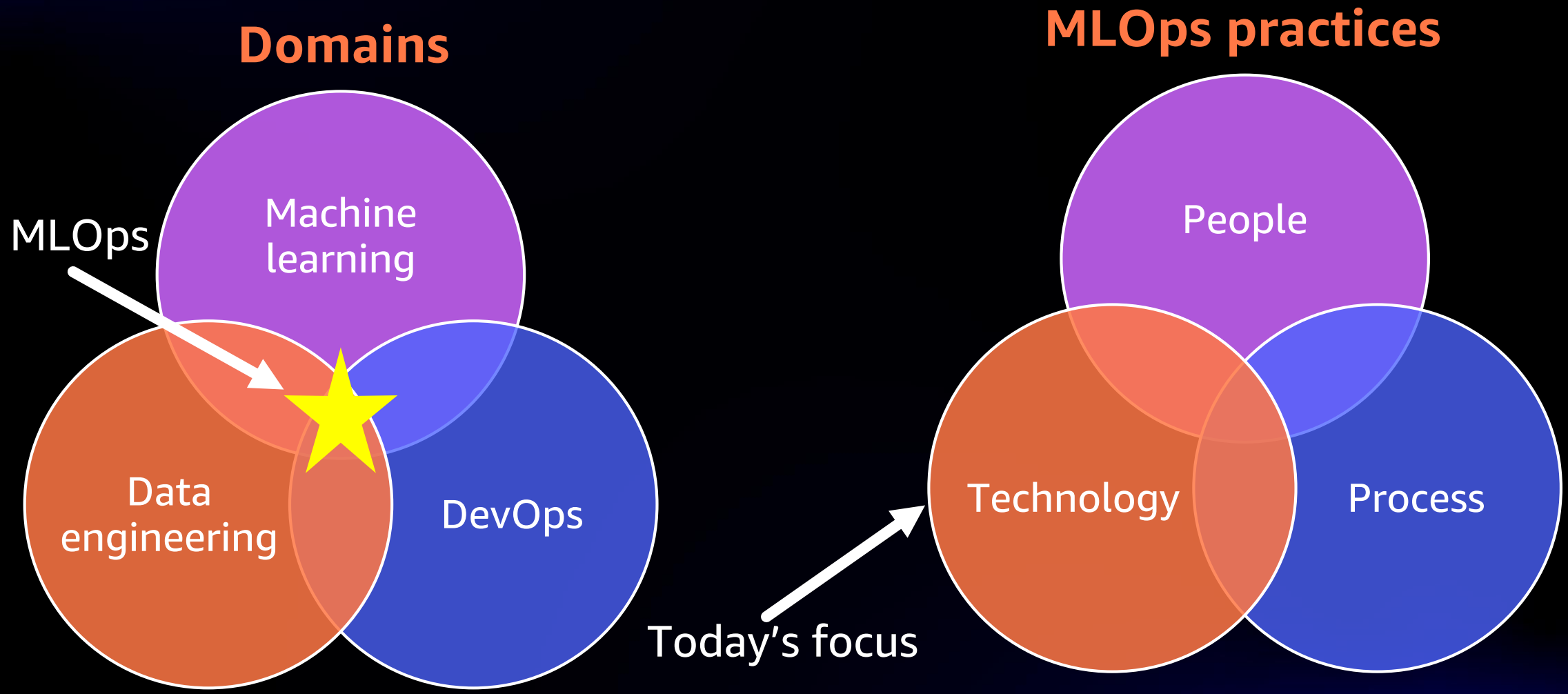
MLOps practices



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

MLOps: High-level view

INTERSECTION OF DOMAINS WITH ENABLING PRACTICES



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

MLOps: The golden pipeline



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

MLOps practices: Tool selection

AUTOMATED WORKFLOWS & CI/CD PIPELINES

SageMaker native

- SageMaker Pipelines
- SageMaker projects
- SageMaker Model Registry

Do-it-yourself

- AWS Developer Tools (AWS CodeCommit, AWS CodePipeline, AWS CodeBuild)
- Amazon ECR
- Amazon Managed Workflows for Apache Airflow (Amazon MWAA)
- AWS Step Functions

Third-party examples

- GitHub, Bitbucket
- Jenkins
- Artifactory
- Kubeflow (SageMaker Components for Kubeflow Pipelines)
- Apache Airflow

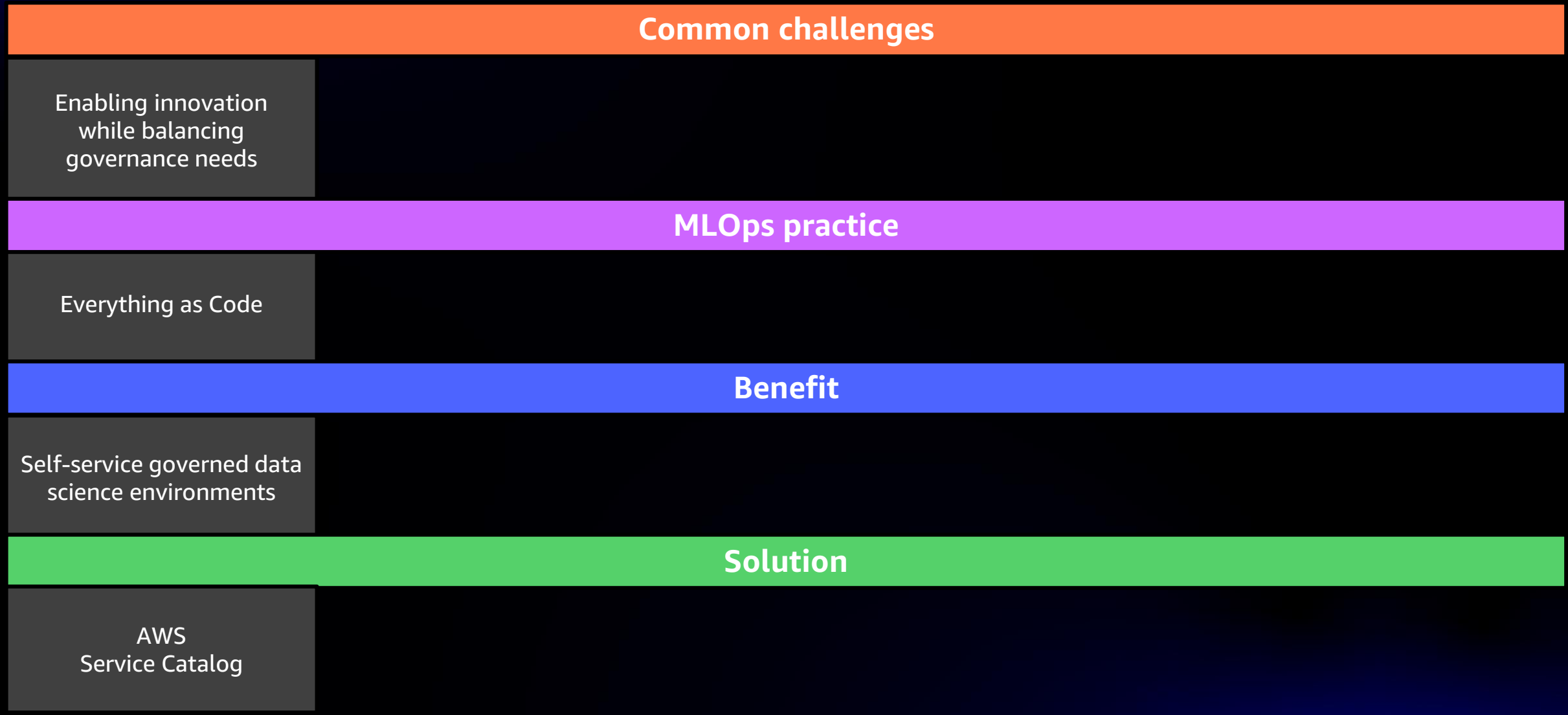
aws

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

SageMaker MLOps-Ready Capabilities

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

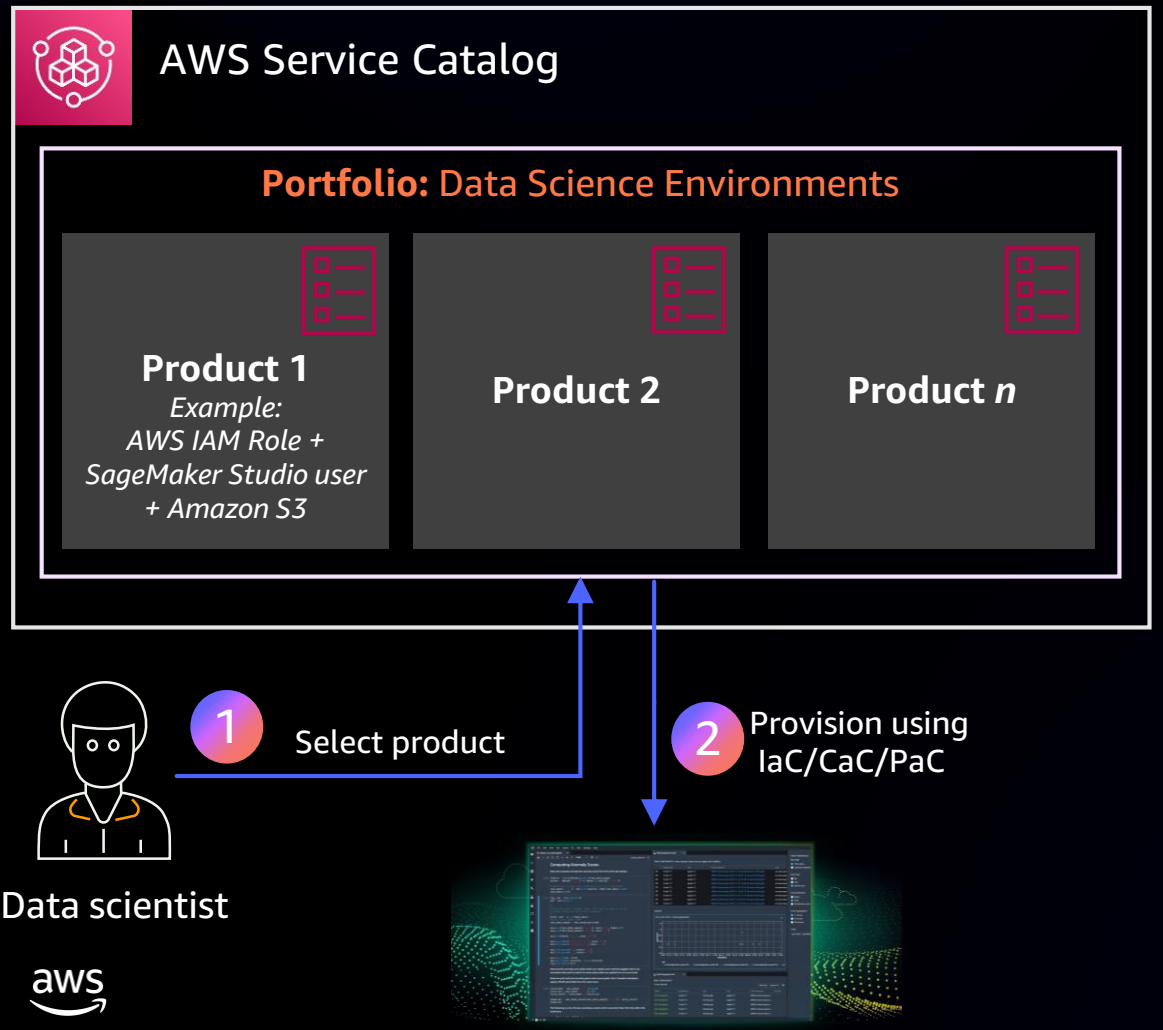
MLOps: Core practices and benefits



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

Data science environments

Challenge: Enabling innovation while balancing governance needs



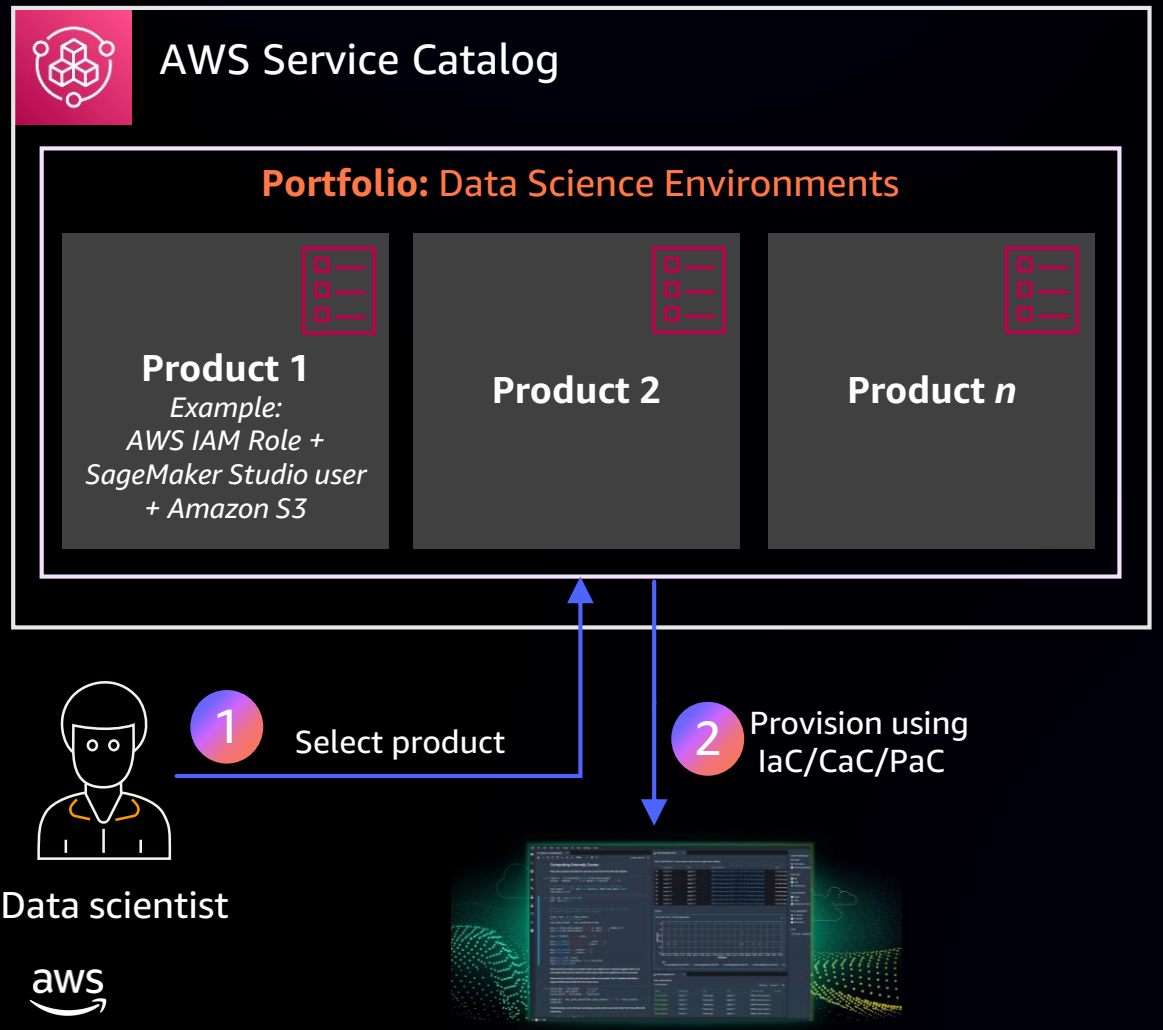
Solution

Utilize AWS Service Catalog to provision environments using infrastructure, configuration, or policy as code

