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

ANT336

Building data mesh architectures on AWS

Ian Meyers

Director, Product Management
AWS Analytics Service

Nivas Shankar

Principal Product Manager
AWS Lake Formation

Travis Muhlestein

Chief Data & Analytics Officer
GoDaddy

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

Agenda

Modern data strategy on AWS

Customer challenges

Why build a data mesh?

Data mesh – Design goals and four core principles

Data mesh architecture on AWS

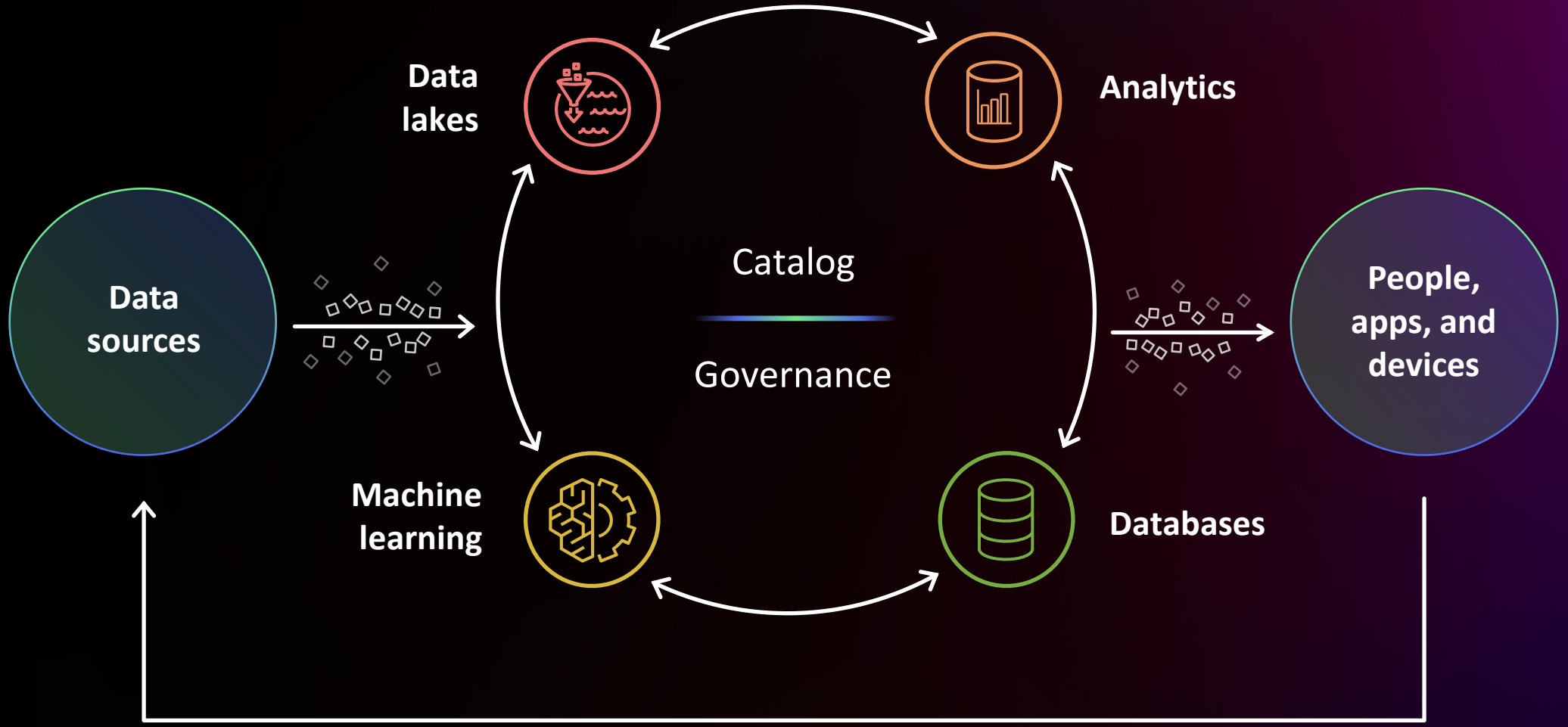
Why GoDaddy chose a data mesh pattern

How GoDaddy built a data mesh using AWS modern data architecture

Conclusion

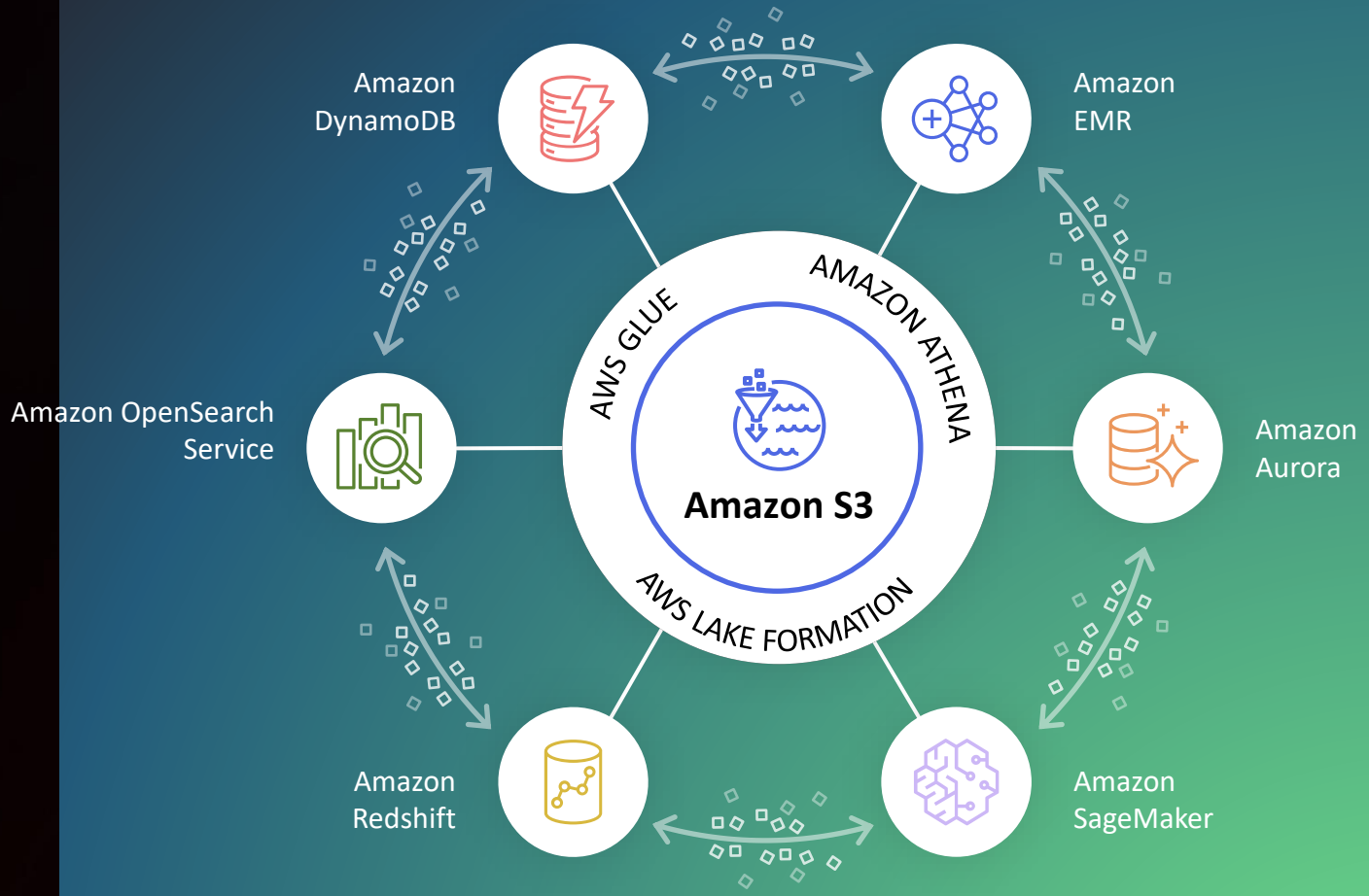
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Modern data strategy



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

Modern data strategy on AWS



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

Key focus areas across all services

Security, durability, availability

Simplicity and ease of use and operations

Price/performance

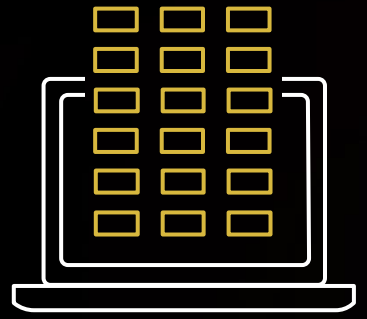
Data connectivity and integration

Data governance

aws

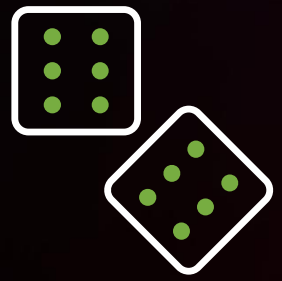
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Customer challenges



MORE SOURCES

Data sources increasing with the pace of the business



MORE BUSINESS UNITS

Data and analytics is table stakes for innovation

All business units are all-in

STRONGER GOVERNANCE

Enterprises need stronger governance to protect their data so that it can be used safely

aws

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

Sharing data in an enterprise can be challenging

Lines of business have made investments in cloud-based data lakes and analytics that are purpose-built to solve their specific business problem

These systems are often unique to the type of data and the algorithms being applied, and don’t always translate to other problems

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

Common data lake challenges we hear from our customers

"Finding the data I need is too difficult"

"I just want to get access to the data I need"

"It's difficult to meet all requirements across differing business units"

"I wish to focus on innovating with data, not on maintaining and administering a data lake"

"My data science team should easily find the datasets they seek and have the ability to share them with others"

"My team needs to own datasets, pipelines, and repositories that are isolated from other teams"

"Why doesn't our organization treat data as a product?"

