
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

BOA203-R

Evolution of the machine learning development environment

Brooke Jamieson (they/them)

Senior Developer Advocate
AWS

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

How it started/how it’s going

What is MLOps?

What does MLOps look like in practice now?

If MLOps is so important, why isn’t it more popular?

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How it started/how it’s going

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The reach of ML is growing



Increased spending

By 2025, global spending on artificial intelligence (AI) will reach $204 billion

– IDC

IDC, “Worldwide spending on artificial intelligence,” http://bit.ly/3tLHX4V.

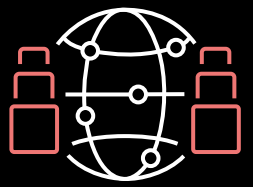


From piloting to operationalizing

By the end of 2024, 75% of enterprises will shift from piloting to operationalizing AI

– Gartner

Gartner, “Gartner identifies Top 10,” http://bit.ly/3XuWik2.



AI transformation

57% said that AI would transform their organization in the next three years

– Deloitte

Deloitte, “Thriving in the era of pervasive AI,” http://bit.ly/3gp9ENY.

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The reach of ML is growing



From piloting to operationalizing

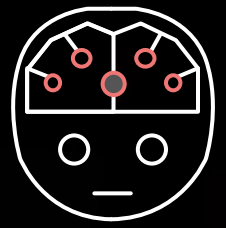
Only 53% of projects make it from AI prototypes to production

– Gartner

Gartner, “Top strategic technology trends for 2021” https://gtnr.it/3Bln3uU.

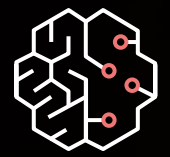
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What is it?



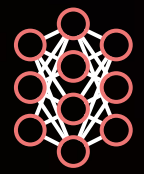
Artificial intelligence (AI)

Any technique that enables computers to mimic human intelligence using logic, if-then statements, and ML



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



Reinforcement learning (RL)

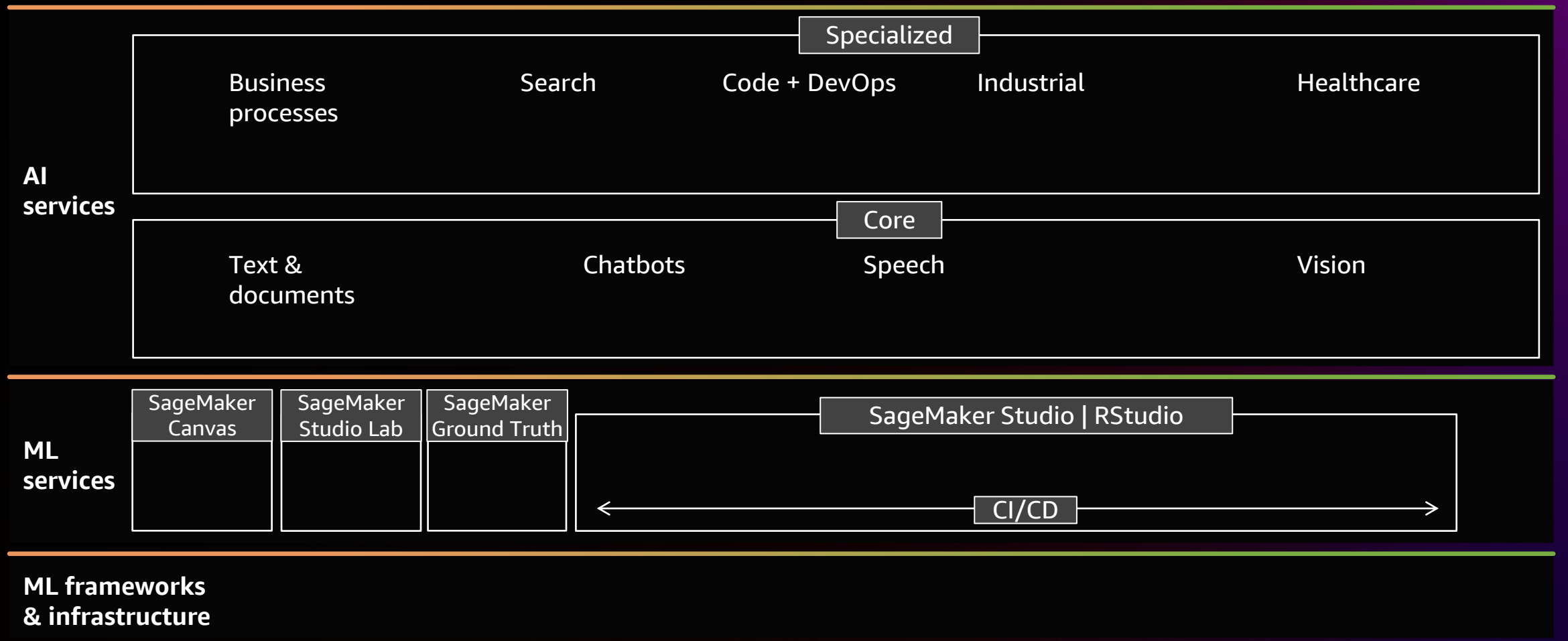
Reinforcement learning trains models by using a reward function of a desired outcome