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

Data science environments

Challenge: Enabling innovation while balancing governance needs



Solution

Utilize AWS Service Catalog to provision environments using infrastructure, configuration, or policy as code

Benefits:

- Ensure repeatability and consistency
- Standardize data science environments
- Implement guardrails via code

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

MLOps: Core practices and benefits

Common challenges	
Enabling innovation while balancing governance needs	Bridging the gap between model build and model deploy tasks
MLOps practice	
Everything as Code	Lineage tracking and configuration management
Benefit	
Self-service governed data science environments	Repeatable process to create workflows across personas
Solution	
AWS Service Catalog	Amazon SageMaker Model Registry

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

MLOps: Core practices and benefits

Common challenges		
Enabling innovation while balancing governance needs	Bridging the gap between model build and model deploy tasks	Providing end-to-end traceability
MLOps practice		
Everything as Code	Lineage tracking and configuration management	CI/CD
Benefit		
Self-service governed data science environments	Repeatable process to create workflows across personas	Auditable ML pipeline
Solution		
AWS Service Catalog	Amazon SageMaker Model Registry	Amazon SageMaker projects

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

MLOps: Core practices and benefits

Common challenges			
Enabling innovation while balancing governance needs	Bridging the gap between model build and model deploy tasks	Providing end-to-end traceability	Building a repeatable and consistent process for deploying models
MLOps practice			
Everything as Code	Lineage tracking and configuration management	CI/CD	Continuous delivery and deployment
Benefit			
Self-service governed data science environments	Repeatable process to create workflows across personas	Auditable ML pipeline	Improve time-to-market
Solution			
AWS Service Catalog	Amazon SageMaker Model Registry	Amazon SageMaker projects	Amazon SageMaker projects and Pipelines

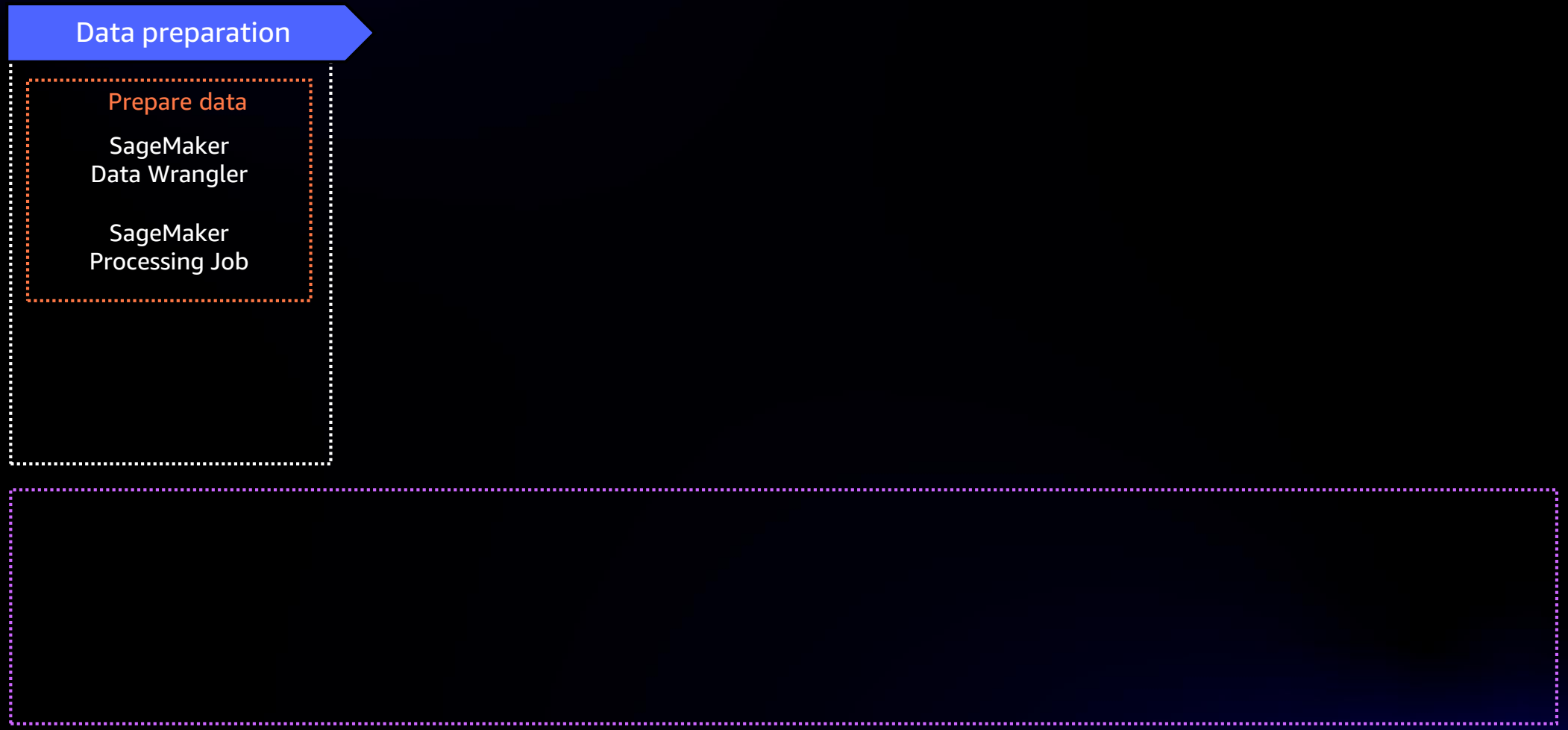
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MLOps: Core practices and benefits

Common challenges				
Enabling innovation while balancing governance needs	Bridging the gap between model build and model deploy tasks	Providing end-to-end traceability	Building a repeatable and consistent process for deploying models	Monitoring models
MLOps practice				
Everything as Code	Lineage tracking and configuration management	CI/CD	Continuous delivery and deployment	Continuous monitoring
Benefit				
Self-service governed data science environments	Repeatable process to create workflows across personas	Auditable ML pipeline	Improve time-to-market	Maintain model performance over time
Solution				
AWS Service Catalog	Amazon SageMaker Model Registry	Amazon SageMaker projects	Amazon SageMaker projects and Pipelines	Amazon SageMaker Model Monitor

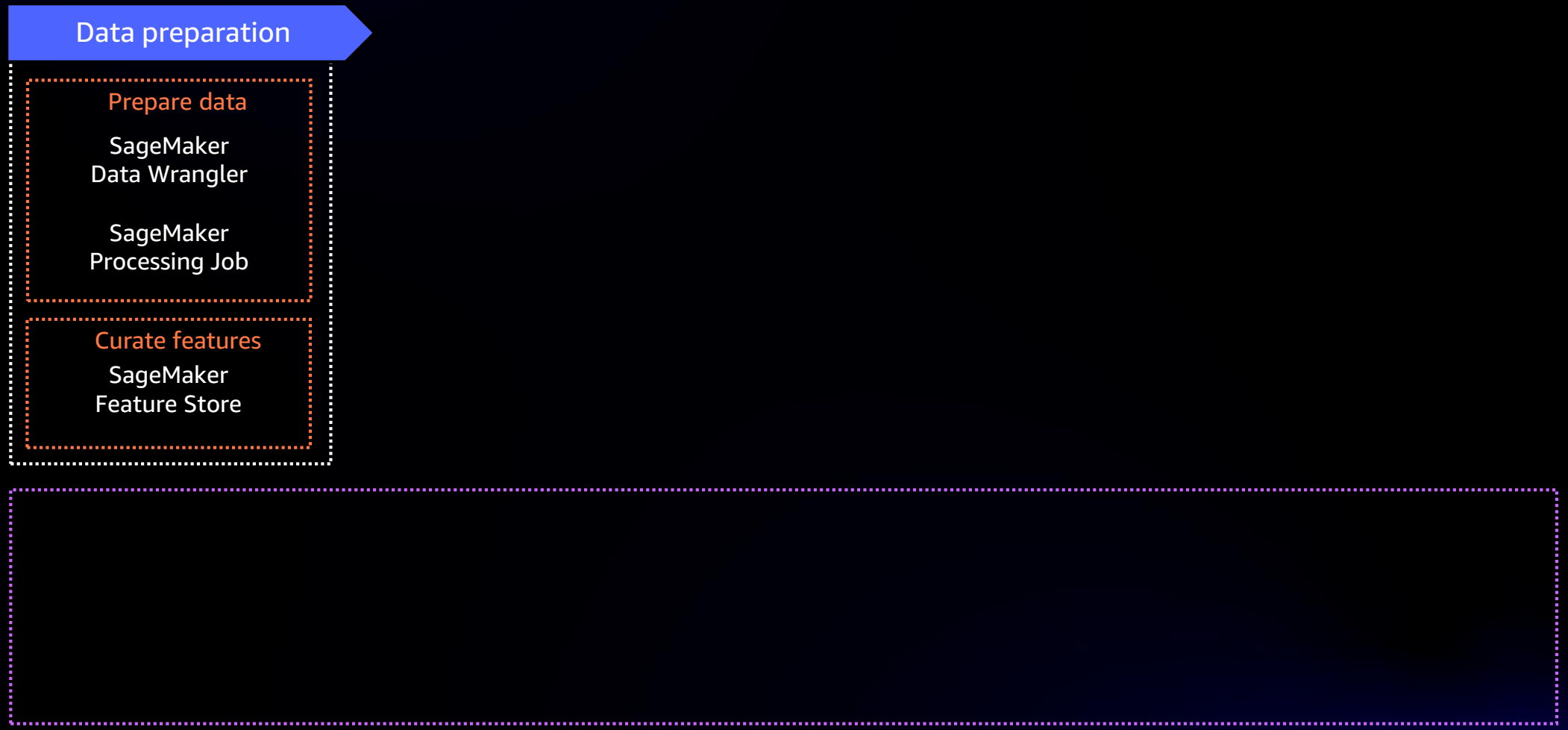
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MLOps-ready features and capabilities



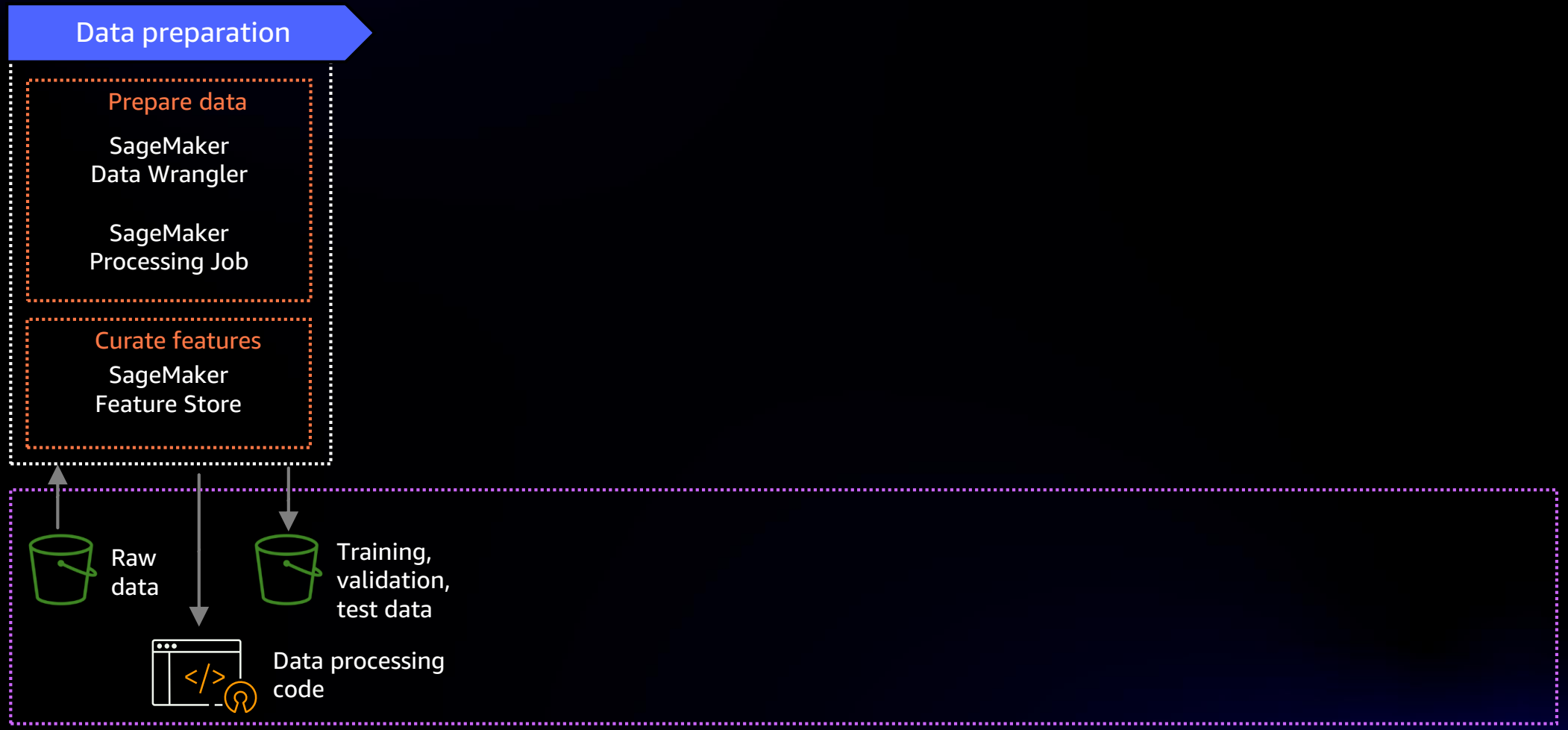
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MLOps-ready features and capabilities