"Current data architecture is complex and monolithic and slow to change"

"If I share data, I've lost control"

"There is a mismatch between executive leadership goals & business line deliverables, and incentives"

"Our internal policies on what can be shared is unclear & there is lack of incentive to share"

"I need to create a model to support sharing from both producers and consumers of data"

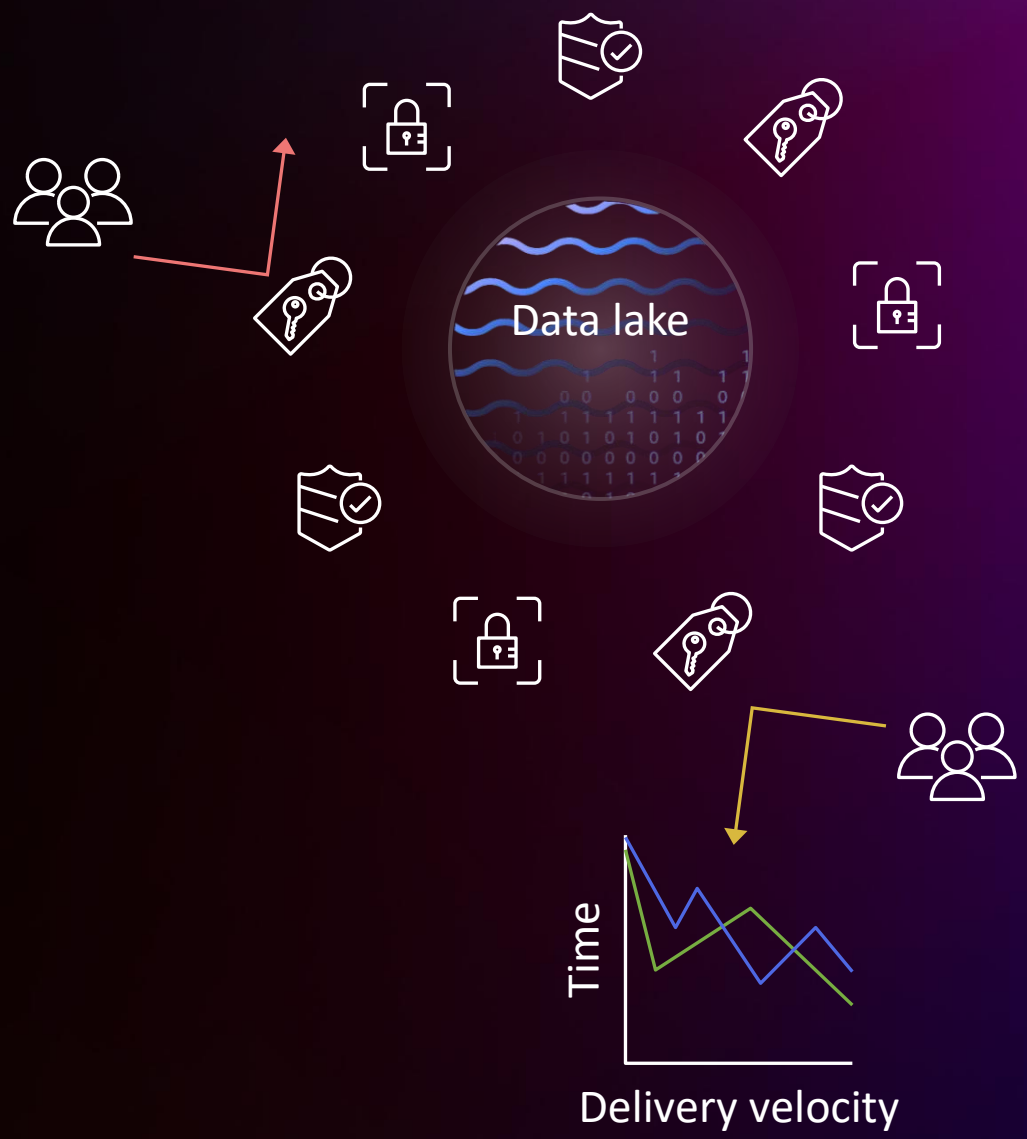
aws

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

Sharing data in an enterprise can be challenging

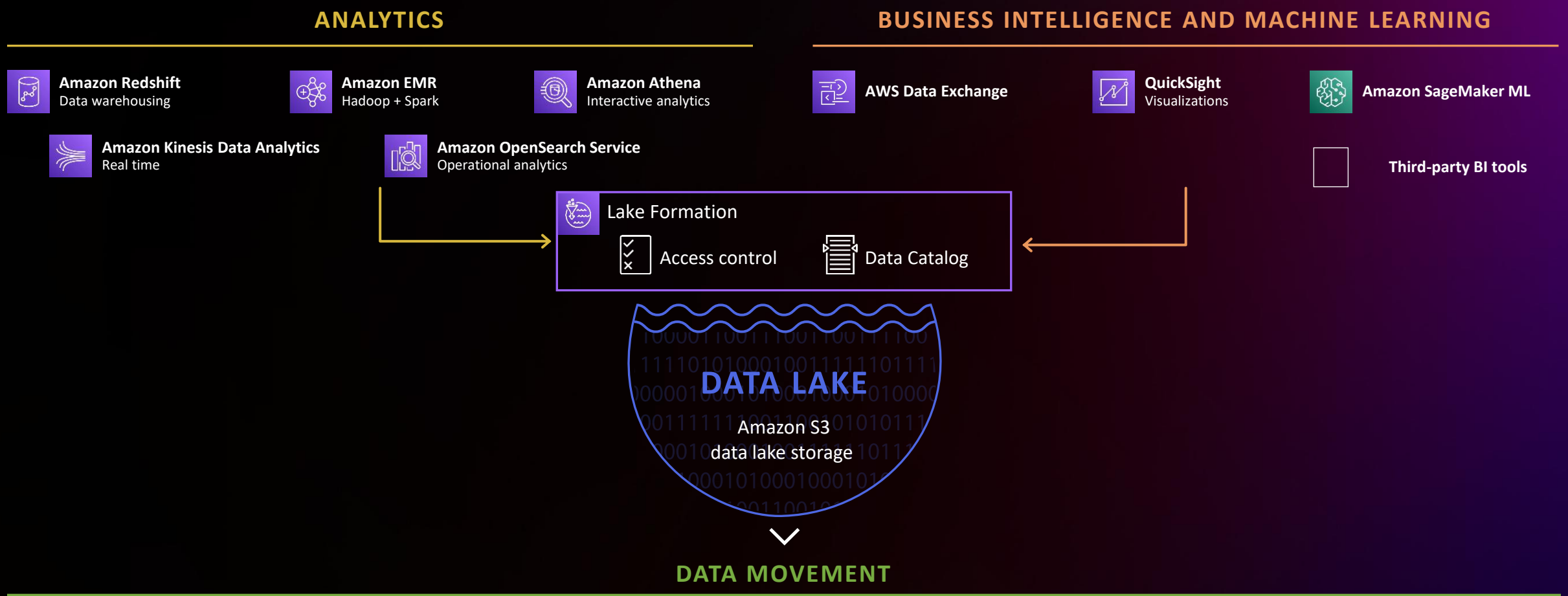
It may seem that the only way to effectively govern data is through a single, centralized architecture and technology

This can increase friction to adopt and reduce velocity of delivering business outcomes



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

Centralized data lakes are complex to scale across business units



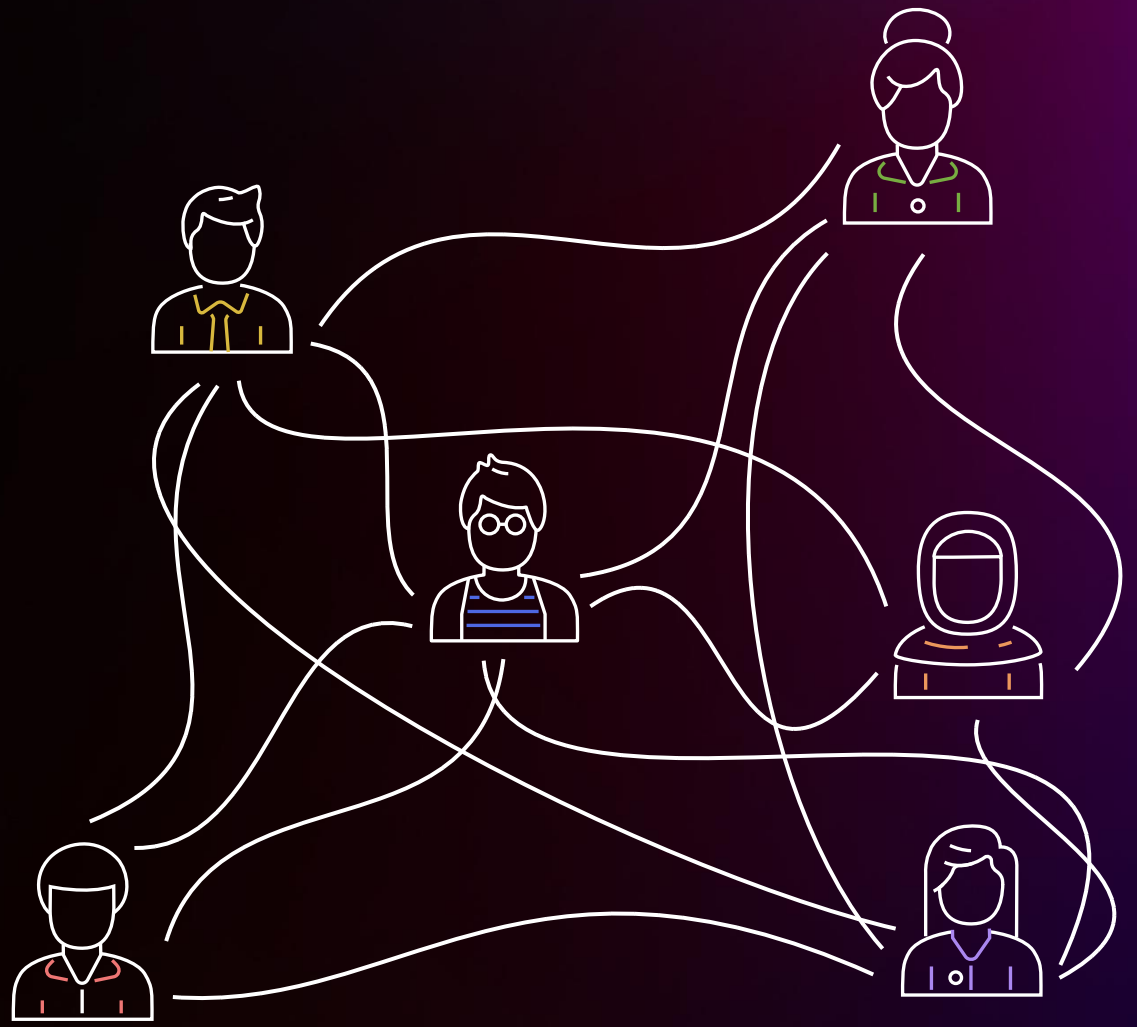
AWS Database Migration Service | AWS Snowball | AWS Snowmobile | Amazon Kinesis Data Firehose | Amazon Kinesis Data Streams Managed Streaming for Kafka