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The AWS ML stack

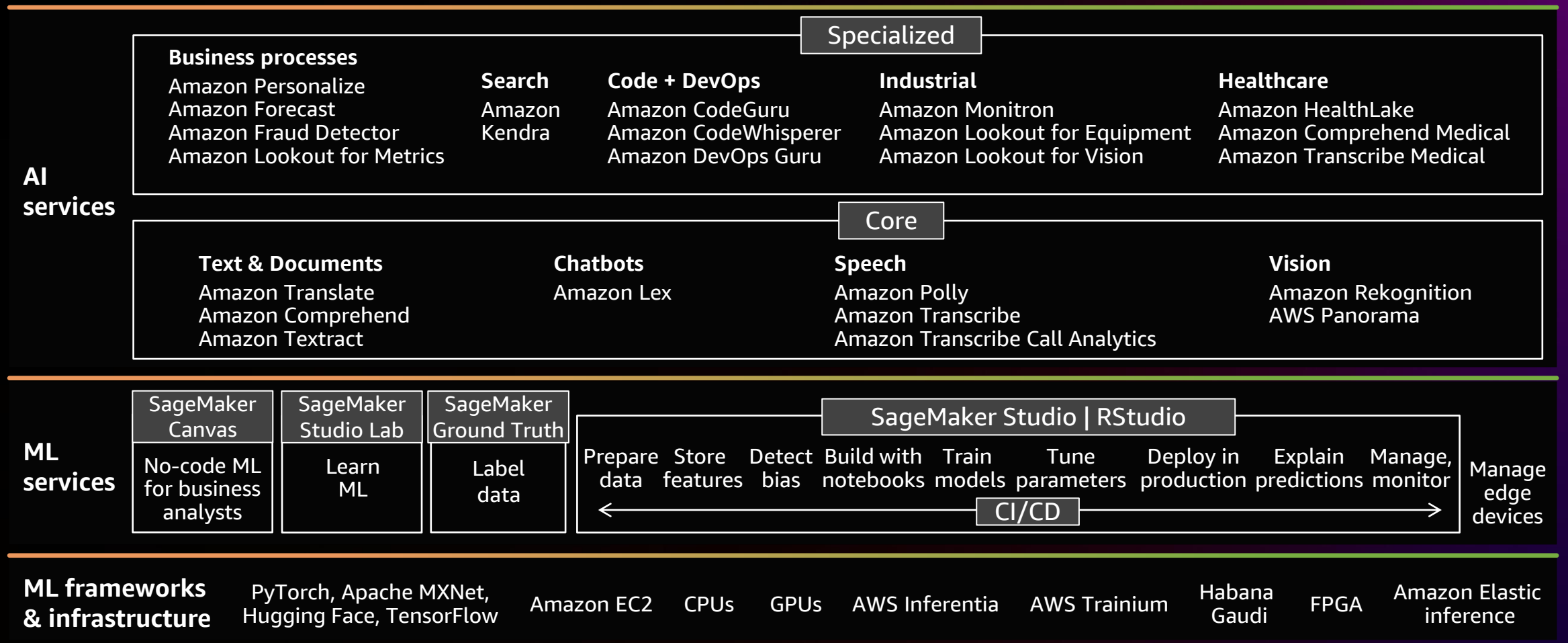
BROADEST AND MOST COMPLETE SET OF MACHINE LEARNING CAPABILITIES



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

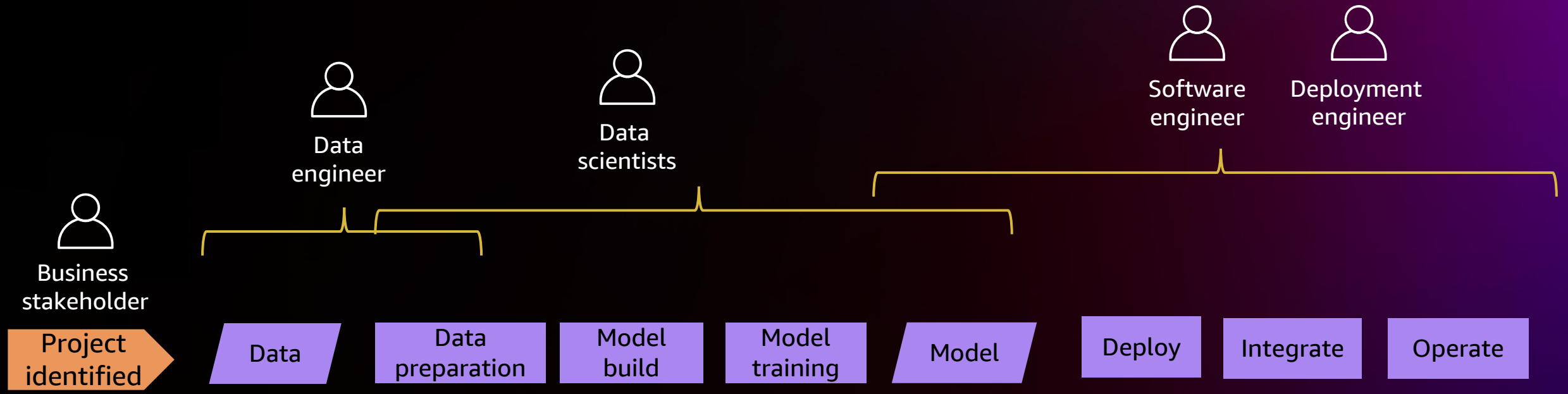
The AWS ML stack

BROADEST AND MOST COMPLETE SET OF MACHINE LEARNING CAPABILITIES



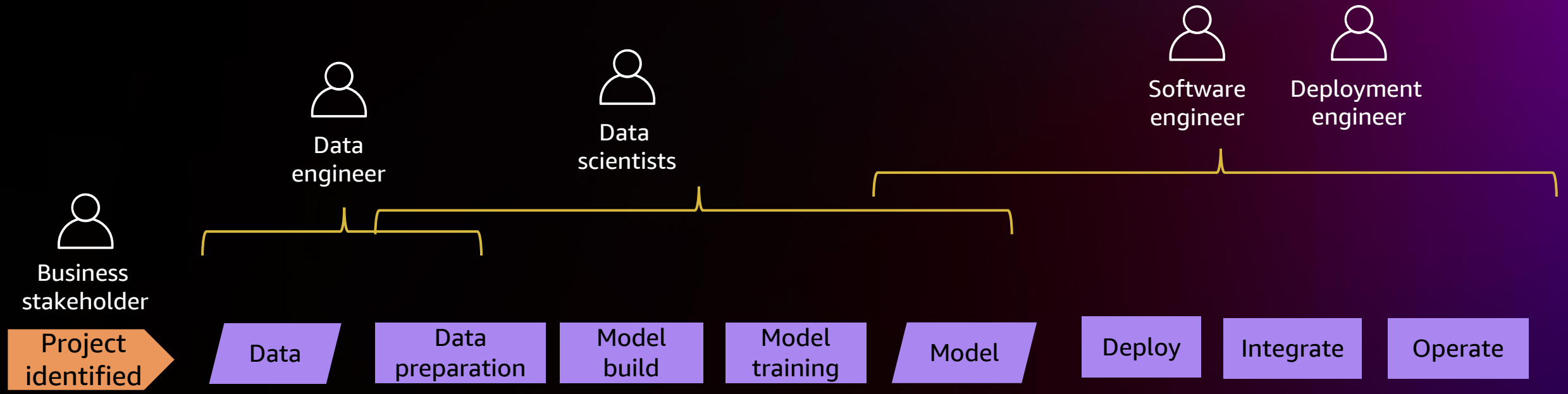
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

First models in production



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

First models in production



Characteristics

- Ad-hoc data strategy
- Manual hand-offs

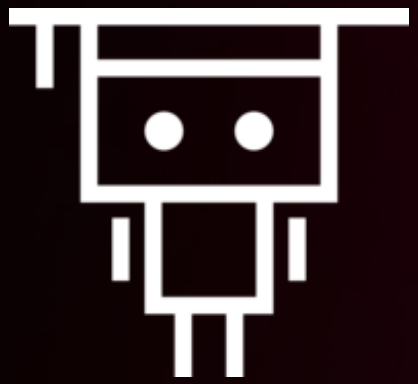
- No/limited automation
- Notebooks
- Manual hand-offs

- Ad-hoc integration
- Monitoring limited to system
- Ad-hoc/manual retraining strategy

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Machine Learning University