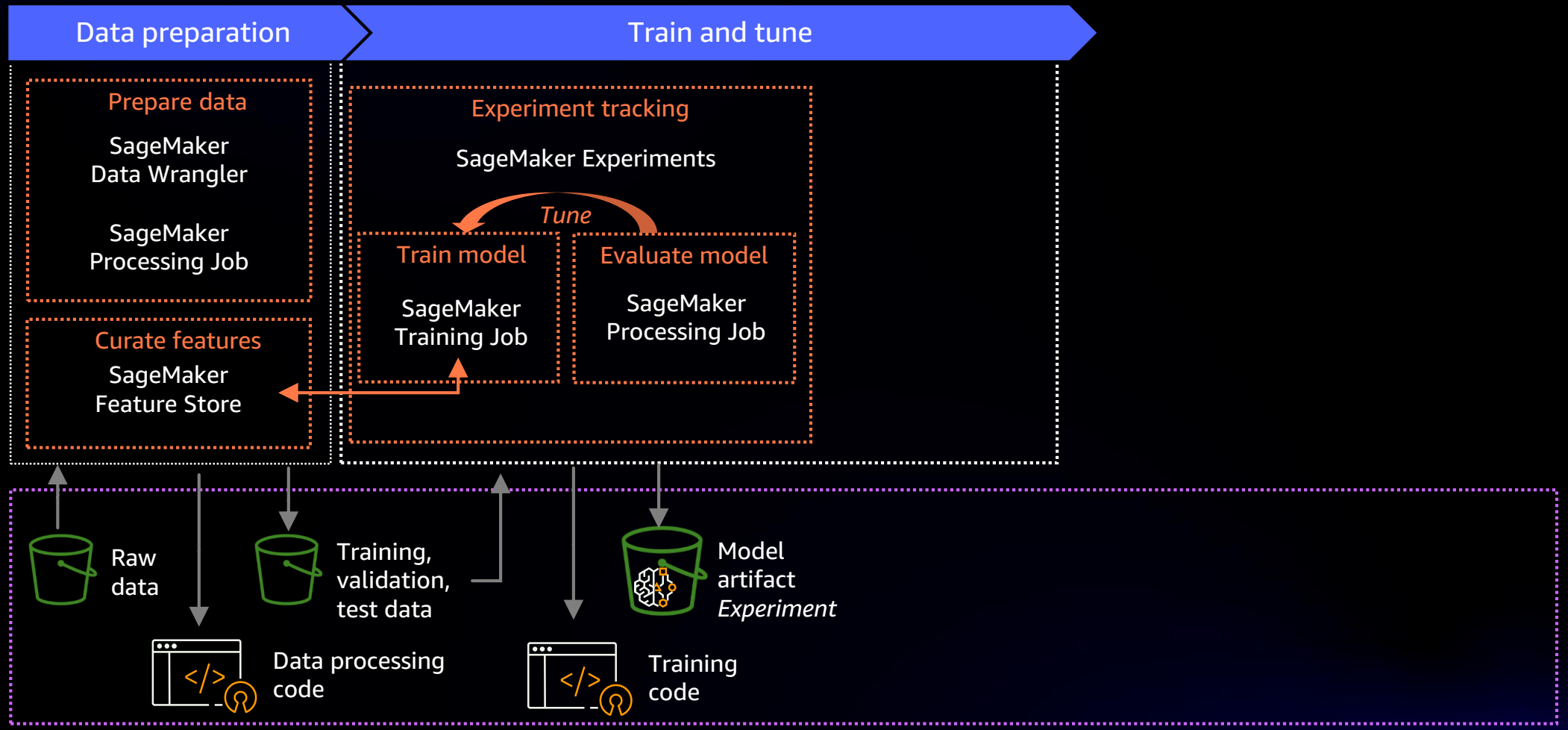
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MLOps-ready features and capabilities



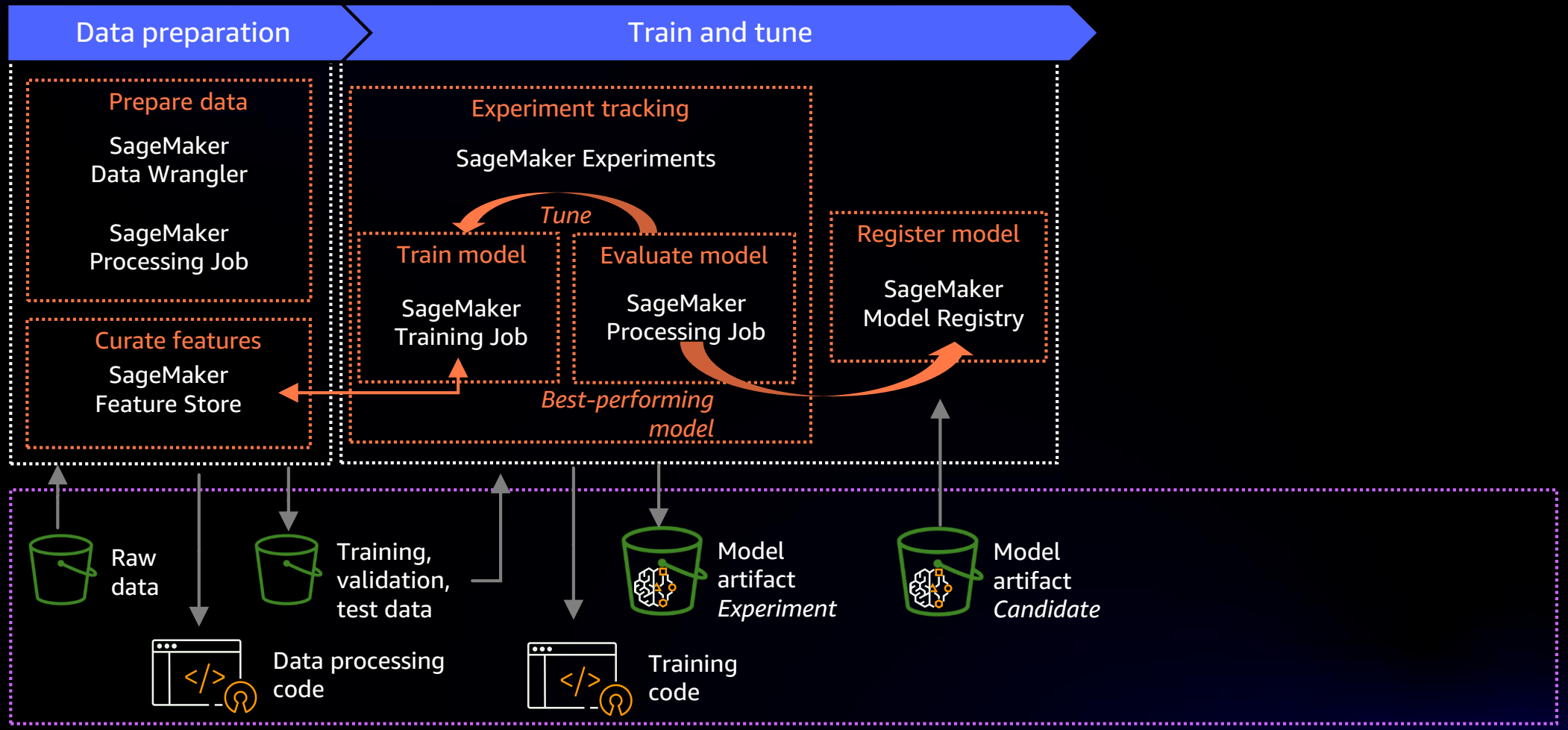
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MLOps-ready features and capabilities



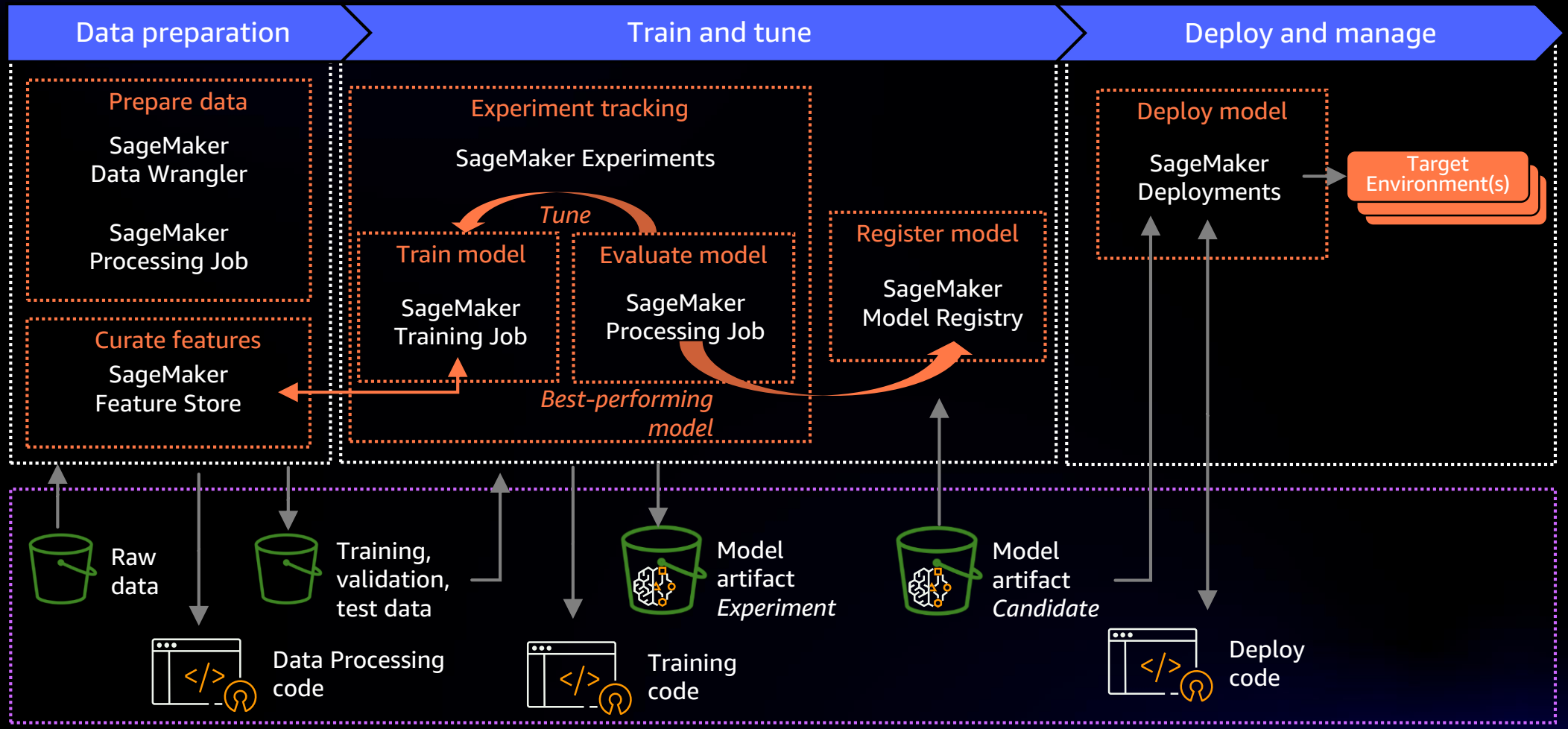
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MLOps-ready features and capabilities



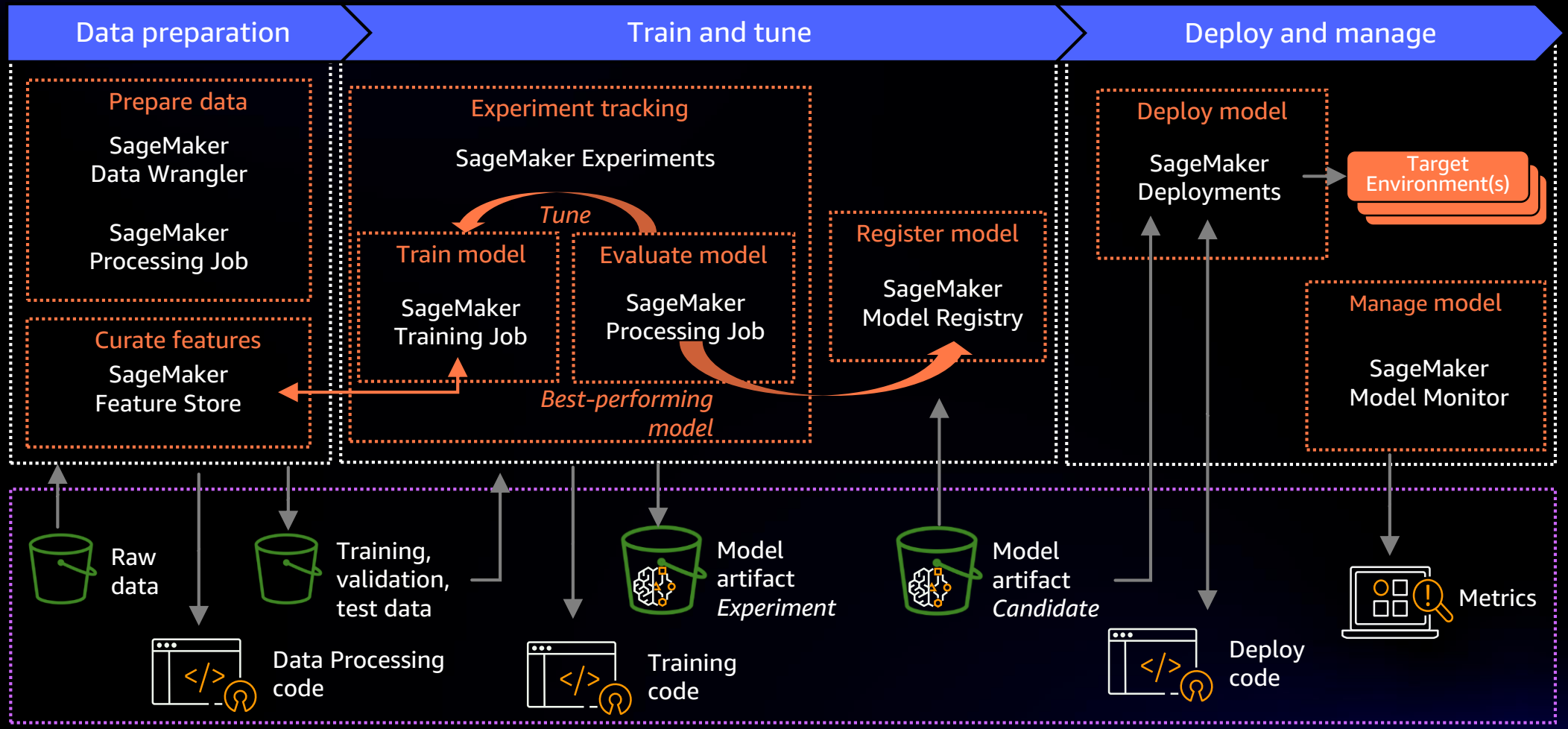
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MLOps-ready features and capabilities



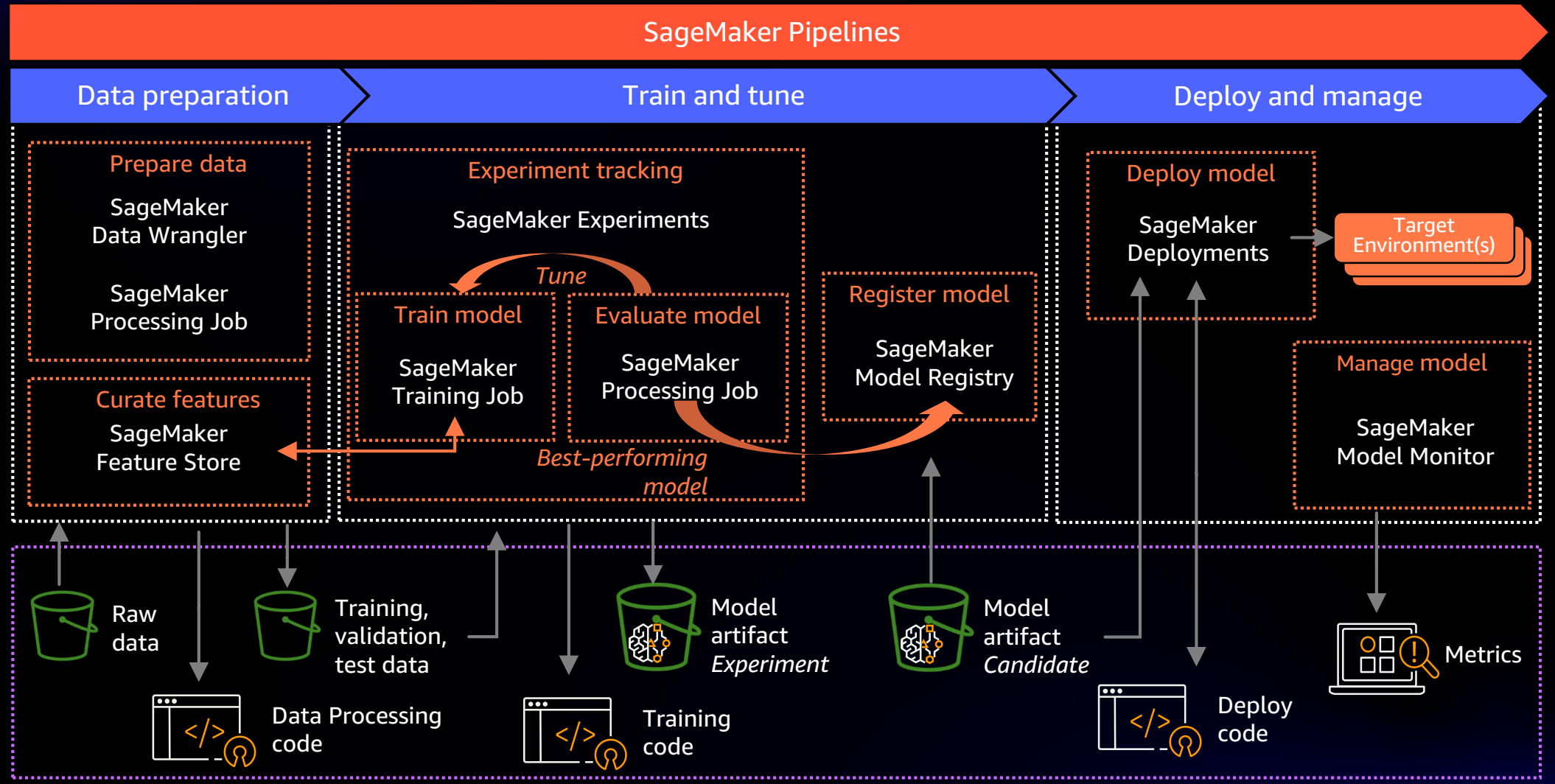
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MLOps-ready features and capabilities