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

Sharing data in an enterprise can be challenging

Often data is being shared (and maybe reshared), but in an ad hoc, ungoverned way based on team connections

This can increase risk associated with protecting sensitive data

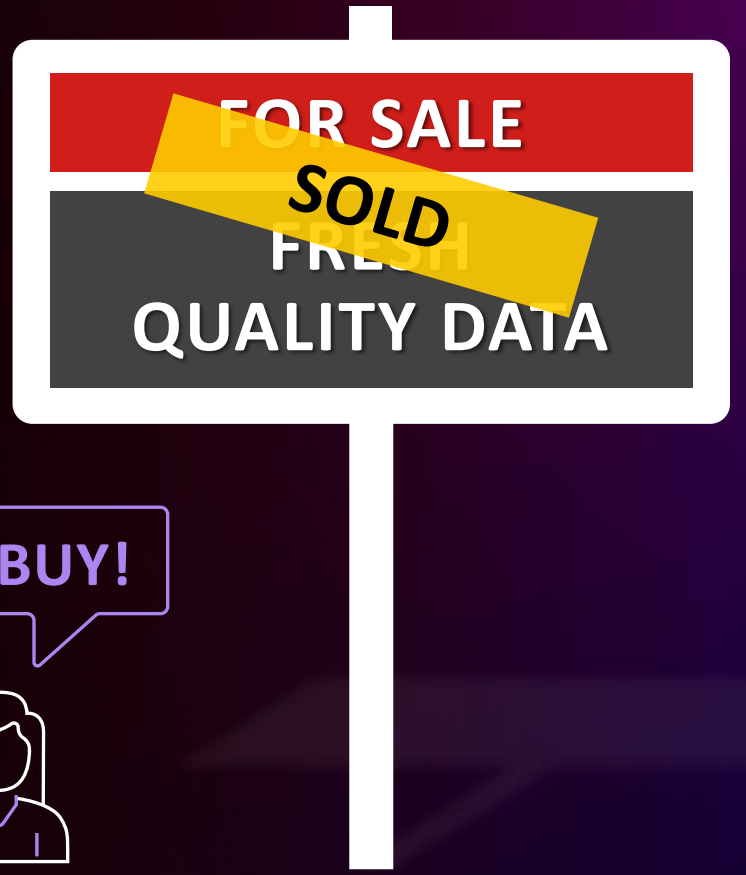


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Sharing data in an enterprise can be challenging

ORGANIZATIONAL INCENTIVES AROUND DATA SHARING CAN BE MISALIGNED

"Everyone wants to be a consumer, no one wants to be a producer"



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

Why data mesh?

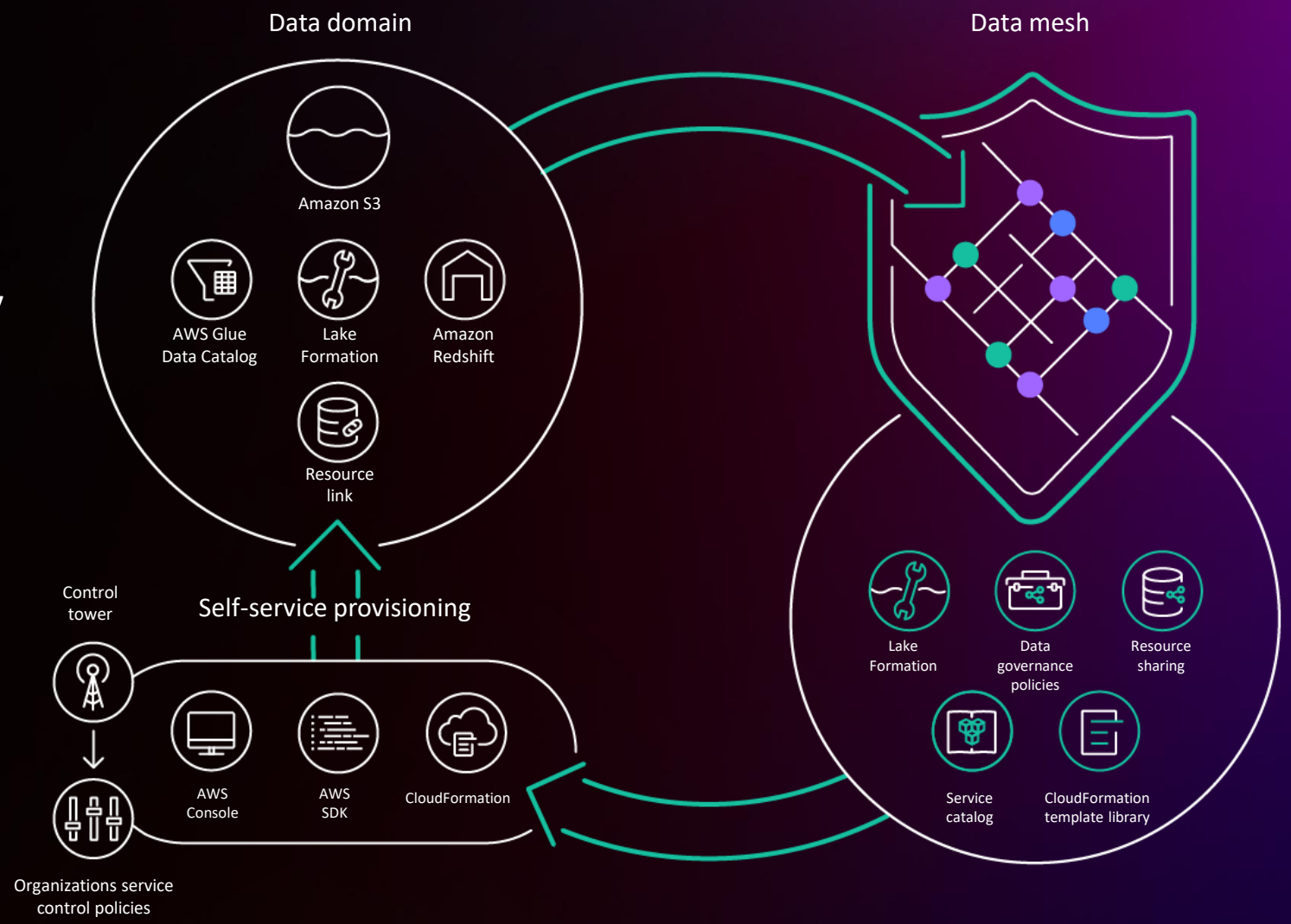
Use existing investments in data platforms and treat them as independent “domains”

Improve data governance by pushing policy down into data domains

Provides a clear mechanism for centralized data discovery

Provides self-service data sharing features to allow domain owners to grant access to consumers

Measure and invest in data products based on usage and business value



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

Data mesh: Design goals and four key principles

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

Data mesh: Four core principles

Data owner

Data domain ownership

A data mesh features data domains as nodes, which exist in data lake accounts; it is founded in decentralization and distribution of data responsibility to people closest to the data

Data engineer

Data as a product

A data producer contributes one or more data products to a central catalog in a data mesh account; DaaP must be autonomous, discoverable, secure, and correct, and useable

Data steward

Federated computational governance

Federated data governance is how data products are shared – delivering discoverable metadata auditability based on federated decision-making and accountability structures

Data consumer

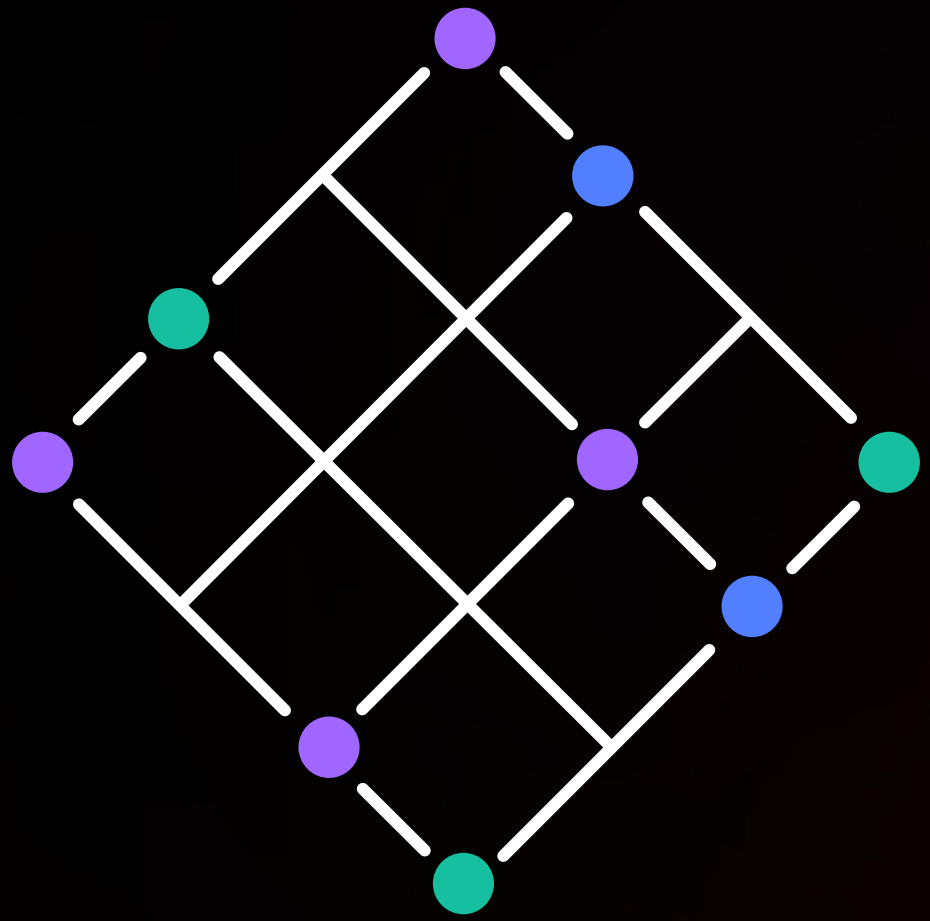
Self-serve sharing

The platform streamlines the experience of data users to discover, access, and use data products; it streamlines the experience of data providers to build, deploy, and maintain data products

aws

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

Data mesh design goals



Enable organizations to get value from data at scale

Create a business-oriented data products that can support the top strategic goals

Allow business domains federated governance through lightweight
centralized policy by removing bottlenecks