Self-service ML training from Amazon's own scientists



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Demo 1

Pet adoption predictor KNN classifier

Predicting the outcome
(0 = not placed, 1 = placed)



aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What is MLOps?

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Think about rotting fruit

I'M SORRY, BUT IT WORKS



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Operational challenges with managing the ML lifecycle

Manual iterative processes slow down ML innovation



Difficult to scale the number of models in production



CI/CD for ML requires writing custom code



Compliance requirements are difficult to meet



aws

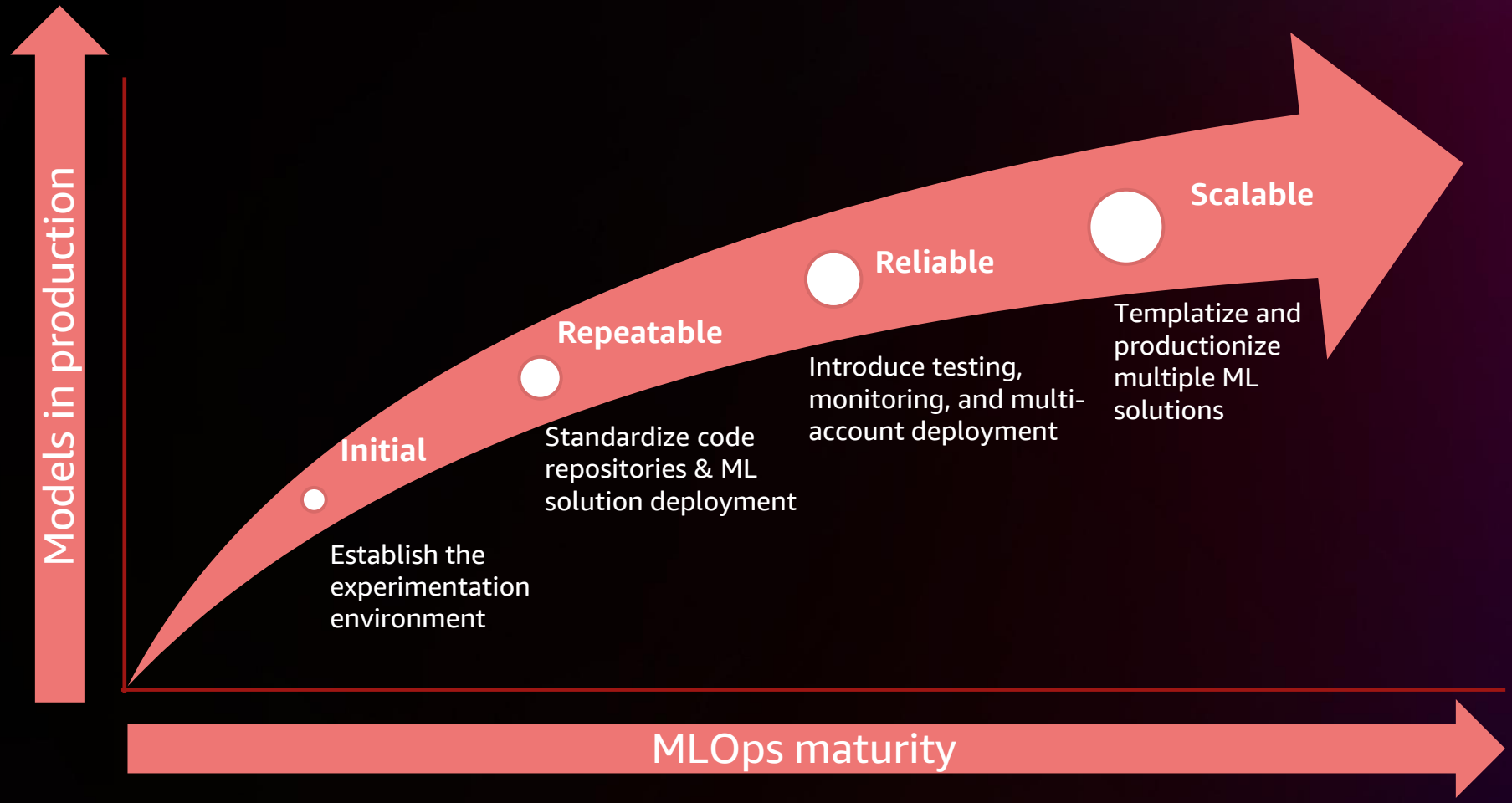
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MLOps key concepts