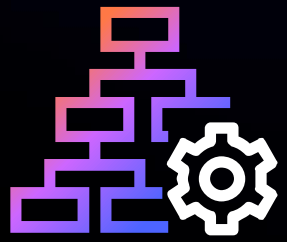
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MLOps-ready features and capabilities



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

SageMaker Pipelines



Python SDK for quickly and easily building ML workflows



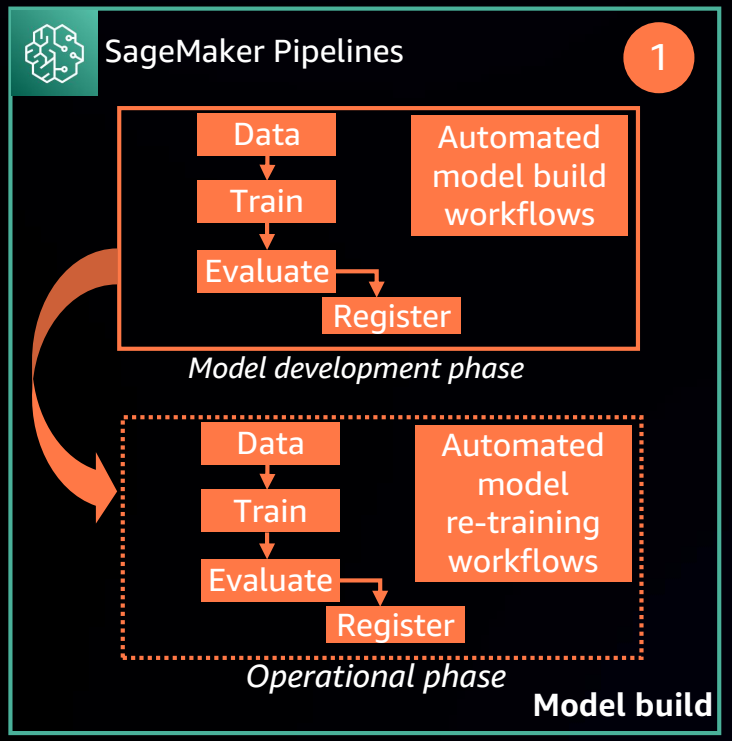
Catalog models to manage models at scale and trigger automated deployment workflows



Built-in support for CI/CD and end-to-end lineage tracking

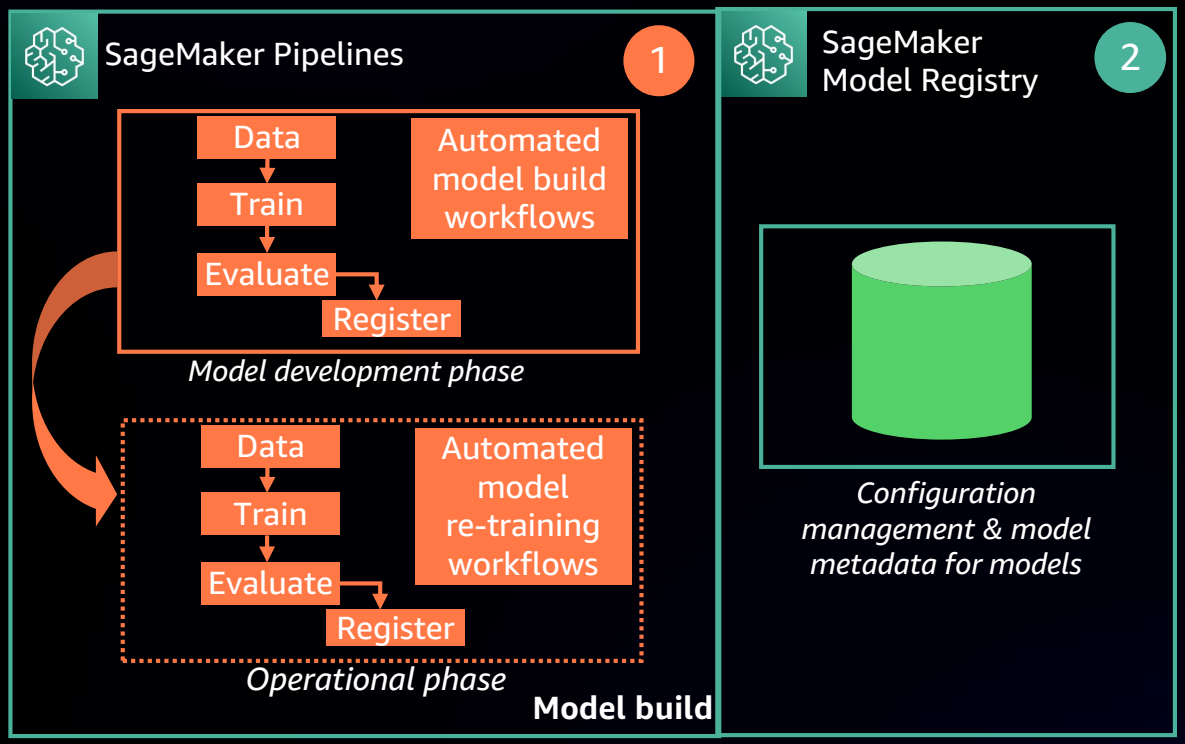
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SageMaker Pipelines: components



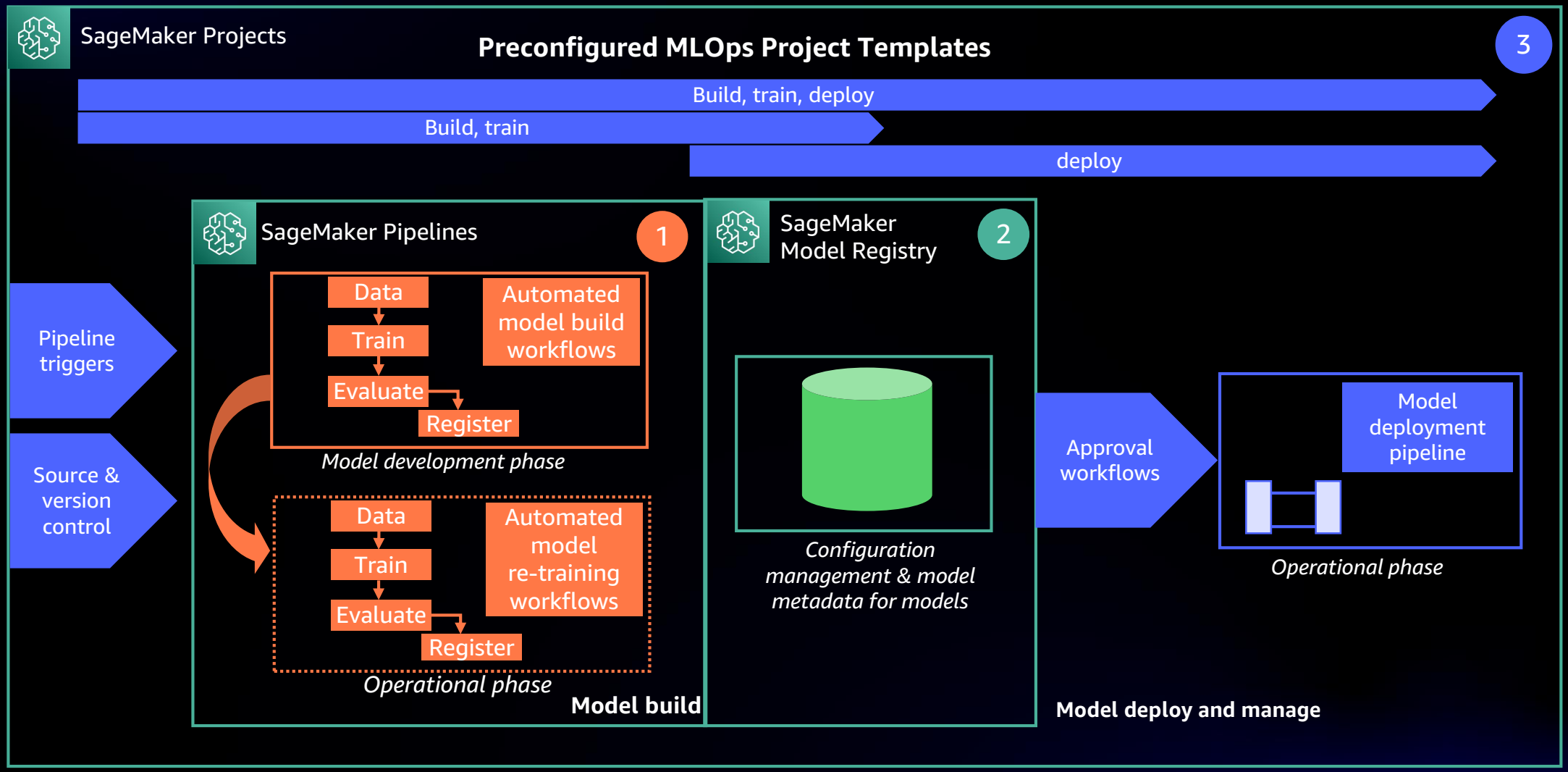
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SageMaker Pipelines: components



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

SageMaker Pipelines: components



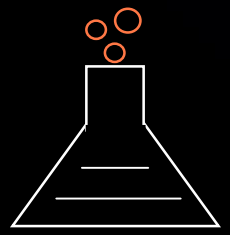
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Demo – SageMaker: Projects and Pipelines

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

SageMaker: Projects and Pipelines

What's new?



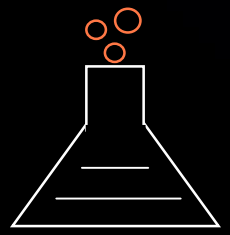
Integration with SageMaker Experiments

Automatically track pipeline executions – no setup required!

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

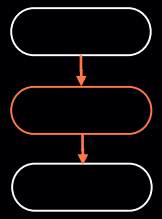
SageMaker: Projects and Pipelines

What's new?



Integration with SageMaker Experiments

Automatically track pipeline executions – no setup required!



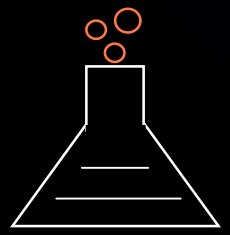
New natively supported pipeline steps

Including: SageMaker automatic model tuning, custom callback, and AWS Lambda

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

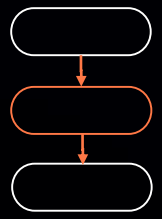
SageMaker: Projects and Pipelines

What's new?



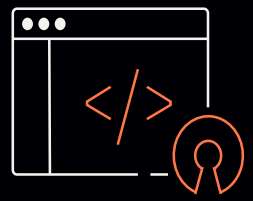
Integration with SageMaker Experiments

Automatically track pipeline executions – no setup required!



New natively supported pipeline steps

Including: SageMaker automatic model tuning, custom callback, and AWS Lambda



Support for third-party tools

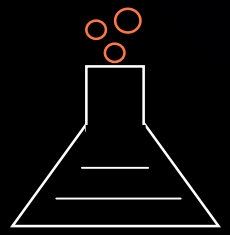
Including: GitHub, GitHub Enterprise, BitBucket Cloud, and Jenkins

aws

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

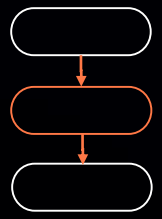
SageMaker: Projects and Pipelines

What's new?



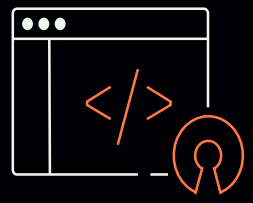
Integration with SageMaker Experiments

Automatically track pipeline executions – no setup required!



New natively supported pipeline steps

Including: SageMaker automatic model tuning, custom callback, and AWS Lambda



Support for third-party tools

Including: GitHub, GitHub Enterprise, BitBucket Cloud, and Jenkins



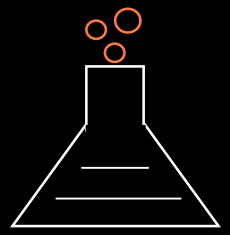
Automatic rollback to previous model version

Rollback deployment to previous version of model

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

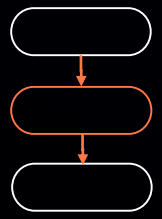
SageMaker: Projects and Pipelines

What's new?



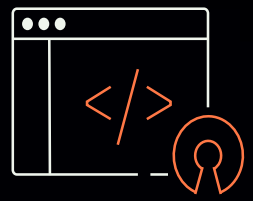
Integration with SageMaker Experiments

Automatically track pipeline executions – no setup required!