Encourage data-driven agility

Support the sharing of data products, with the goal of delighting the experience of data users

aws

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

Modern data architecture consists of five parts

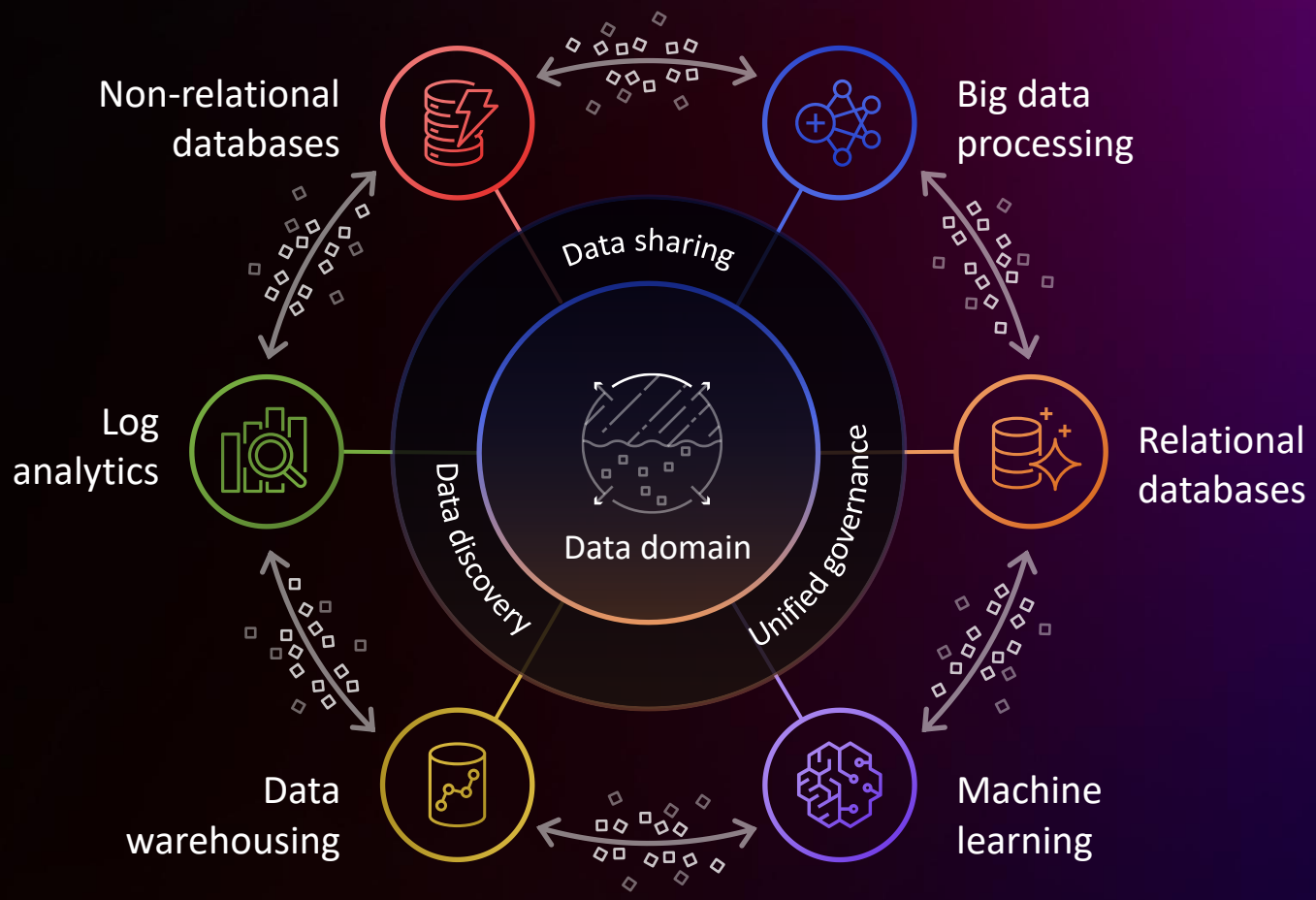
Scalable data lake and warehouses

Purpose-built data services

Data sharing

Unified governance

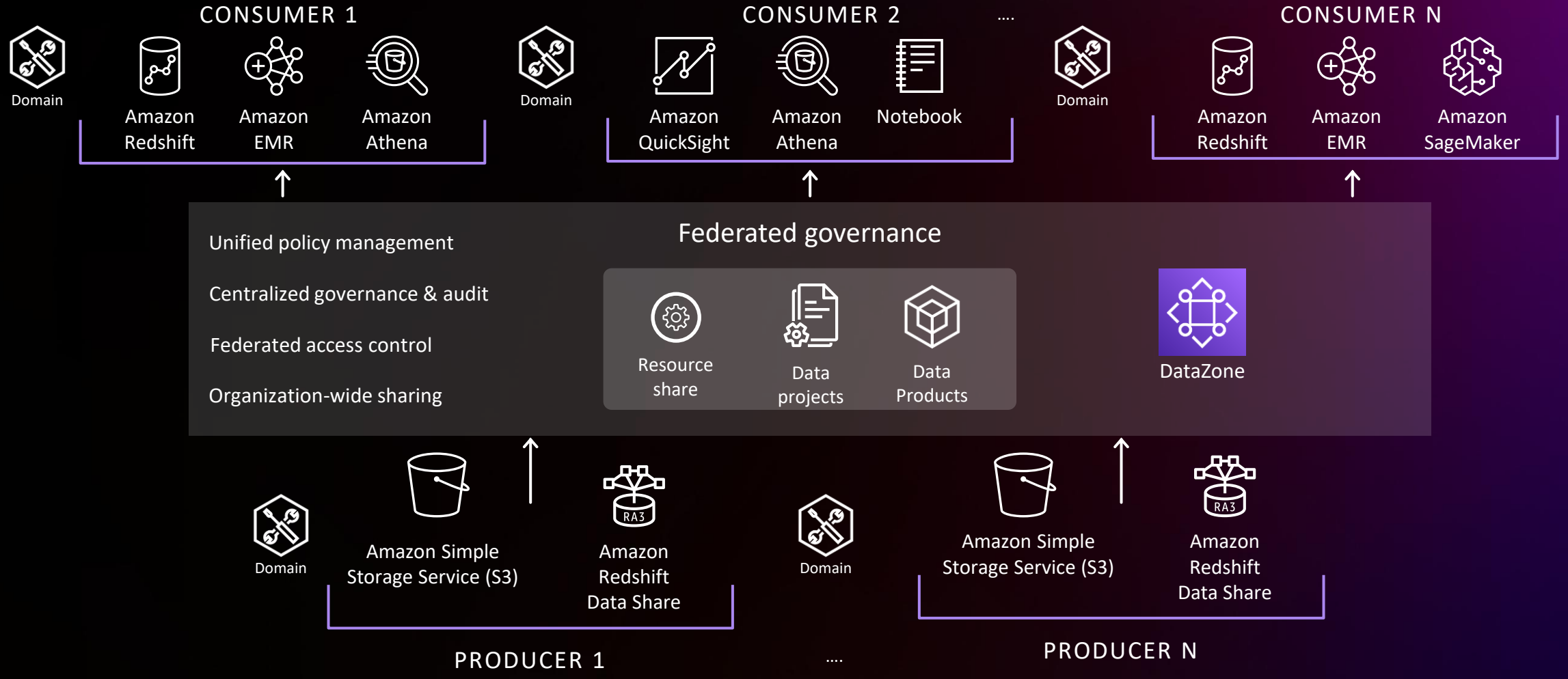
Data discovery



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

Data mesh architecture

DECENTRALIZED, LIGHTWEIGHT FEDERATED GOVERNANCE ACROSS DOMAIN-ORIENTED DATA SYSTEMS TO DRIVE GOVERNED SHARING



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

Data mesh principle #1: Data domain ownership

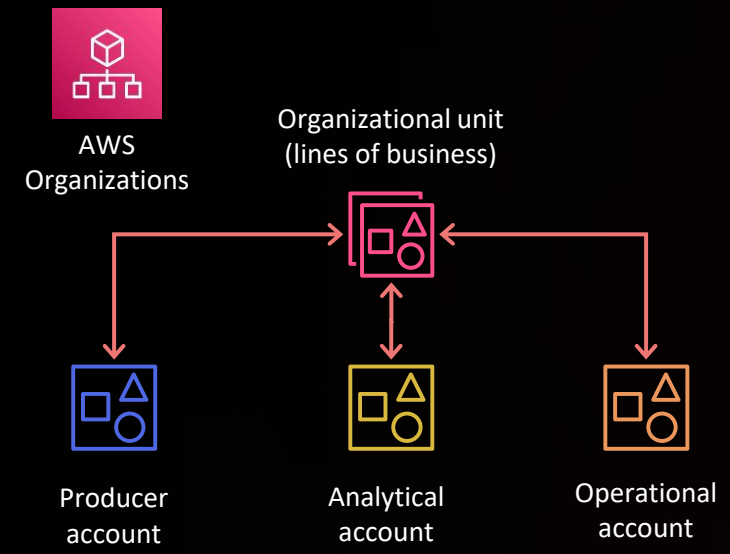
DATA OWNERS ARE ACCOUNTABLE FOR THEIR DATA PRODUCTS TO BE RELIABLE, AVAILABLE, AND ACCURATE

Data owner

Accountability for data domain and consumption of data products

Create

Owns data **created** or **transformed** in my organization



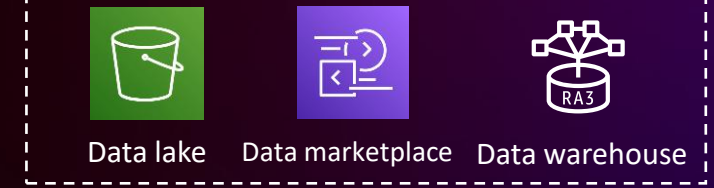
Store

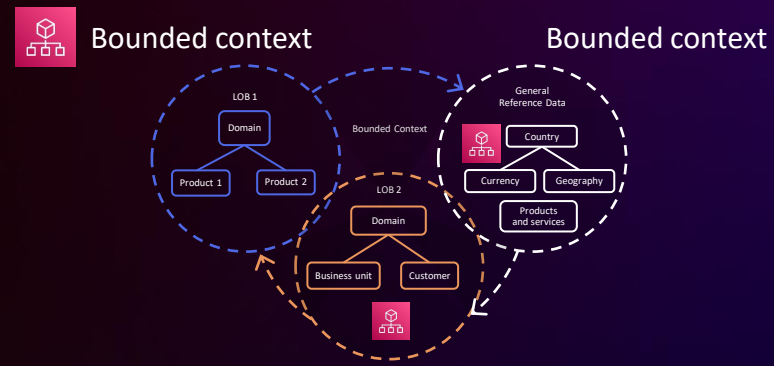
Owns data **stored** in my organization

- Register and catalog data assets
- Protect and secure organization data
- Manage data quality
- Maintain ease of use and atomic integrity