Flexibility	• Accommodate any framework
Reproducibility	• Need to recreate past experiments
Reusability	• Make ML components reusable
Scalability	• Being able to scale resources on demand
Auditability	• Logs, versions, and dependencies of artifacts
Consistency	• Minimize variance between environments

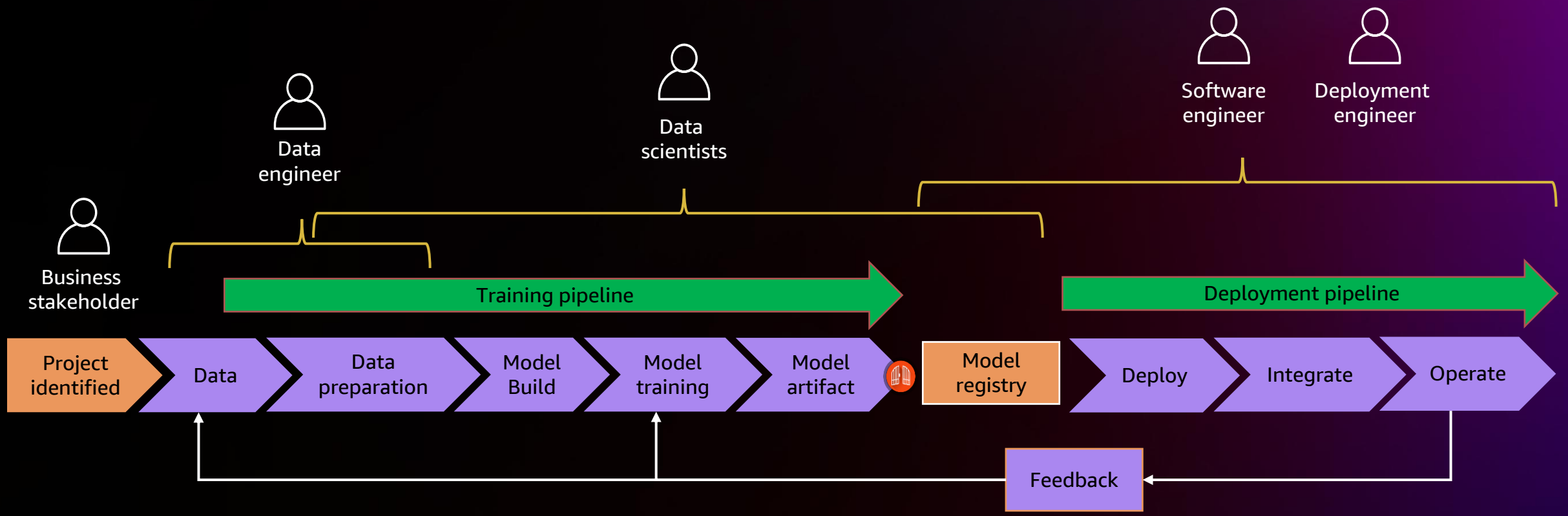
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MLOps maturity model