New natively supported pipeline steps

Including: SageMaker automatic model tuning, custom callback, and AWS Lambda



Support for third-party tools

Including: GitHub, GitHub Enterprise, BitBucket Cloud, and Jenkins



Automatic rollback to previous model version

Rollback deployment to previous version of model



New model registry features

Cross-account model approvals and deployments

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

Vanguard’s MLOps journey

aws

Vanguard by the numbers

Mission

"To take a stand for all investors, to treat them fairly, and give them the best chance for investment success"

Founded: May 1, 1975

Global AUM: $8.3 trillion

Funds worldwide: 417

Global offices: 19

Employees (crew): 17,300

Clients: 30+ million

Note: Information as of 09/30/2021

Objectives to adopt ML at scale



Reduce lead time to set up usable platform

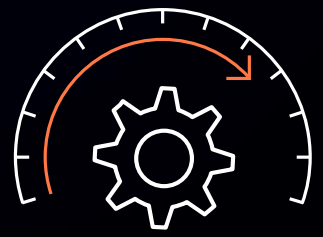
Less than 24 hours to set up usable platform

Objectives to adopt ML at scale



Reduce lead time to set up usable platform

Less than 24 hours to set up usable platform



Reduce lead time to enable consumption

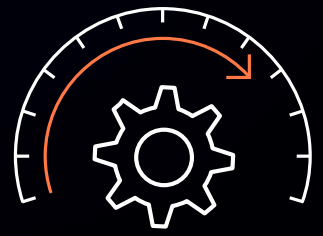
Automate 100% of deployment

Objectives to adopt ML at scale



Reduce lead time to set up usable platform

Less than 24 hours to set up usable platform



Reduce lead time to enable consumption

Automate 100% of deployment



Reduce risk due to use of ML

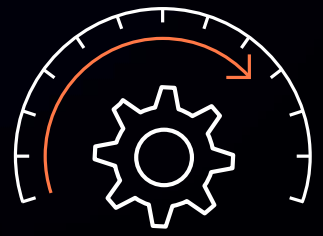
Automate validation of controls adherence

Objectives to adopt ML at scale



Reduce lead time to set up usable platform

Less than 24 hours to set up usable platform



Reduce lead time to enable consumption

Automate 100% of deployment



Reduce risk due to use of ML

Automate validation of controls adherence



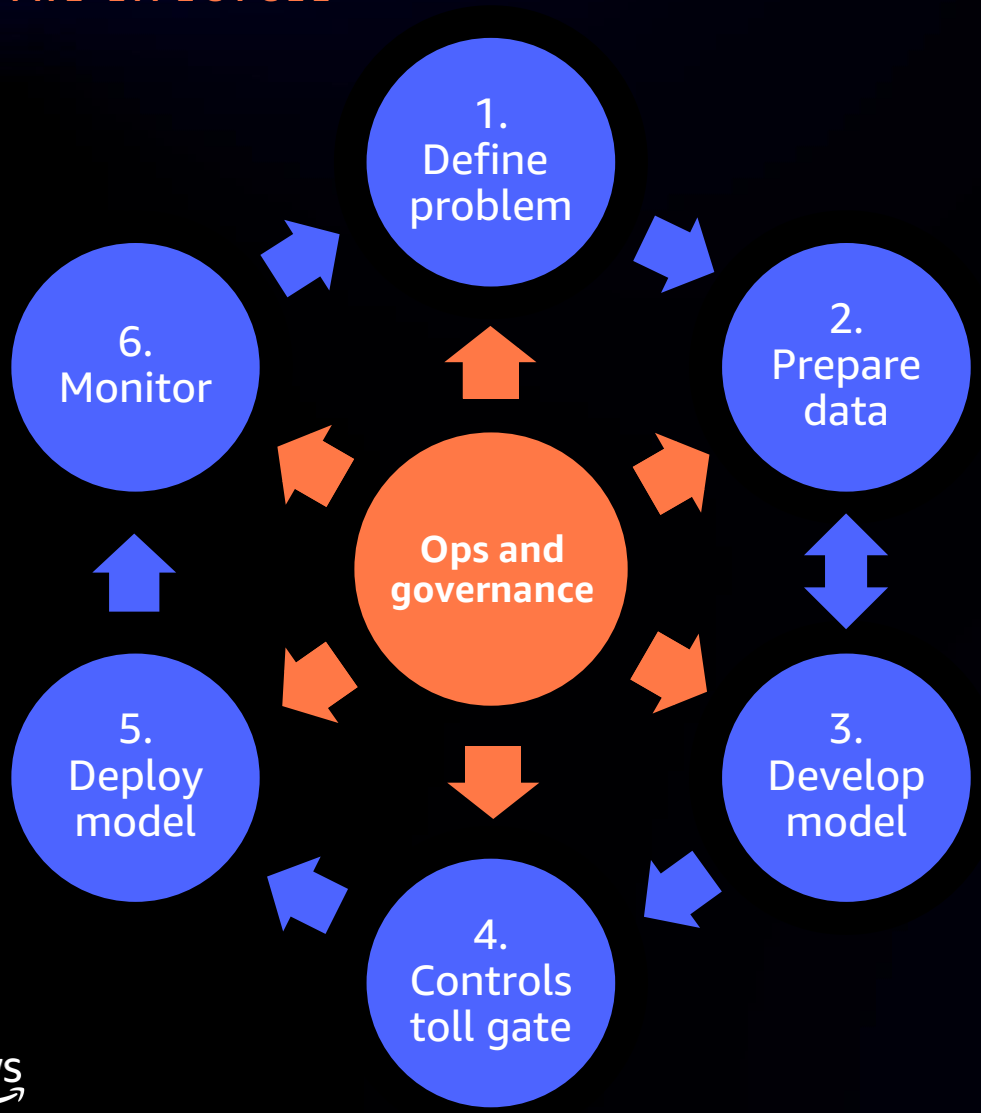
Increase client satisfaction for platforms

10% YoY improvement to satisfaction

aws

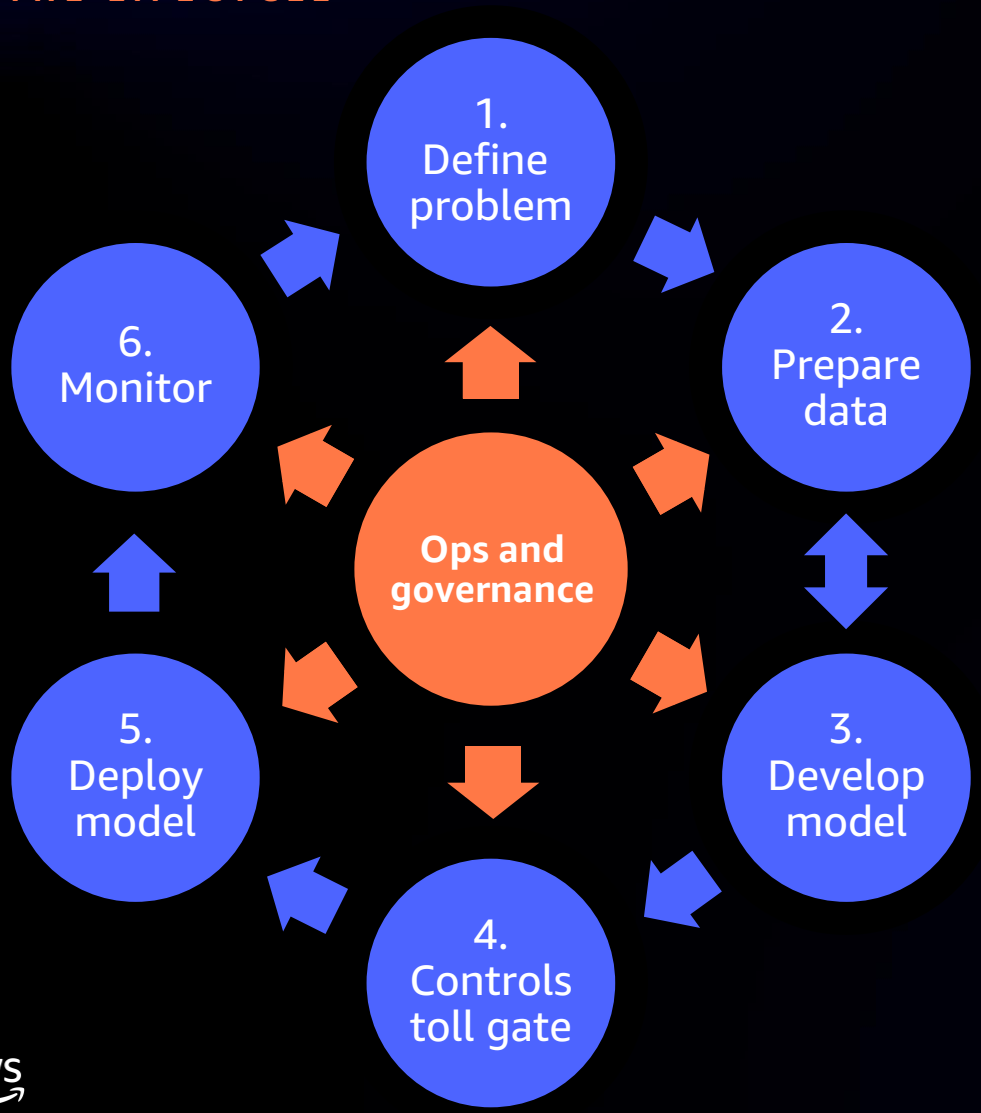
MDLC and capabilities related to MLOps

OPS AND GOVERNANCE CAPABILITIES ARE NOT AN ISOLATED PHASE BUT EMBEDDED IN EACH PHASE OF THE LIFECYCLE



MDLC and capabilities related to MLOps

OPS AND GOVERNANCE CAPABILITIES ARE NOT AN ISOLATED PHASE BUT EMBEDDED IN EACH PHASE OF THE LIFECYCLE



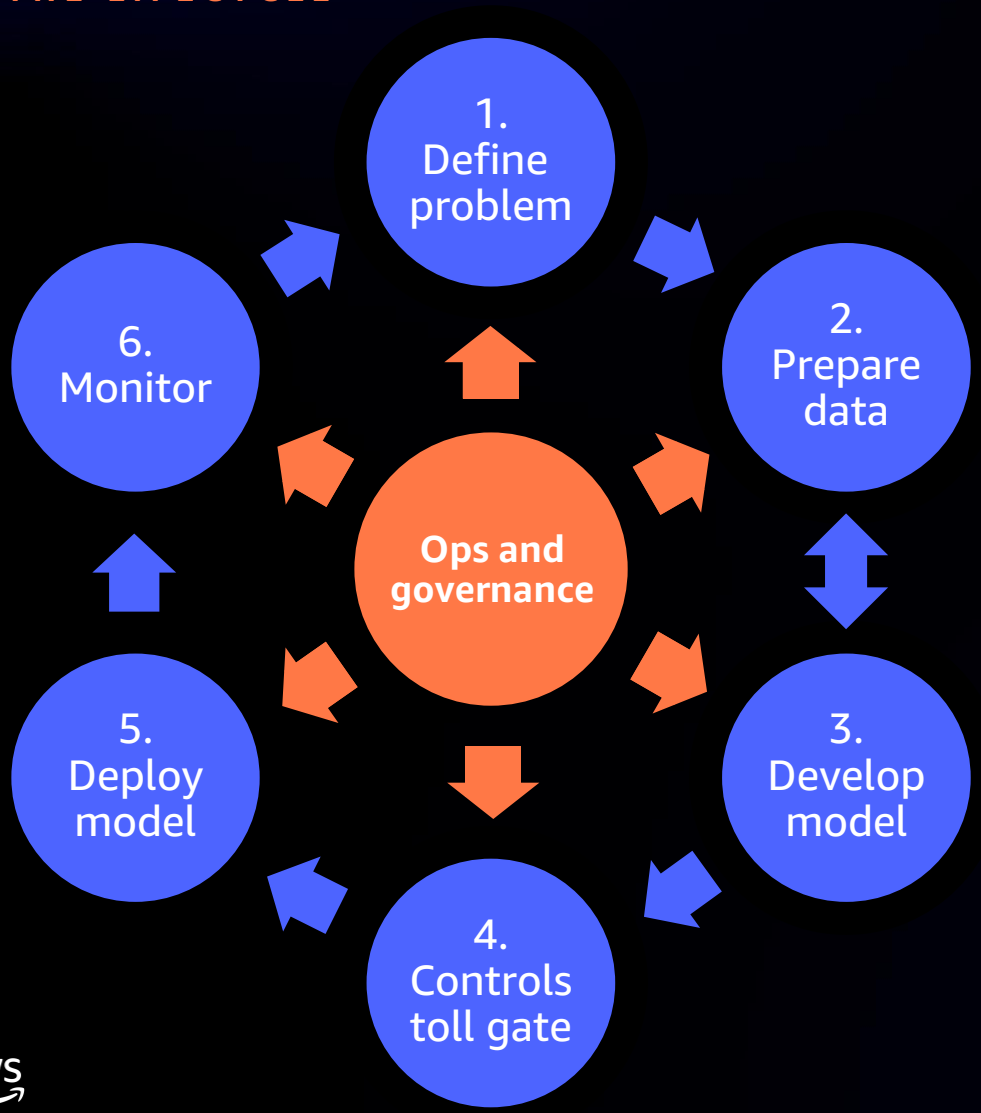
Define problem:

- Document problem statement(s)
- Establish goals
- Establish and capture ROI
- Identify potential risk
- Establish team (product owner, model owner, DS, MLOps engineer, …)

aws

MDLC and capabilities related to MLOps

OPS AND GOVERNANCE CAPABILITIES ARE NOT AN ISOLATED PHASE BUT EMBEDDED IN EACH PHASE OF THE LIFECYCLE



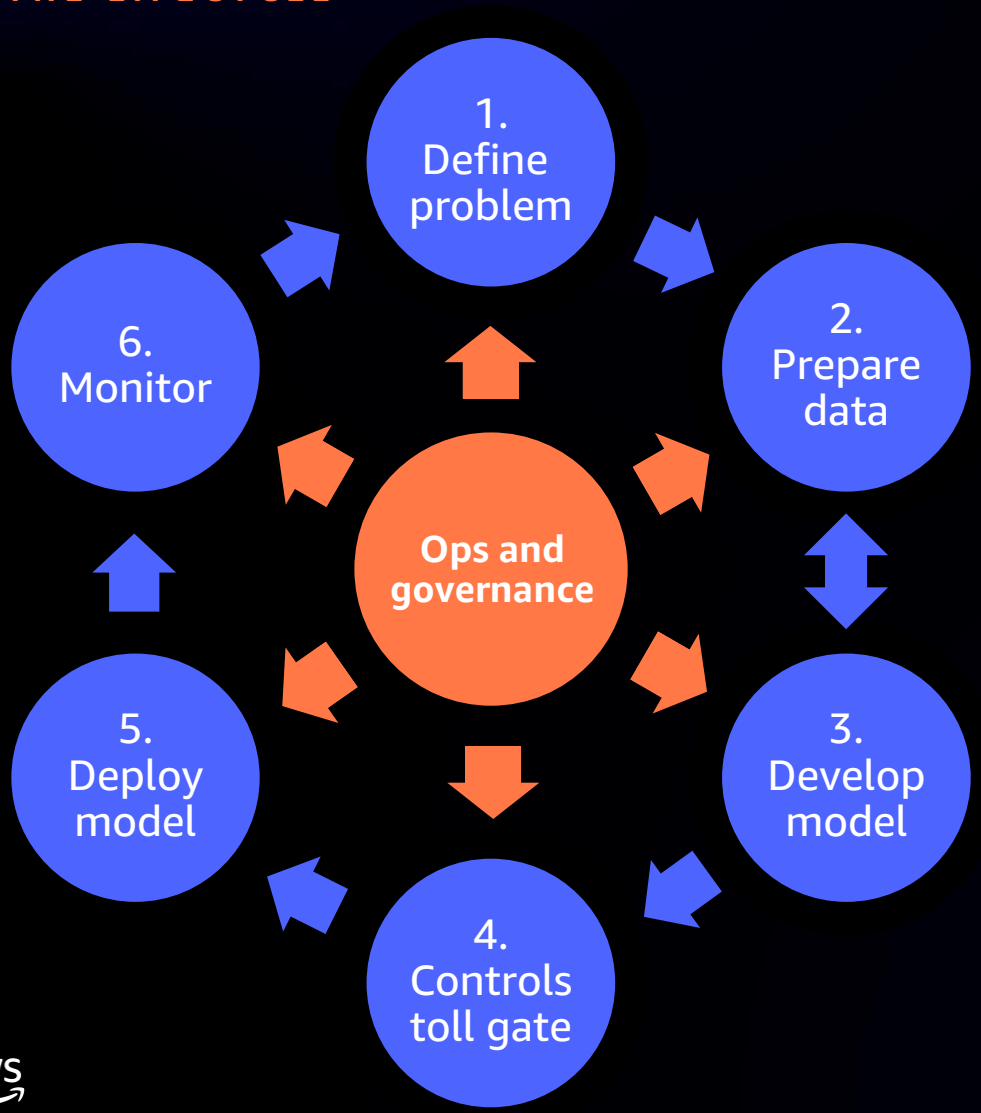
Define problem:
- Document problem statement(s)
- Establish goals
- Establish and capture ROI
- Identify potential risk
- Establish team (product owner, model owner, DS, MLOps engineer, …)