Share

Owns how my organization **uses** the data

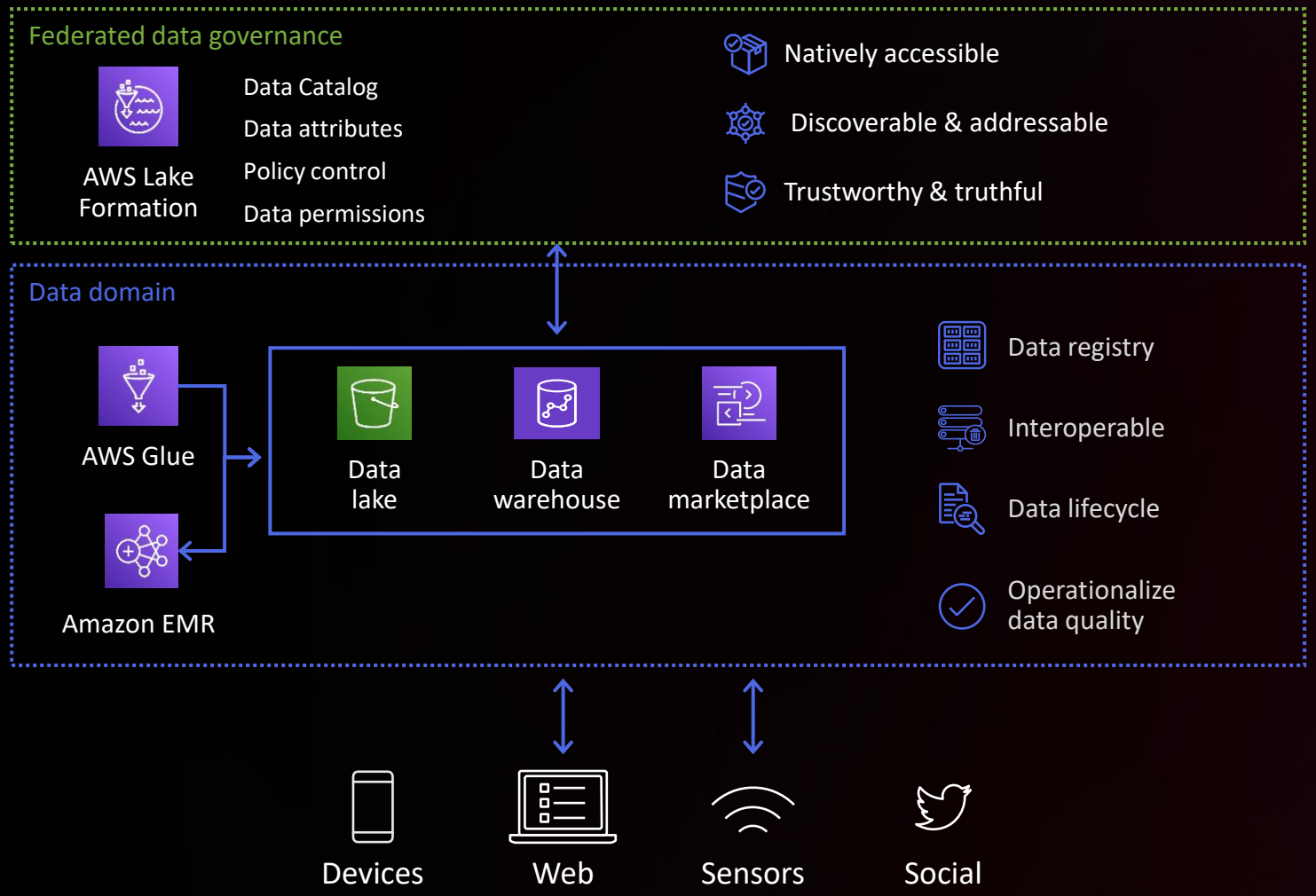




© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh principle #2: Data as a product

DOMAIN-DRIVEN DESIGN TECHNIQUES TO FORMULATE AND ESTABLISH BOUNDED CONTEXTS FOR DATA PRODUCTS



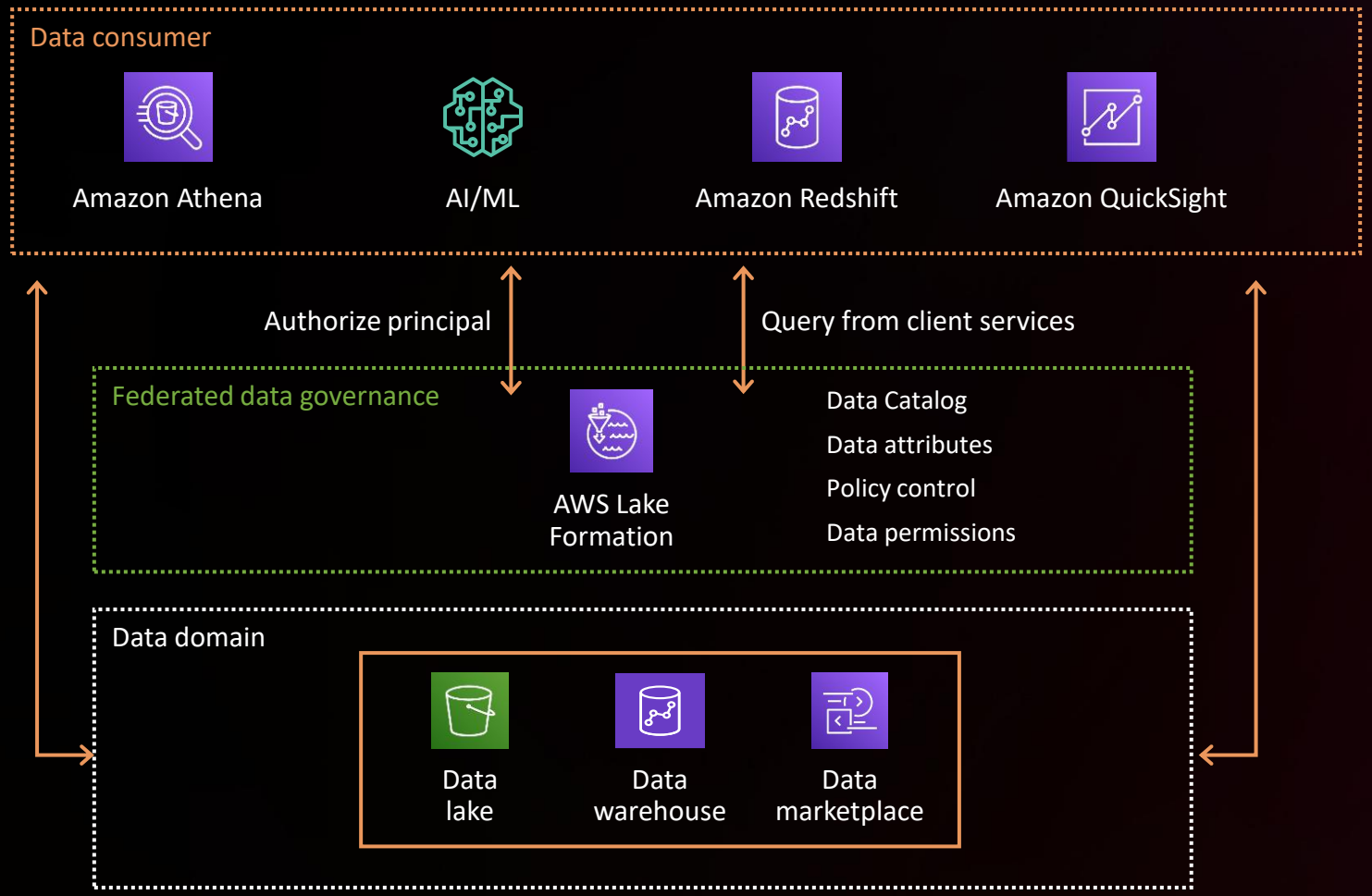
Data engineer

- Each team manages their data and organizes it as data products
- Each product provides an interface(s) to allow others to interact (e.g., APIs, SQL, reports)
- Remove usability frictions, meet the user where they are
- Provide all supporting metadata, lineage
- Data products are valuable on their own

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

Data mesh principle #3: Self-serve sharing

ECOSYSTEM OF SELF-SERVE DATA INFRASTRUCTURE WITH OPEN PROTOCOLS



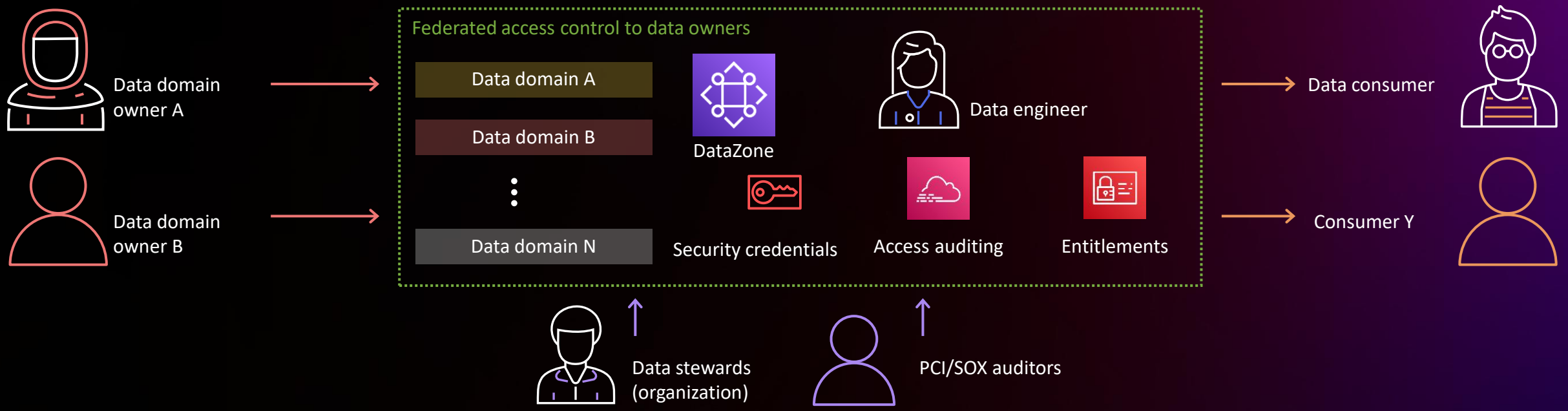
Data consumer

- Design for generalist majority (i.e., make it easy to use and adopt with no specialist skills needed)
- Enable personas to discover, learn, understand, consume, and maintain data products
- Collection of interoperable data products, which enable cross-functional domains to produce and consume data easily and with autonomy and will allow it to scale
- Data products must include data, metadata, code, and policy all as single unit of value
- Abstract complexity through automation