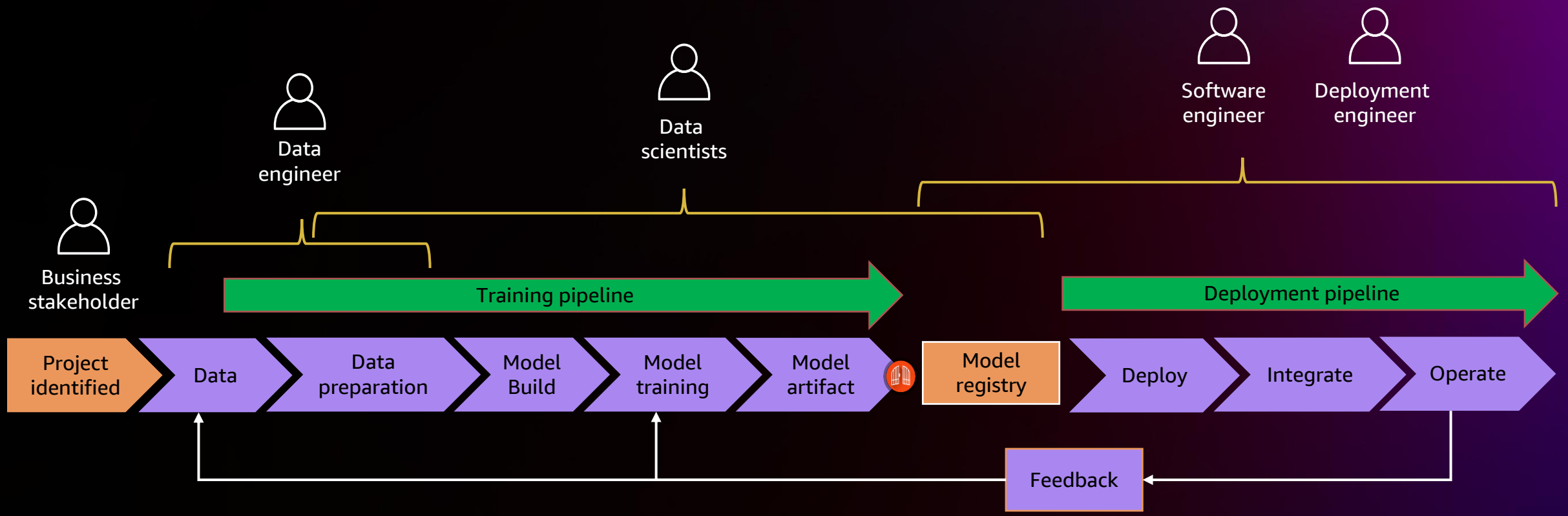
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Increased MLOps maturity



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Increased MLOps maturity



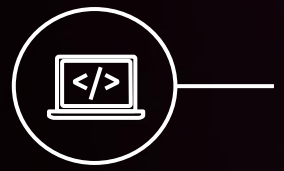
Characteristics

- Pipeline-as-code
- Automated training
- Continuous integration
- Experiment management
- Automated quality gate
- Model monitoring (data/model drift)
- Model governance
- Continuous delivery
- Advanced deployment strategies (such as A/B)
- Model consumption as a microservice
- Business outcome evaluation

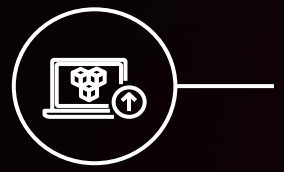
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon SageMaker MLOps

Streamline the ML lifecycle



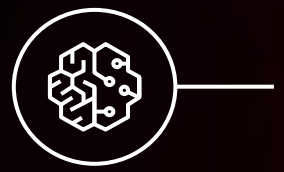
Automate ML workflows to scale model development



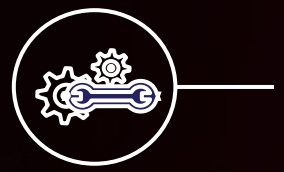
Build CI/CD pipelines for ML to accelerate model deployment



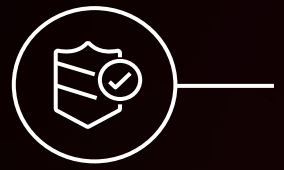
Catalog model versions, metadata, metrics, and approvals for traceability and reusability