Prepare data:
- Identify, understand, and obtain usage approval to data
- Set up access to data
- Engineer data and features
- Check for bias and imbalances
- Develop and adjust data pipelines

aws

MDLC and capabilities related to MLOps

OPS AND GOVERNANCE CAPABILITIES ARE NOT AN ISOLATED PHASE BUT EMBEDDED IN EACH PHASE OF THE LIFECYCLE



Define problem:
- Document problem statement(s)
- Establish goals
- Establish and capture ROI
- Identify potential risk
- Establish team (product owner, model owner, DS, MLOps engineer, …)

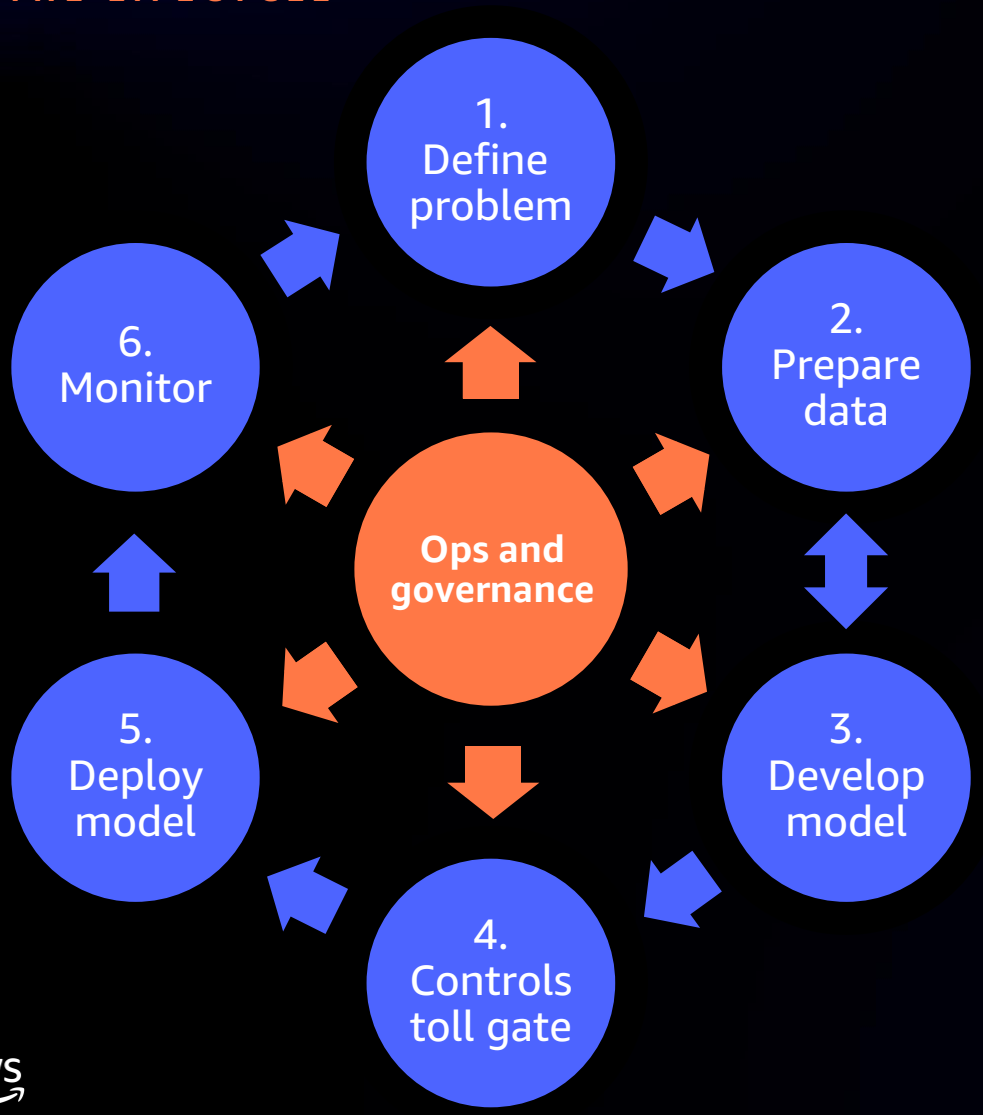
Prepare data:
- Identify, understand, and obtain usage approval to data
- Set up access to data
- Engineer data and features
- Check for bias and imbalances
- Develop and adjust data pipelines

Develop model:
- Explore algorithms and frameworks
- Build and train models via experiments
- Test, tune, optimize, and validate models via trials
- Check for biases introduced by model
- Establish model explainability
- Archive code, configurations, and model artifact in repository

aws

MDLC and capabilities related to MLOps

OPS AND GOVERNANCE CAPABILITIES ARE NOT AN ISOLATED PHASE BUT EMBEDDED IN EACH PHASE OF THE LIFECYCLE



Define problem:
- Document problem statement(s)
- Establish goals
- Establish and capture ROI
- Identify potential risk
- Establish team (product owner, model owner, DS, MLOps engineer, …)

Prepare data:
- Identify, understand, and obtain usage approval to data
- Set up access to data
- Engineer data and features
- Check for bias and imbalances
- Develop and adjust data pipelines

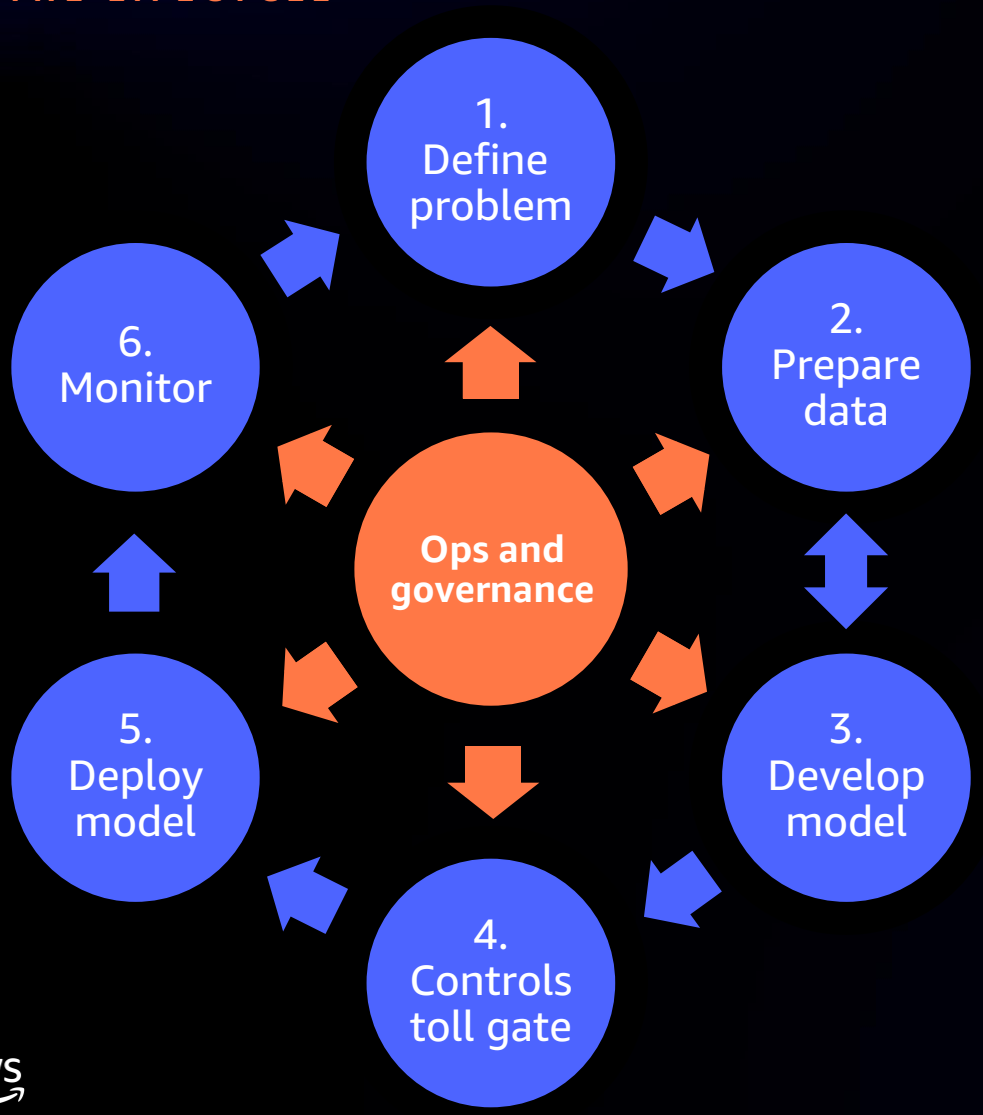
Develop model:
- Explore algorithms and frameworks
- Build and train models via experiments
- Test, tune, optimize, and validate models via trials
- Check for biases introduced by model
- Establish model explainability
- Archive code, configurations, and model artifact in repository

Controls toll gate:
- Establish risk tiers
- Document adherence to controls for established risk tier
- Obtain approval to proceed with deployment

aws

MDLC and capabilities related to MLOps

OPS AND GOVERNANCE CAPABILITIES ARE NOT AN ISOLATED PHASE BUT EMBEDDED IN EACH PHASE OF THE LIFECYCLE



Define problem:
- Document problem statement(s)
- Establish goals
- Establish and capture ROI
- Identify potential risk
- Establish team (product owner, model owner, DS, MLOps engineer, …)

Prepare data:
- Identify, understand, and obtain usage approval to data
- Set up access to data
- Engineer data and features
- Check for bias and imbalances
- Develop and adjust data pipelines

Develop model:
- Explore algorithms and frameworks
- Build and train models via experiments
- Test, tune, optimize, and validate models via trials
- Check for biases introduced by model
- Establish model explainability
- Archive code, configurations, and model artifact in repository

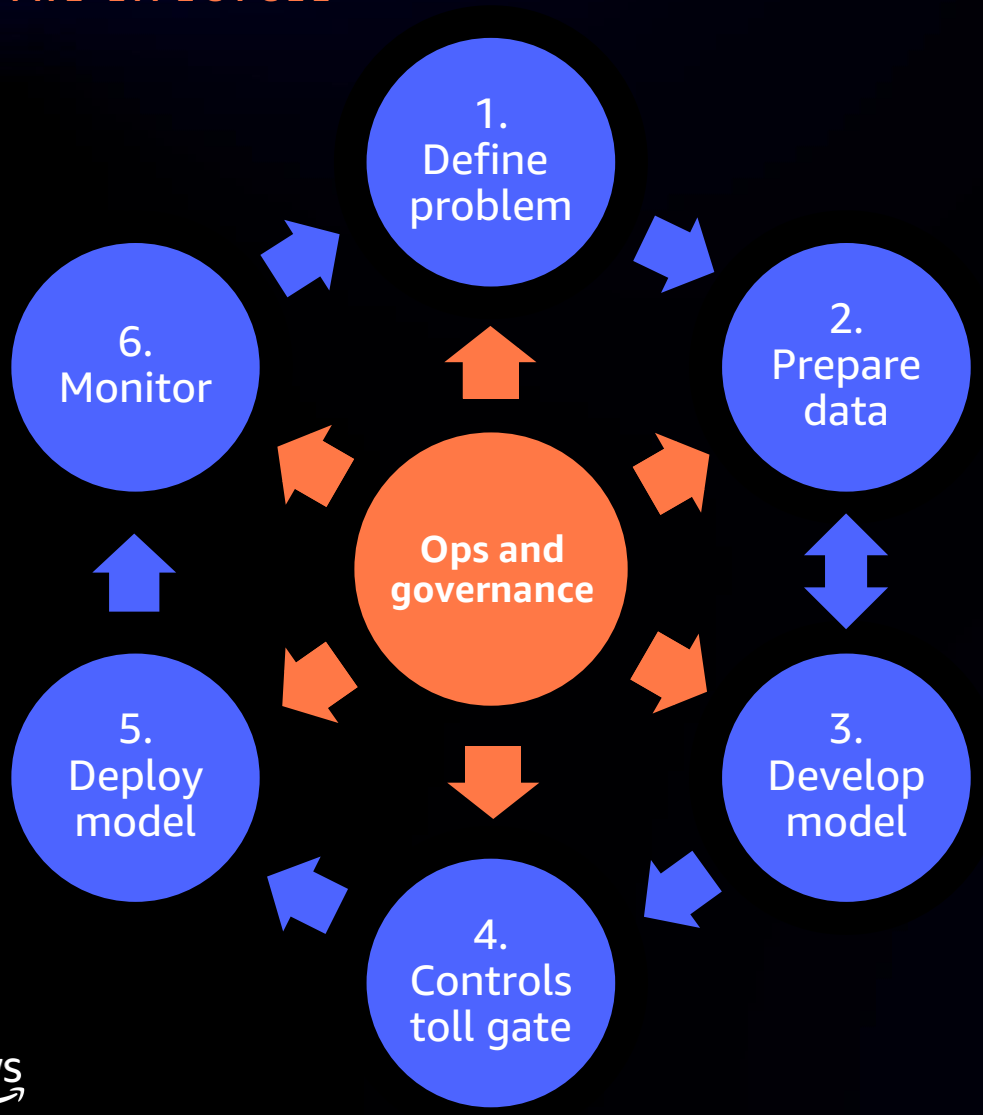
Controls toll gate:
- Establish risk tiers
- Document adherence to controls for established risk tier
- Obtain approval to proceed with deployment

Deploy model and monitor:
- Validate code quality and security standards adherence
- Validate controls and ITIL adherence
- CI/CD pipeline to deploy models and feature pipeline
- Setup monitoring for models (for data drift, model drift, model bias, quality and performance, etc.)
- Demonstrate ROI

aws

MDLC and capabilities related to MLOps

OPS AND GOVERNANCE CAPABILITIES ARE NOT AN ISOLATED PHASE BUT EMBEDDED IN EACH PHASE OF THE LIFECYCLE



Define problem:
- Document problem statement(s)
- Establish goals
- Establish and capture ROI
- Identify potential risk
- Establish team (product owner, model owner, DS, MLOps Engineer, …)

Prepare data:
- Identify, understand, and obtain usage approval to data
- Set up access to data
- Engineer data and features
- Check for bias and imbalances
- Develop and adjust data pipelines

Develop model:
- Explore algorithms and frameworks
- Build and train models via experiments
- Test, tune, optimize, and validate models via trials
- Check for biases introduced by model
- Establish model explainability
- Archive code, configurations, and model artifact in repository

Controls toll gate:
- Establish risk tiers
- Document adherence to controls for established risk tier
- Obtain approval to proceed with deployment

Deploy model:
- Validate code quality and security standards adherence
- Validate controls and ITIL adherence
- CI/CD pipeline to deploy models and feature pipeline
- Various strategies for deployment (blue/green, rolling, canary, etc.)