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

Data mesh principle #4: Federated data governance

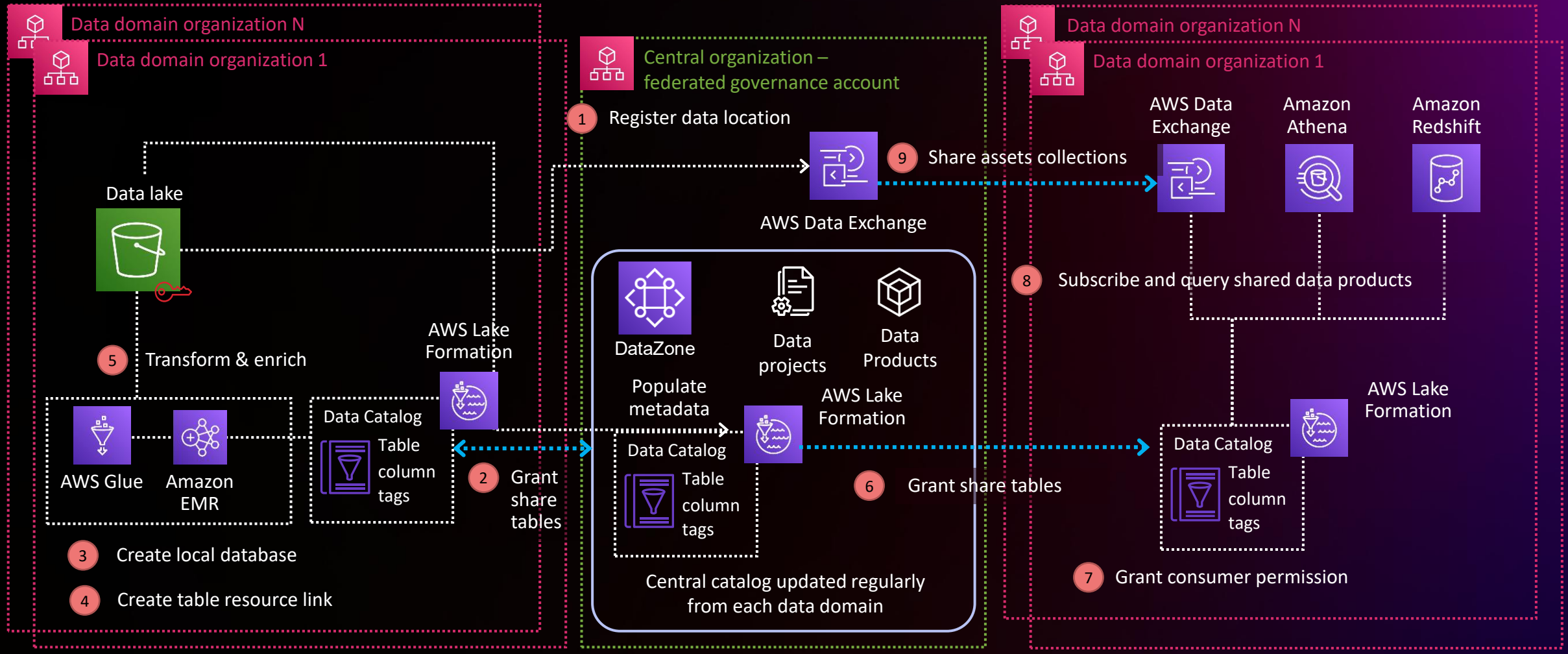
GOVERNANCE MODEL THAT EMBRACES DECENTRALIZATION AND DOMAIN SELF-SOVEREIGNTY THROUGH DECISION-MAKING MODEL LED BY FEDERATION OF DATA PRODUCT OWNERS



- Decentralization implementation of governance team and standard authorization
- Governance team = a guild consisting of representatives of all teams taking part in the data mesh
- They create global policies and standardization to achieve interoperability
- Automated execution of policies by the data domains (e.g., data classification and privacy, compliance, security, documentation, and interoperability)

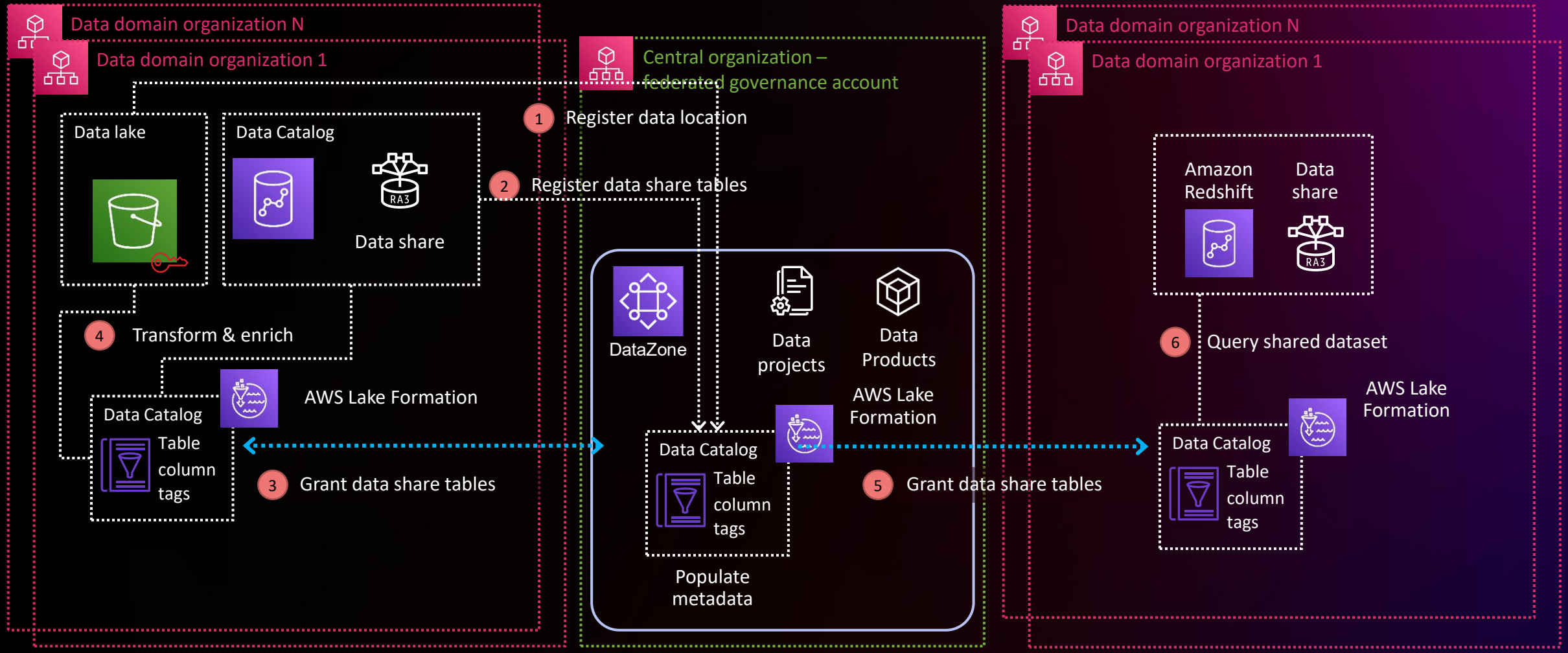
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh architecture pattern: Data lake data products sharing



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

Data mesh architecture pattern: Data lake and data warehouse data product sharing



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

GoDaddy’s journey with AWS

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

Our vision

is to radically shift the global economy toward life-fulfilling entrepreneurial ventures

Our mission

is to empower entrepreneurs everywhere, making opportunity more inclusive for all

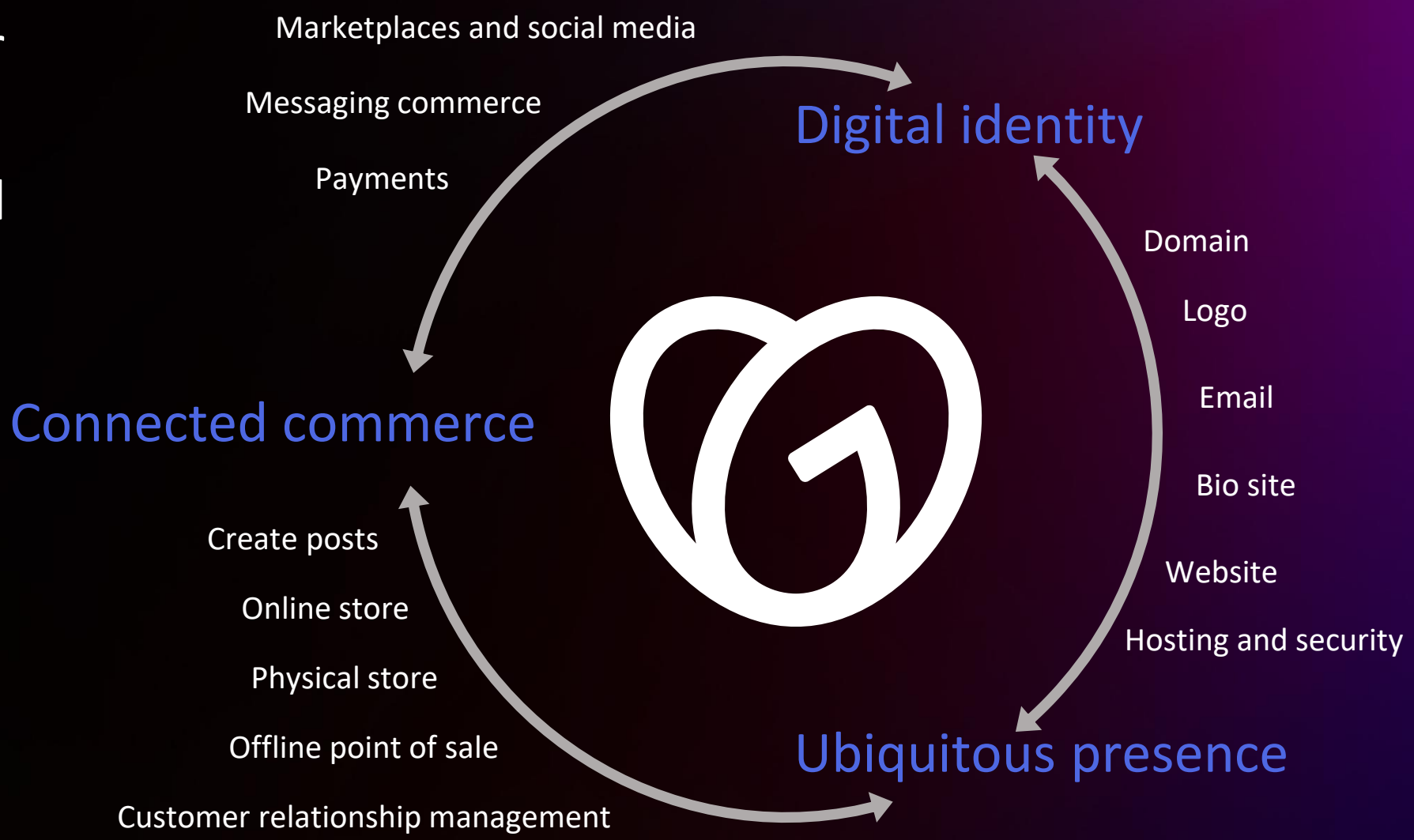
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Our strategy