Track lineage for troubleshooting and compliance



Maintain accuracy of predictions after models are deployed



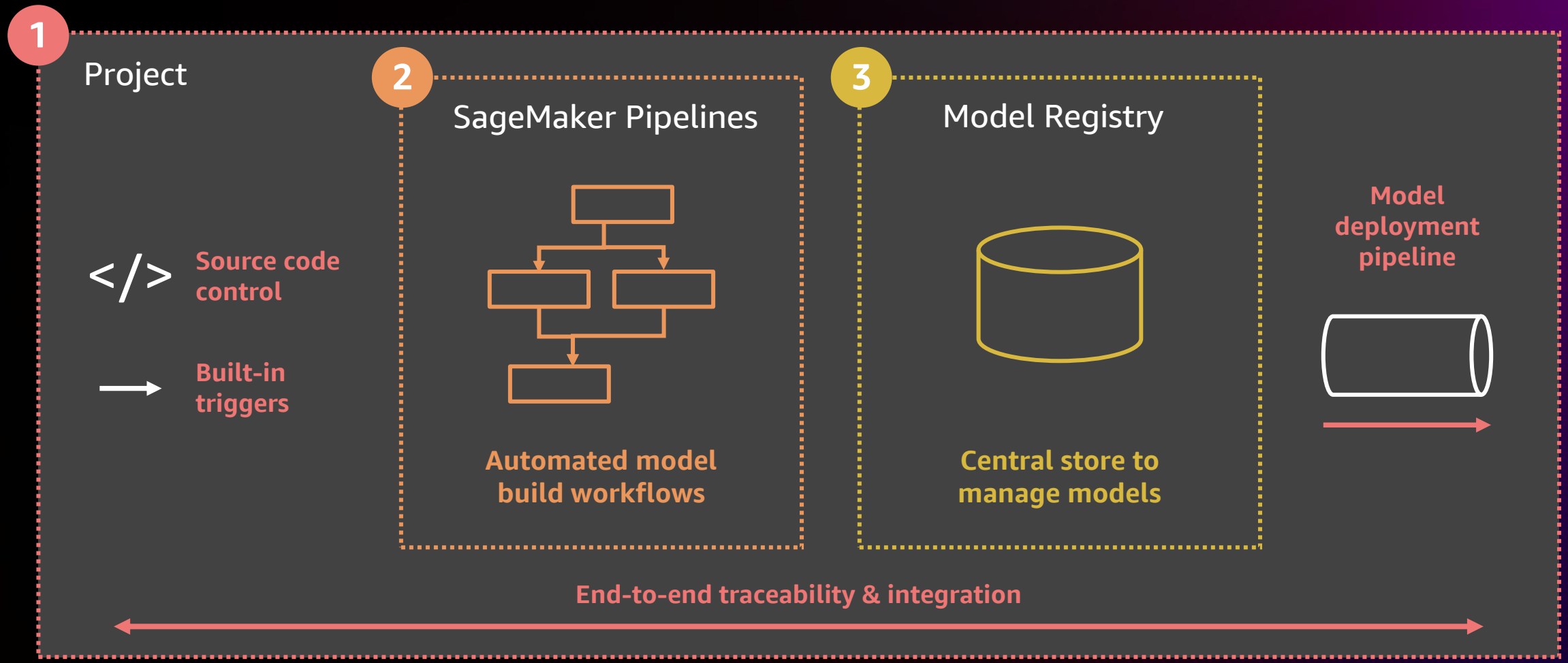
Enhance governance and security

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon SageMaker MLOps

COMPONENTS



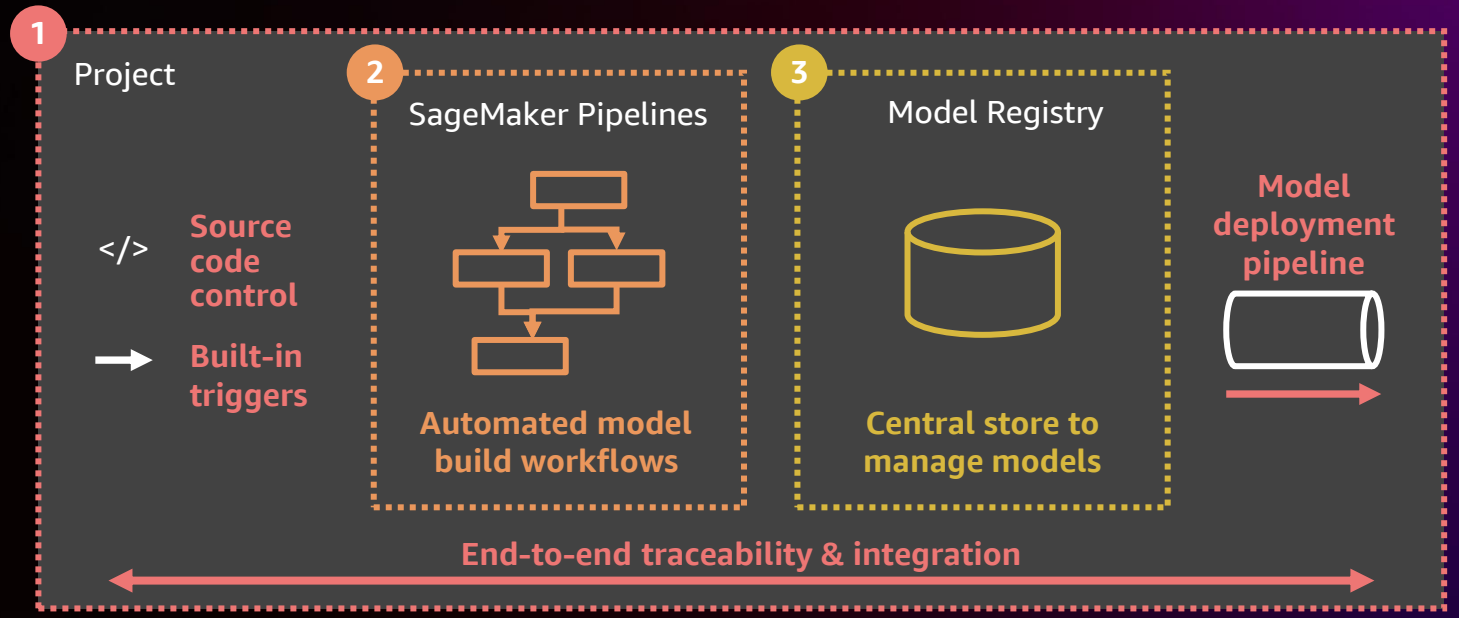
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon SageMaker MLOps

COMPONENTS

Flexible

- ✓ Use projects for end-to-end pipelines that incorporate CI/CD practices
- ✓ Optionally, use SageMaker Pipelines or model registry without projects to meet the needs of your use case