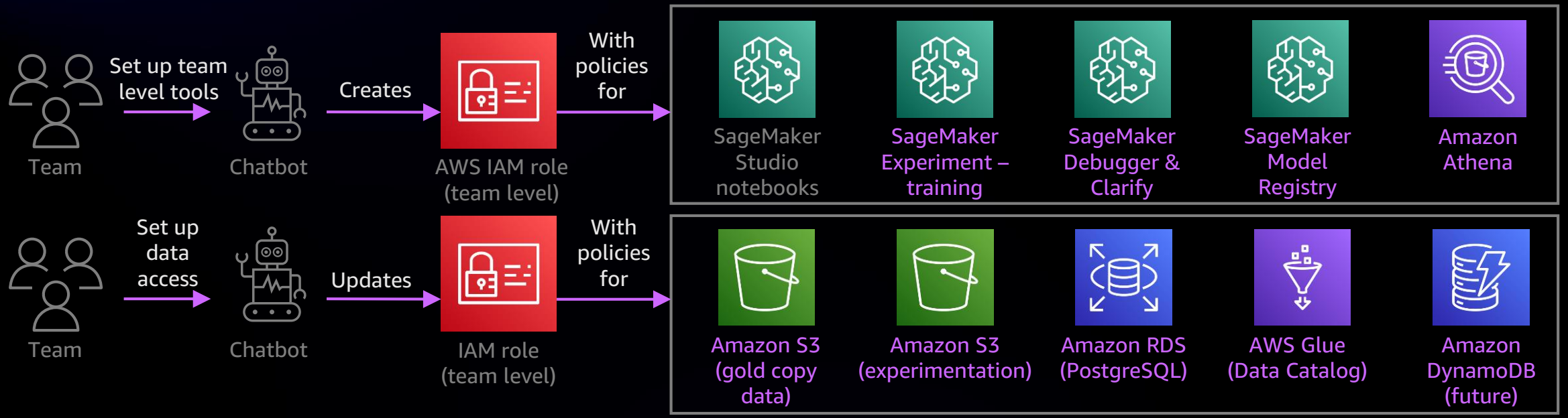
Monitor:
- Setup monitoring for models (for data drift, model drift, model bias, quality and performance, etc.)
- Alerts and notifications setup for monitoring drifts
- Demonstrate ROI

aws

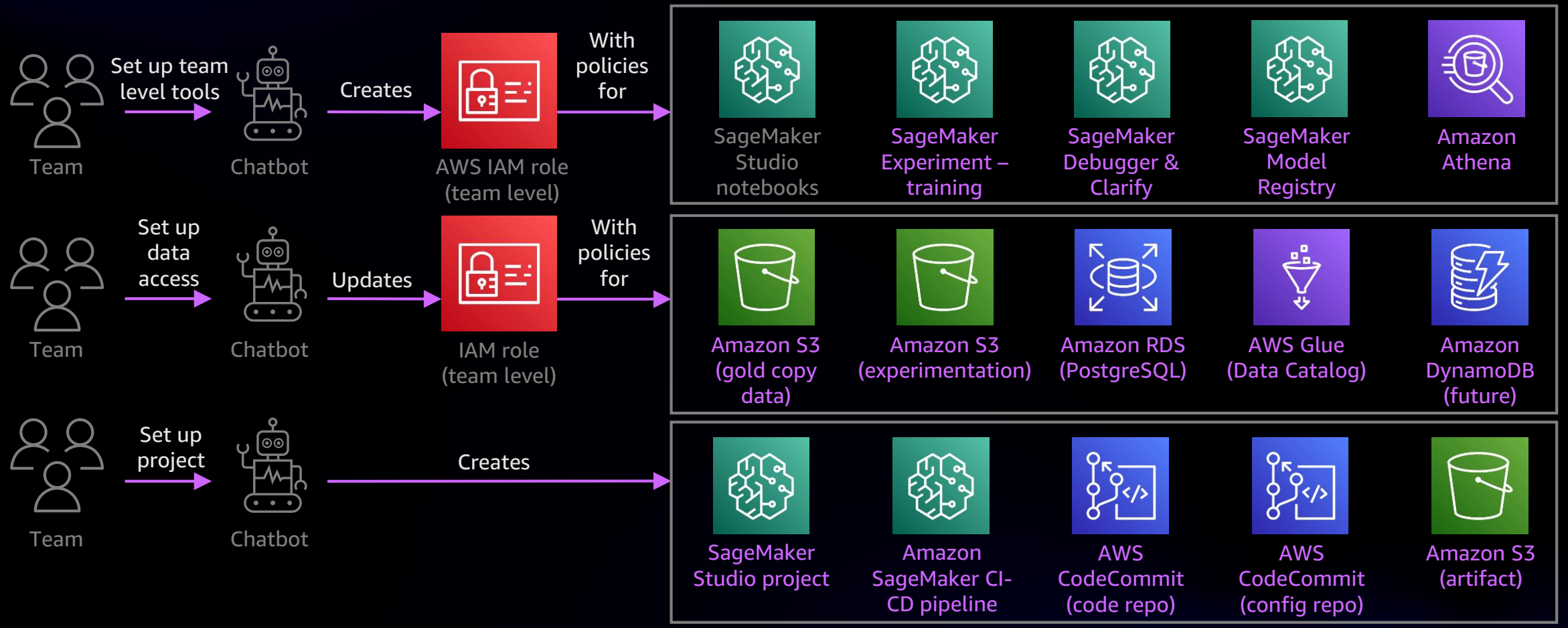
MLOps: Define problem and prepare data phase



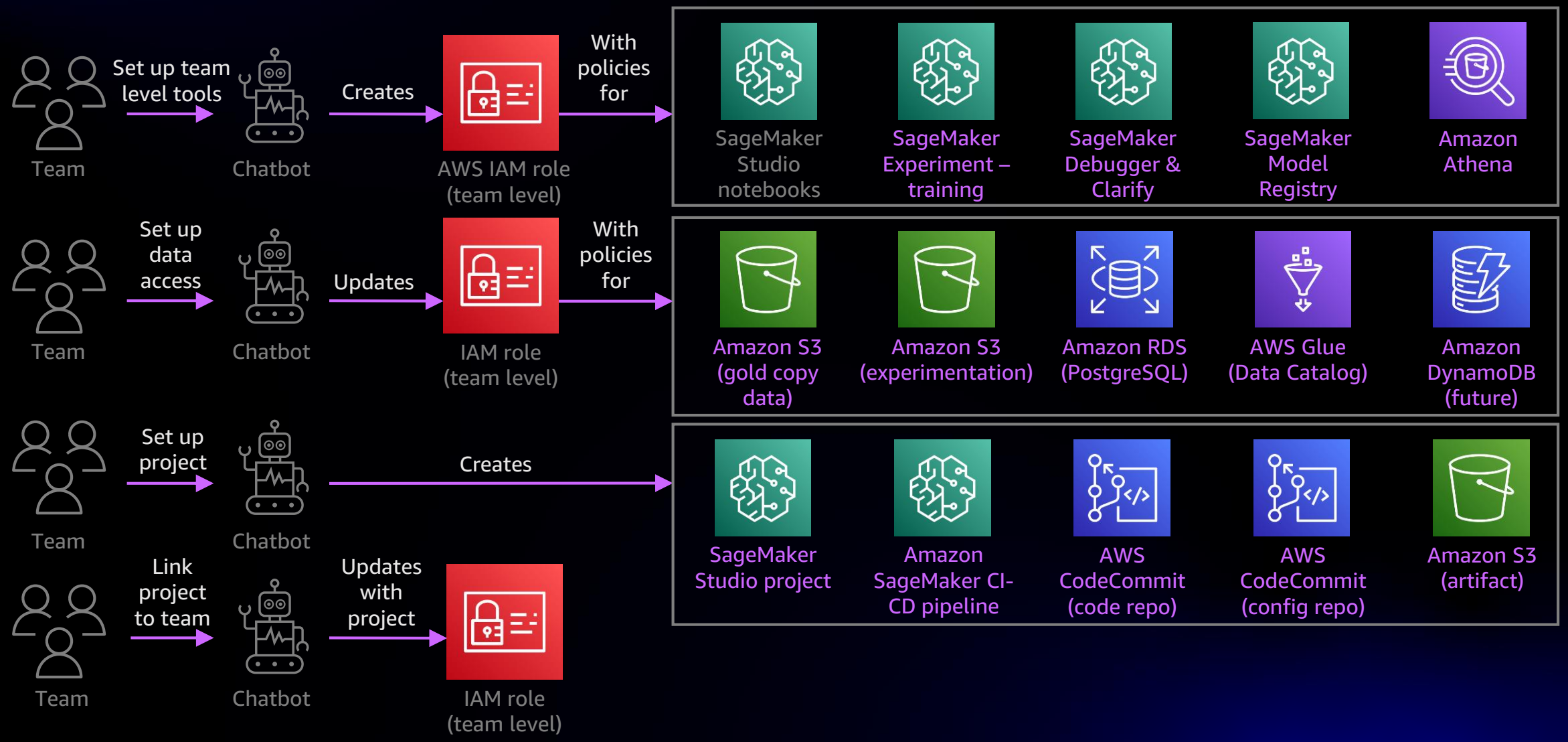
MLOps: Define problem and prepare data phase



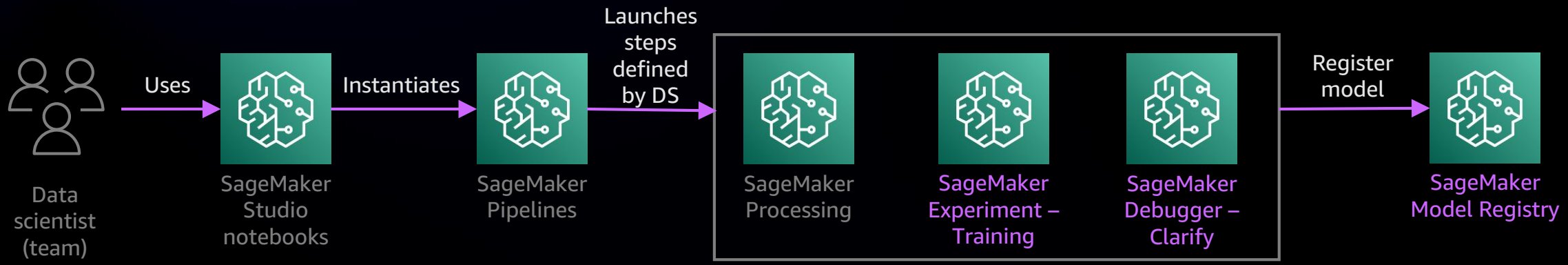
MLOps: Define problem and prepare data phase



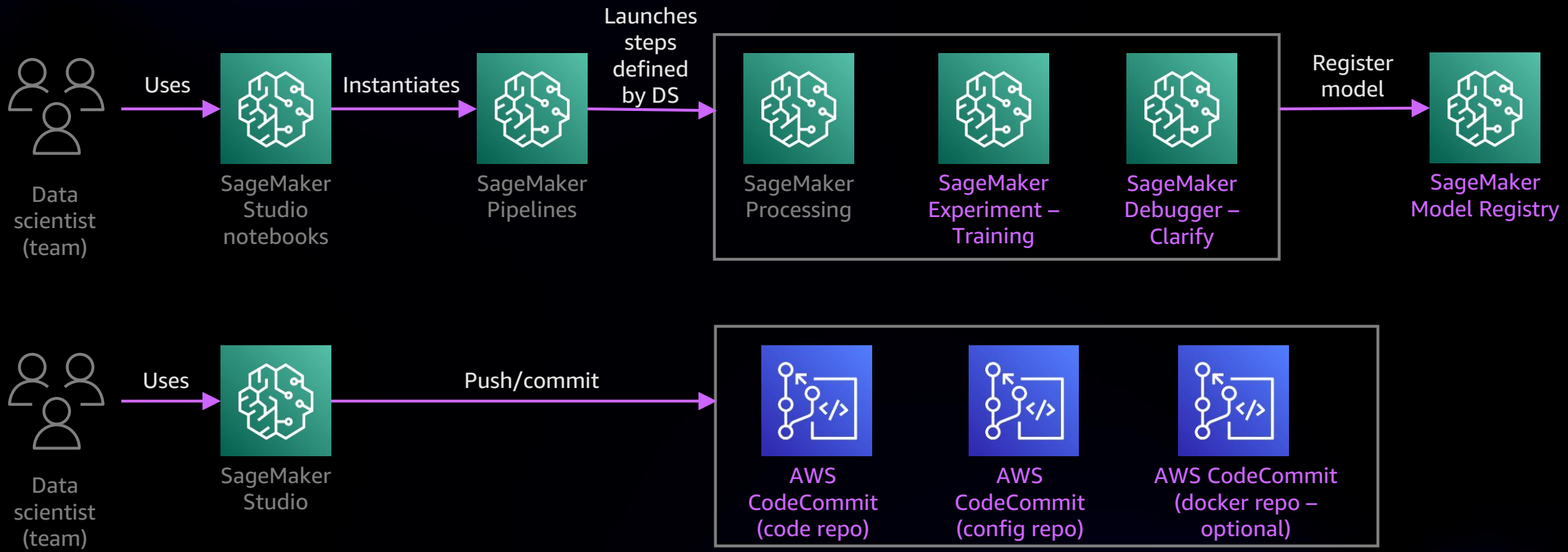
MLOps: Define problem and prepare data phase