We **champion** everyday entrepreneurs by empowering them with **sage guidance** set in **seamlessly intuitive experiences** to securely name, create, and grow their ventures in select markets; leveraging the **exponential power of our community** at global scale to deliver profitable revenue growth

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

At GoDaddy our goal is to partner with our customers at every point on this wheel



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

The largest domain registrar in the world

One of the world's top WordPress hosting brands

One of the top-branded email providers in the world

21M+ customers

84M+ domains

15M+ human-guided moments*

85% customer retention

12% website share

65+ Care net promoter score (NPS)

55 global sites

10M+ mailboxes

$26B gross merchandise volume (GMV)

aws

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

GoDaddy

GoDaddy & AWS – 5 years of strategic collaboration

2018

GoDaddy selects AWS for its scale, performance, and acceleration

App services established & first TLZs

27 teams onboarded to AWS (on goal of 7)

2019

Public cloud portal launches

130+ teams and 800 developers migrate to AWS

100% new product development on AWS (7/2019)

2020

GoDaddy on-premises Hadoop migration begins

150 teams & 97 prod workloads on AWS

Amazon Prime collaboration with websites and marketing

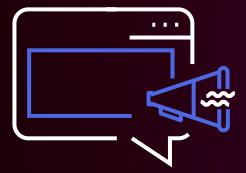
2021

Neustar migration to AWS

Amazon registrar migrates 113K domains to GD

MAIT Team established

Security scans, golden AMIs, node rotation



2022

Websites & marketing migration begins

Migrated from Hadoop to AWS

300+ teams, 165 prod workloads, and 651 projects on AWS

2023

Amazon registrar plans to migrate 1 million domains to GoDaddy

Migration of 100K customer website builder sites into AWS

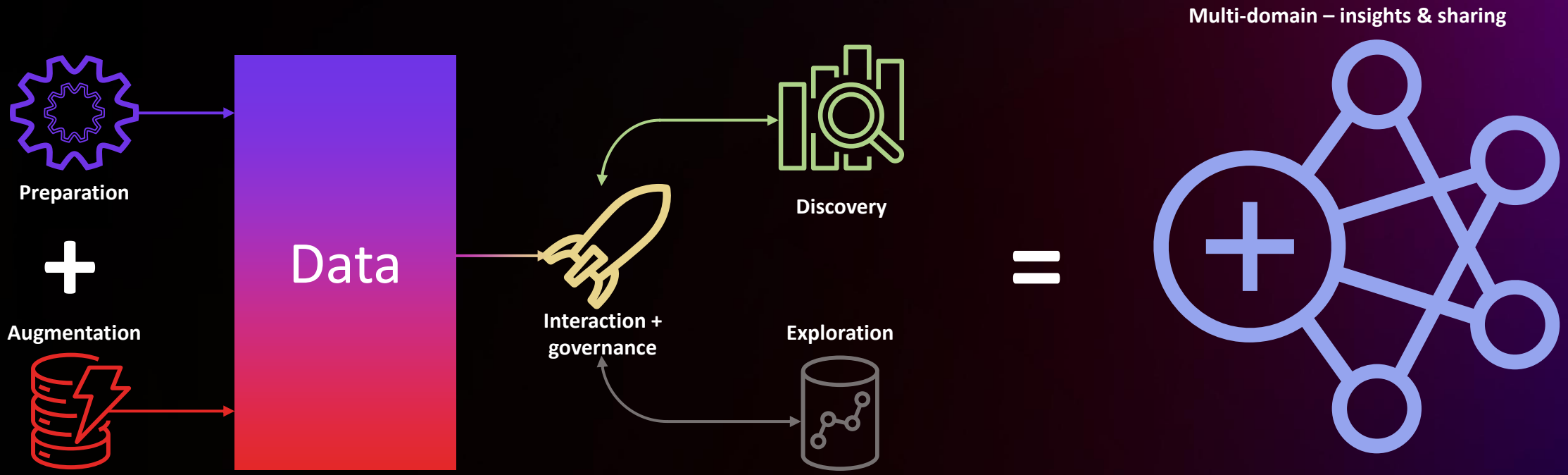
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Why GoDaddy chose a data mesh pattern

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

Data mesh:

Can a data mesh create customer clarity?