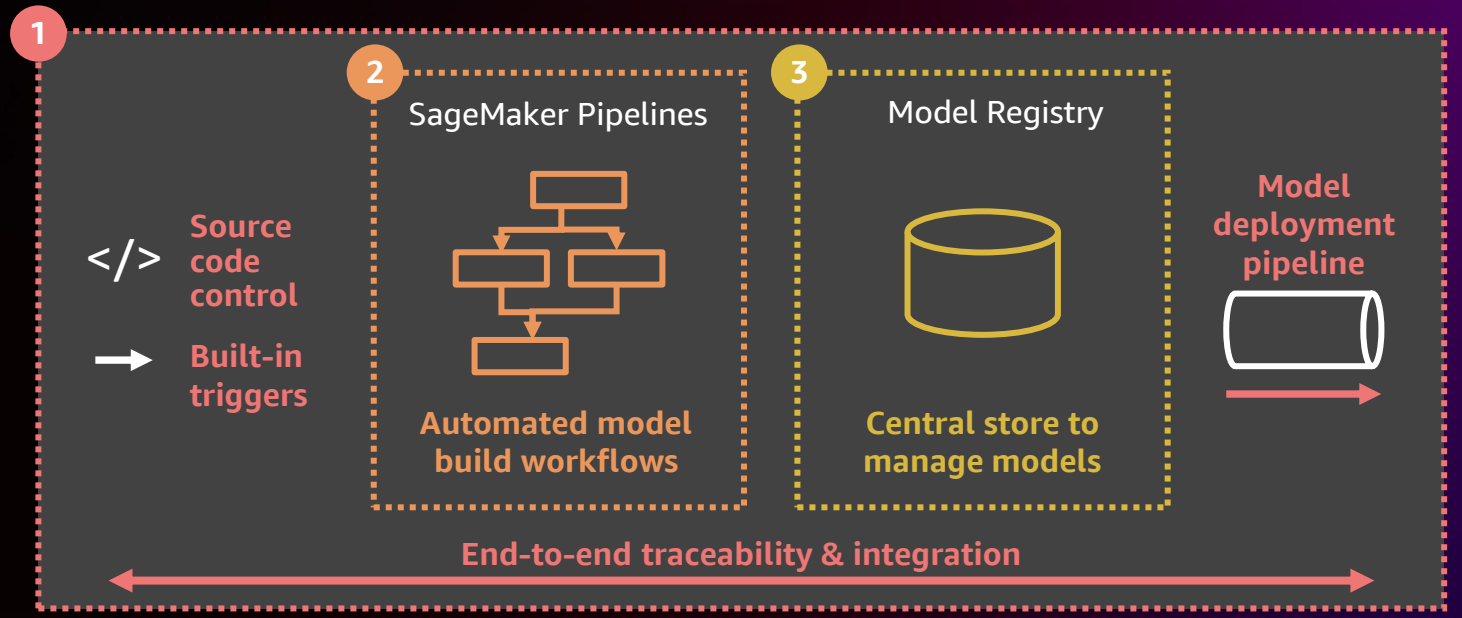
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon SageMaker Pipelines

COMPONENTS

Extensible

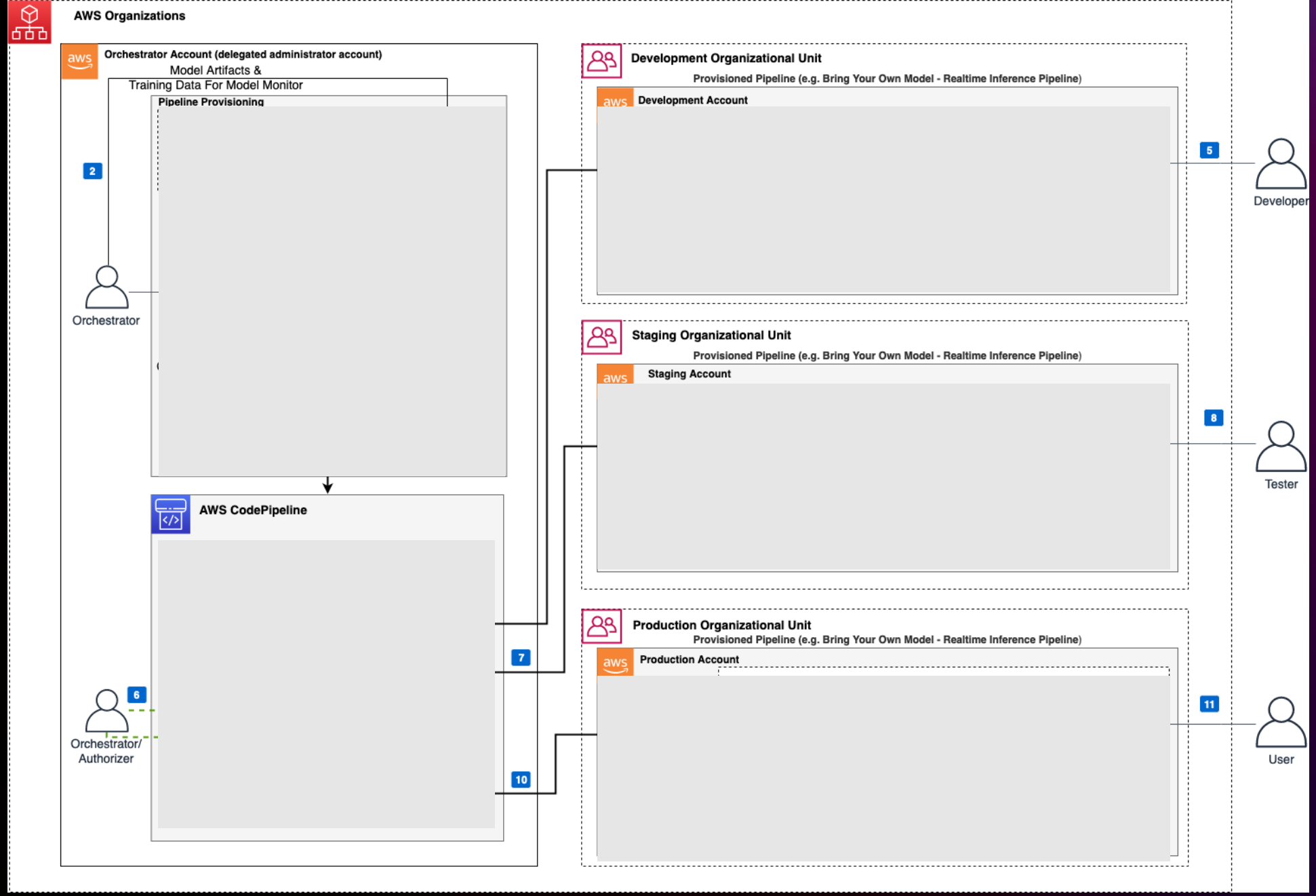
- ✓ Take advantage of built-in MLOps project templates or…
- ✓ Create custom MLOps project templates to meet the unique demands of your corporate or regulatory requirements