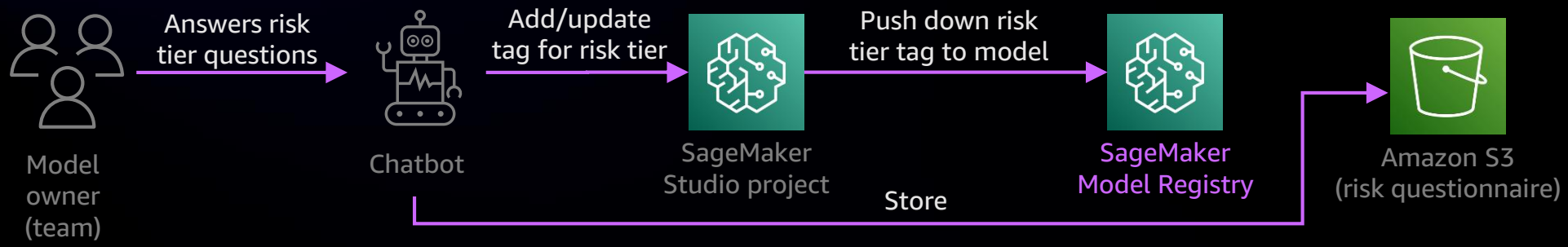
MLOps: Develop phase



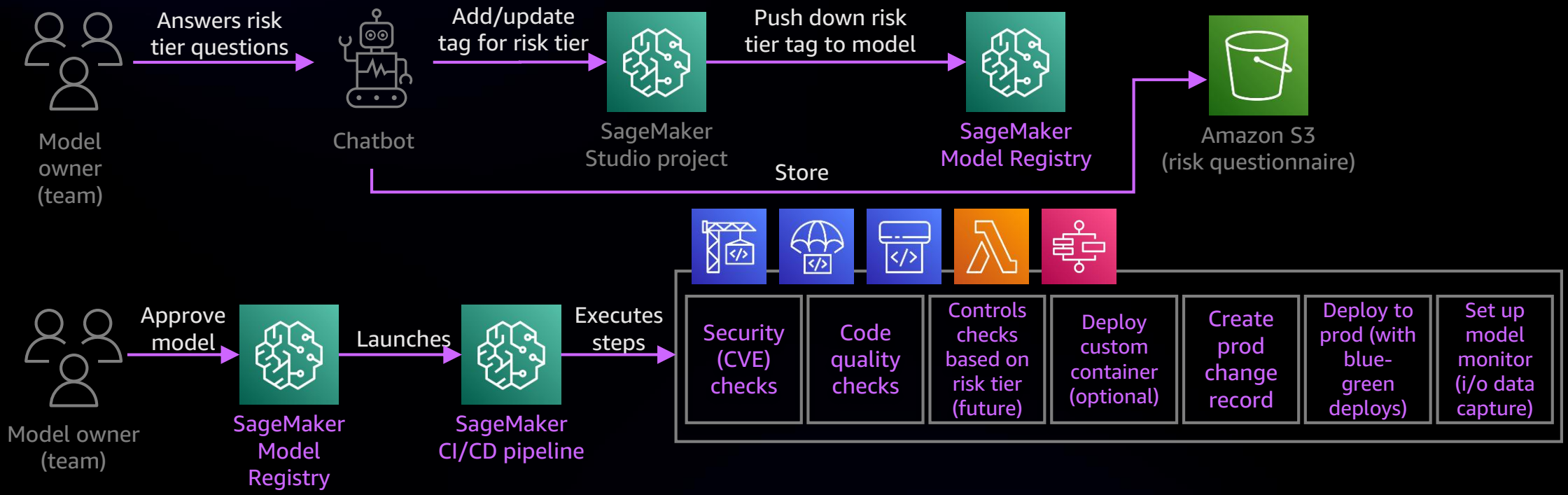
MLOps: Develop phase



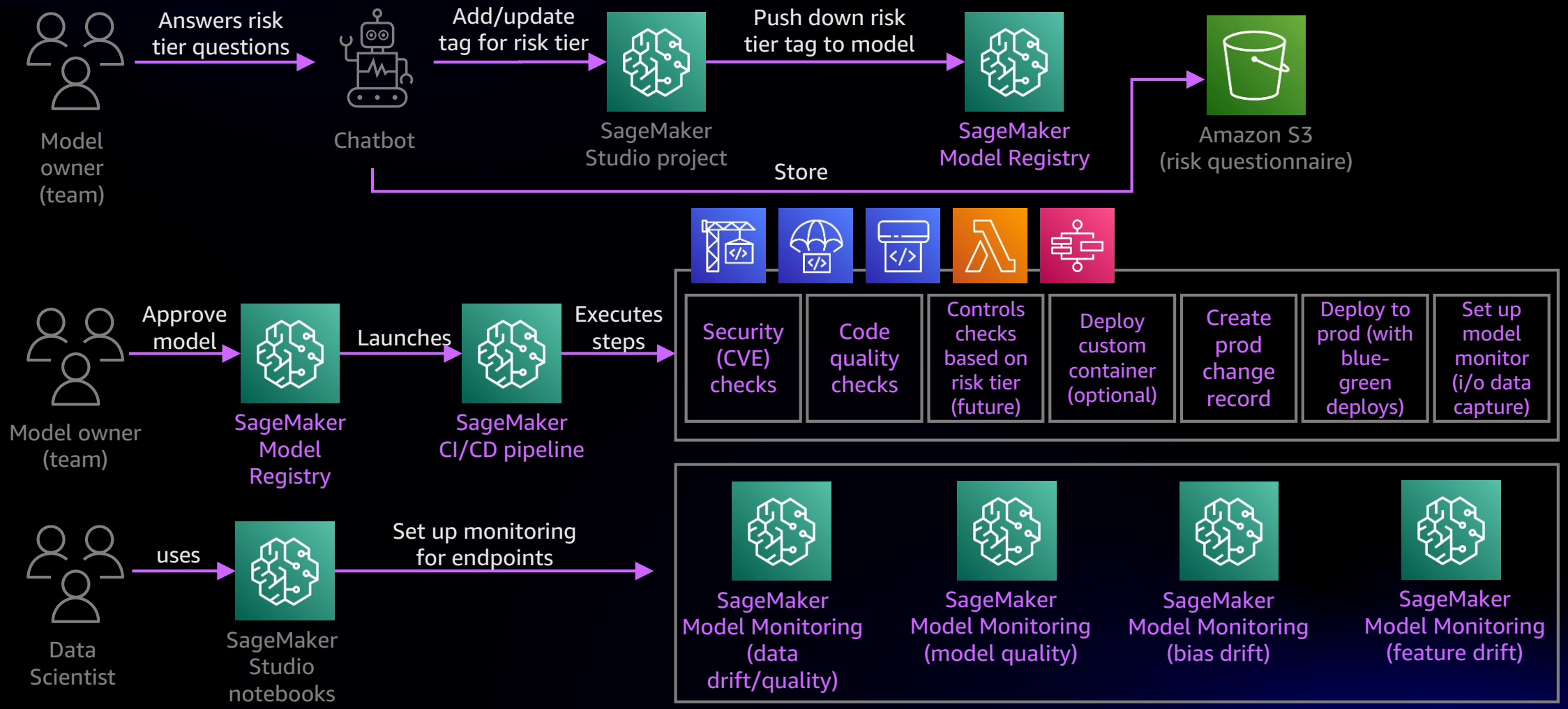
MLOps: Controls, deploy, and monitor phase



MLOps: Controls, deploy, and monitor phase



MLOps: Controls, deploy, and monitor phase



Key takeaways

Building MLOps and governance platforms is a journey

Key takeaways

Building MLOps and governance platforms is a journey



Many capabilities already available in Amazon SageMaker

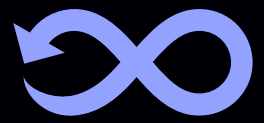
Invest in enabling these capabilities

aws

Key takeaways

Building MLOps and governance platforms is a journey





Many capabilities already available in Amazon SageMaker

Invest in enabling these capabilities

Enabled CI-CD for AWS native and BYO containers

Expand CI-CD for pre-trained models

aws

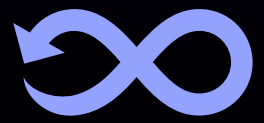
Key takeaways

Building MLOps and governance platforms is a journey



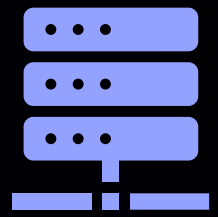
Many capabilities already available in Amazon SageMaker

Invest in enabling these capabilities



Enabled CI-CD for AWS native and BYO containers

Expand CI-CD for pre-trained models



Enabled feature engineering capabilities

Expand to enable Feature Store and CI-CD for feature management

aws

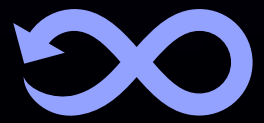
Key takeaways

Building MLOps and governance platforms is a journey



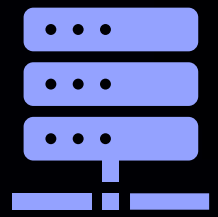
Many capabilities already available in Amazon SageMaker

Invest in enabling these capabilities



Enabled CI-CD for AWS native and BYO containers

Expand CI-CD for pre-trained models



Enabled feature engineering capabilities

Expand to enable Feature Store and CI-CD for feature management



Enabled model monitoring for endpoints

Expand model monitoring for batch and operational health

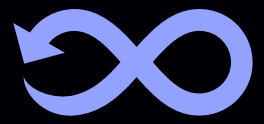
Key takeaways

Building MLOps and governance platforms is a journey



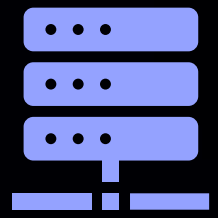
Many capabilities already available in Amazon SageMaker

Invest in enabling these capabilities



Enabled CI-CD for AWS native and BYO containers

Expand CI-CD for pre-trained models



Enabled feature engineering capabilities

Expand to enable Feature Store and CI-CD for feature management



Enabled model monitoring for endpoints

Expand model monitoring for batch and operational health



Enabled CI-CD and automation for consumption

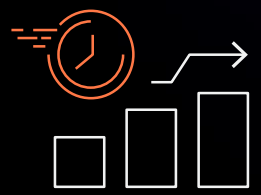
Expand to continuous training (CT) and continuous monitoring (CM)

aws

Wrap-up

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

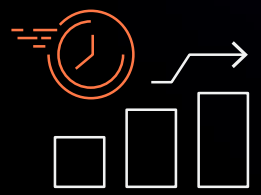
SageMaker MLOps: summary



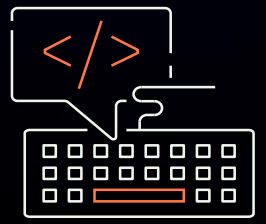
Start small
and iterate

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

SageMaker MLOps: summary



Start small and iterate



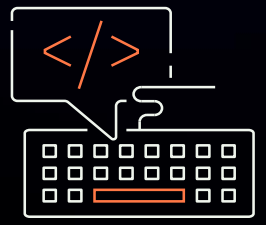
Practice Everything-as-Code for repeatability, consistency, and governance

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

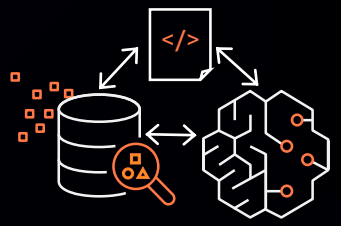
SageMaker MLOps: summary



Start small and iterate



Practice Everything-as-Code for repeatability, consistency, and governance



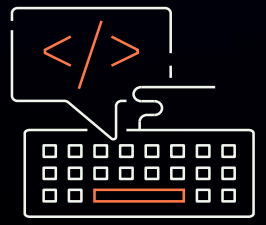
Consider ML specific inputs, outputs, and artifacts for model and pipeline lineage

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

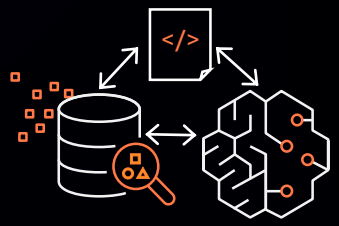
SageMaker MLOps: summary



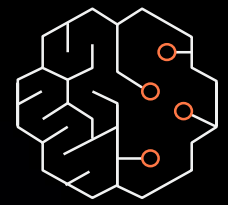
Start small and iterate



Practice Everything-as-Code for repeatability, consistency, and governance



Consider ML specific inputs, outputs, and artifacts for model and pipeline lineage



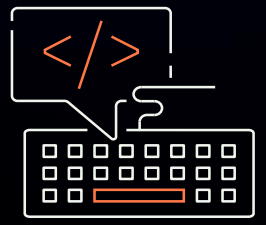
Use integrated SageMaker services when possible to reduce management and setup time

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

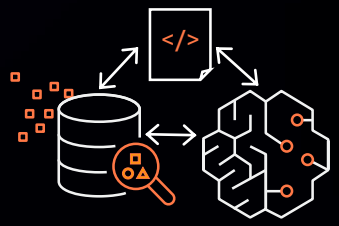
SageMaker MLOps: summary



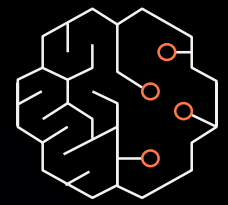
Start small and iterate



Practice Everything-as-Code for repeatability, consistency, and governance



Consider ML specific inputs, outputs, and artifacts for model and pipeline lineage



Use integrated SageMaker services when possible to reduce management and setup time



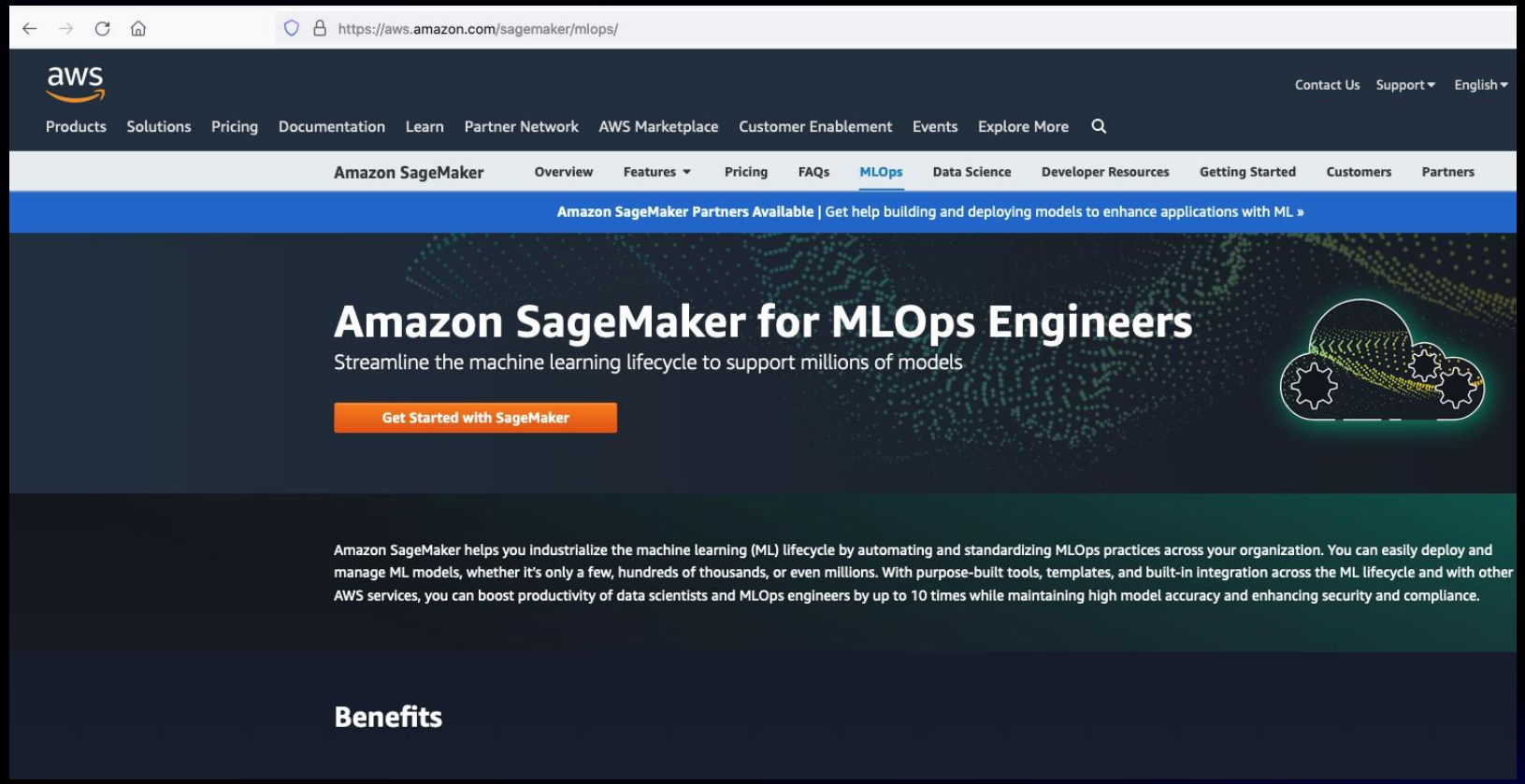
Implement a Model Registry as part of your ML pipeline for managing models

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

SageMaker MLOps: Next Steps

Check out the new MLOps Page

https://aws.amazon.com/sagemaker/mlops/



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

Thank you!

Shelbee Eigenbrode

linkedin.com/in/shelbee-eigenbrode

Ritesh Shah

linkedin.com/in/riteshbshah

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