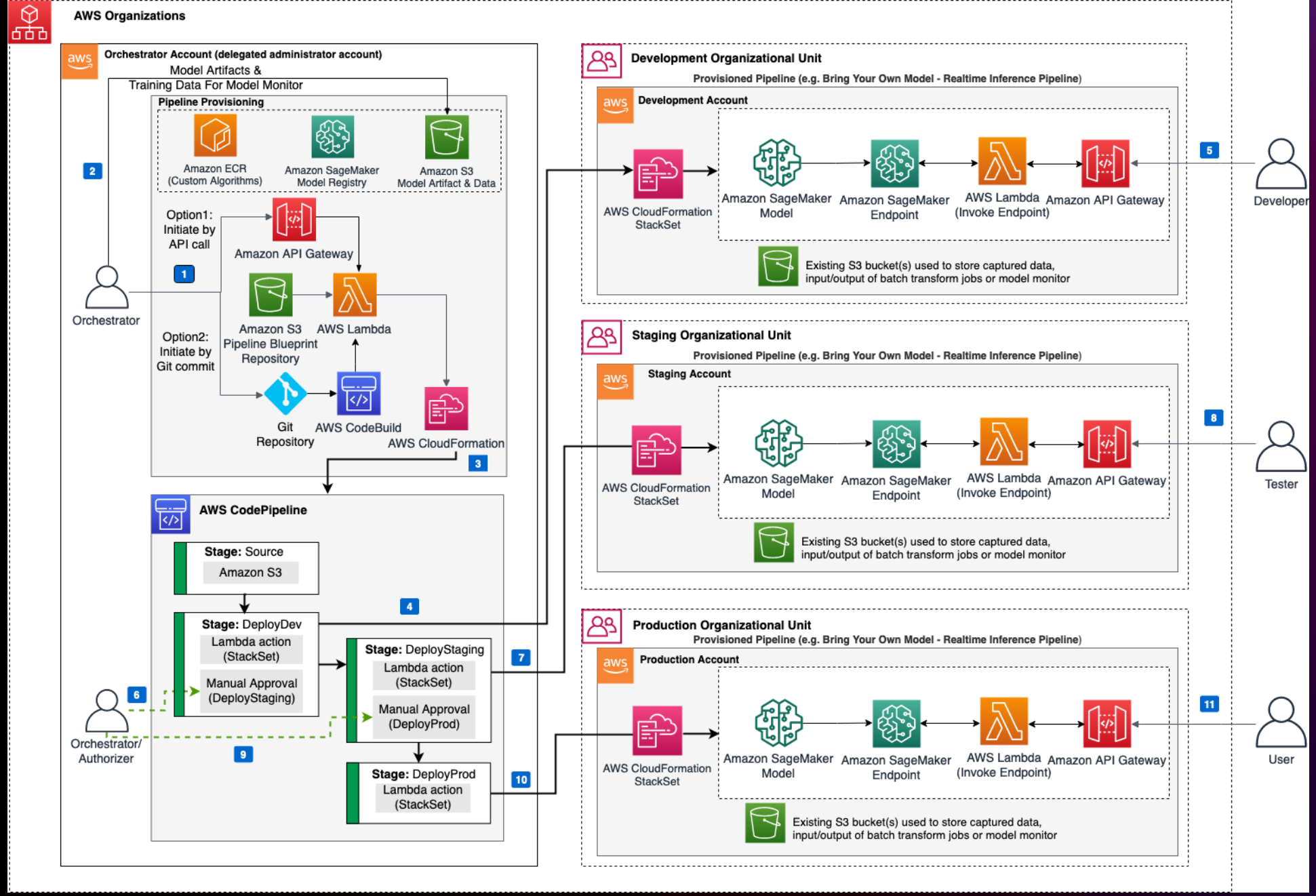
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What does MLOps look like in practice now?

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

If MLOps is so important, why isn’t it more popular?

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Think about rotting fruit

I'M SORRY, BUT IT WORKS



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Is this milk ok to drink?

I'M SORRY BUT IT WORKS



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Regardless of where it starts, it should stop with you

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Brooke Jamieson
linkedin.com/in/brookejamieson
twitter.com/Brooke_Jamieson
tiktok.com/@brookebytes



Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.