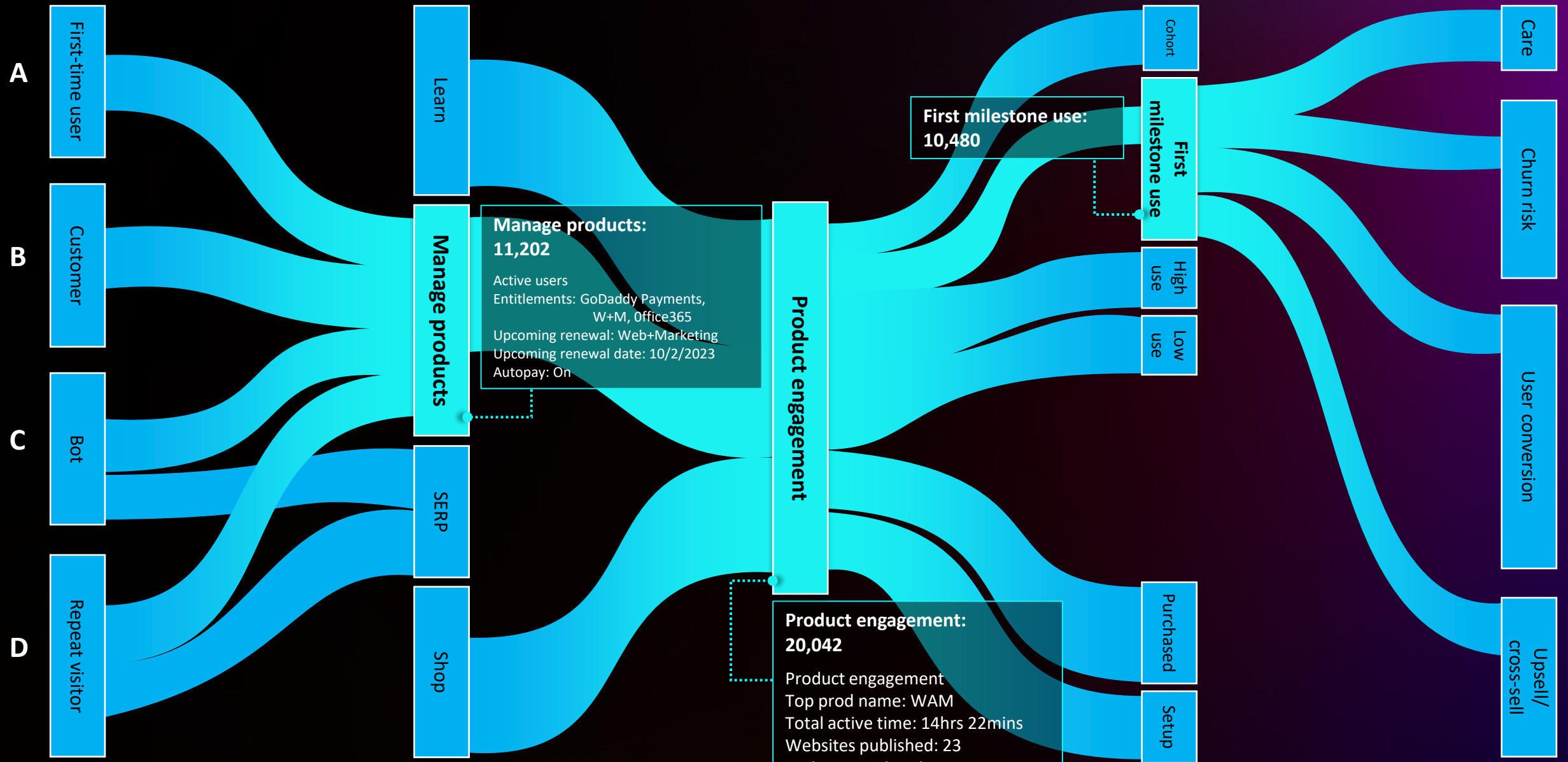


Data consumer

Data discovery | Unified data experience | Data sharing

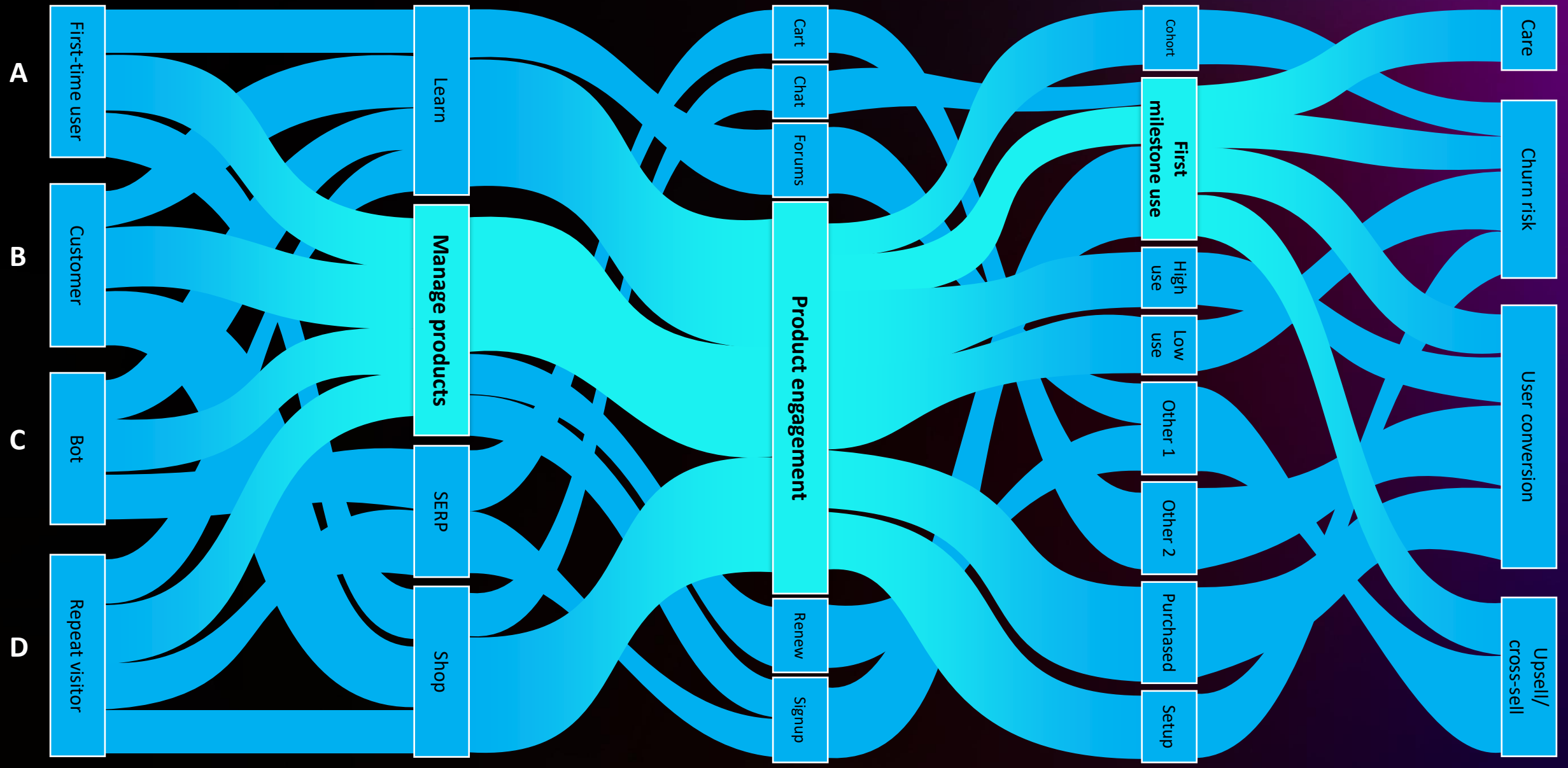
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Multi-domain federated customer actions



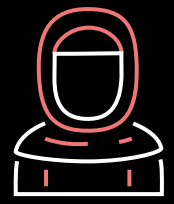
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Multi-domain federated customer actions



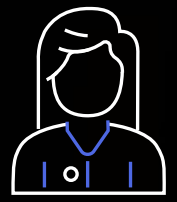
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GoDaddy data mesh – Customer layers



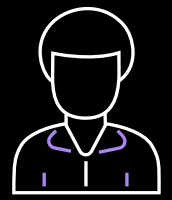
Data owner

Data domain ownership



Data engineer

Data as a product



Data steward

Data governance council

Data consumer

Self-serve sharing

CUSTOMER TIERS

Multi-domain products

Visiting user (tier 0)

Prospective customer (tier 1)

Highly or lowly engaged (tier 2)

Conversion (tier 3)

Account (tier 4)

High-value account (tier 5)

aws

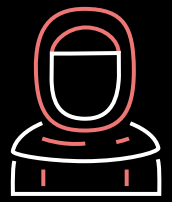
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GoDaddy

How GoDaddy built a data mesh using AWS modern data architecture

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

Data mesh



Data owner

Data domain ownership

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

Data mesh

Data consumer

Self-serve sharing

Producer domains

Systems of record

1st party
3rd party
Services
Products

Self-service*

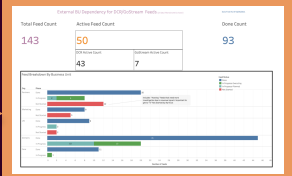
Utility domains

Data ingress

S3

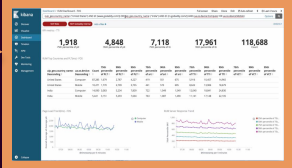
Data egress

Consumer domains



Data products

Business insights

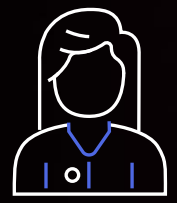


Self-service*

AWS native and third-party services

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

Data mesh



Data engineer

Data as a product

Producer domains

Systems of record

1st party
3rd party
Services
Products

Self-service*

Utility domains

Data ingress

Data processing
(streaming & batch)

S3

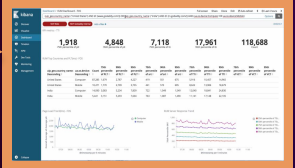
Data lake
(storage & metadata)

Data egress

Consumer domains

Data products

Business insights



Self-service*

Amazon Redshift

Amazon EMR

Amazon Athena

AWS native and third-party services
Spark, Tecton, Tableau Cloud,
Alation Searchability (Catalog)

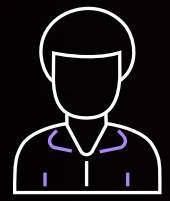
SageMaker

AWS Glue

AWS Lake Formation

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

Data mesh



Data steward

Data governance council

Producer domains

Systems of record

1st party
3rd party
Services
Products

Self-service*

Utility domains

Data ingress

Data processing
(streaming & batch)

S3

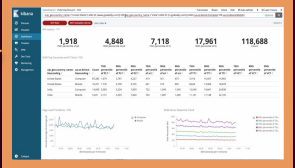
Data lake
(storage & metadata)

Data egress

Consumer domains

Data products

Business insights



Self-service*

Enhanced governance

Data governance* and the data interfaces

Amazon Redshift

Amazon EMR

Amazon Athena

AWS native and third-party services
Spark, Tecton, Tableau Cloud,
Alation Searchability (Catalog)

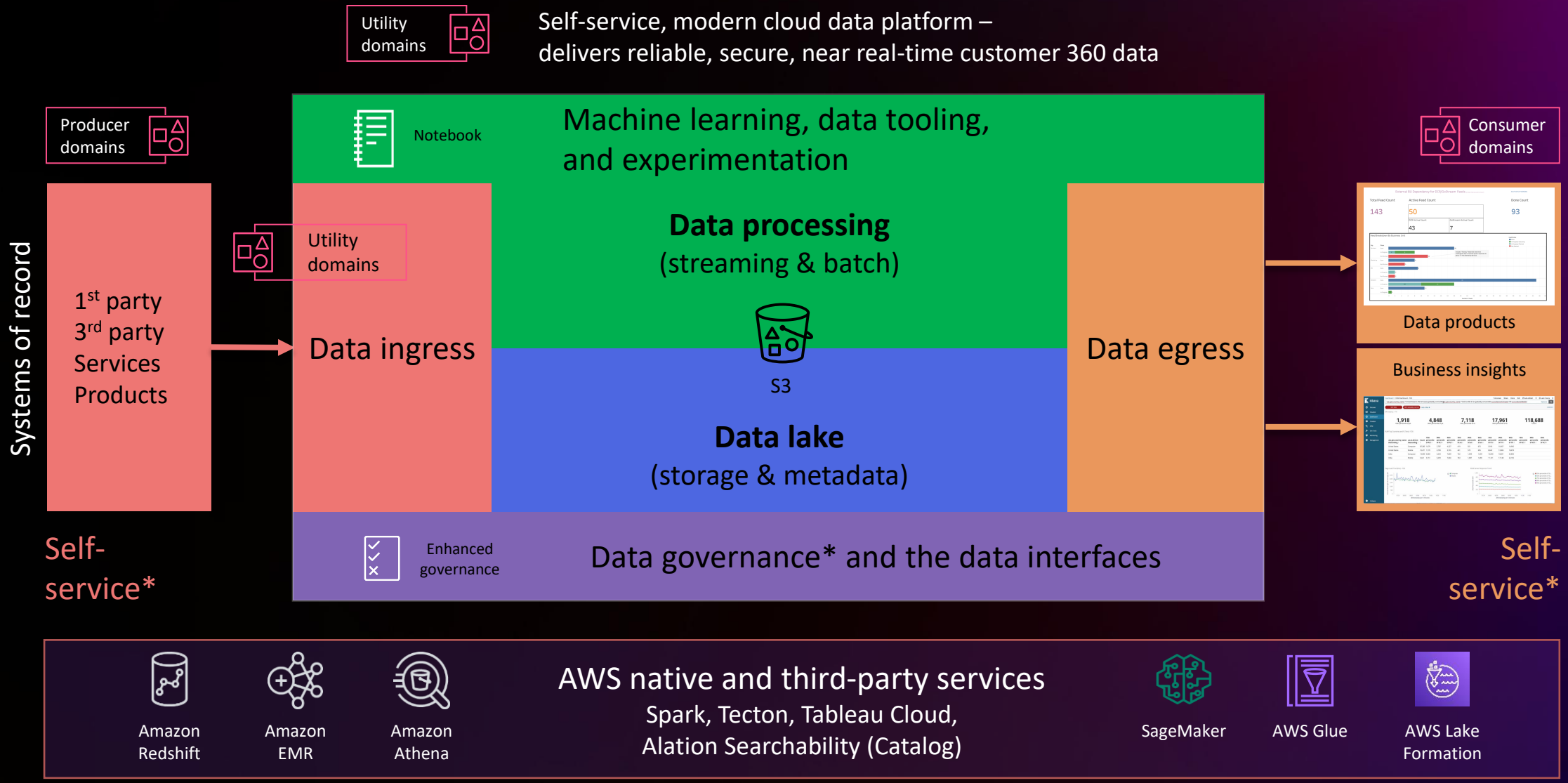
SageMaker

AWS Glue

AWS Lake Formation

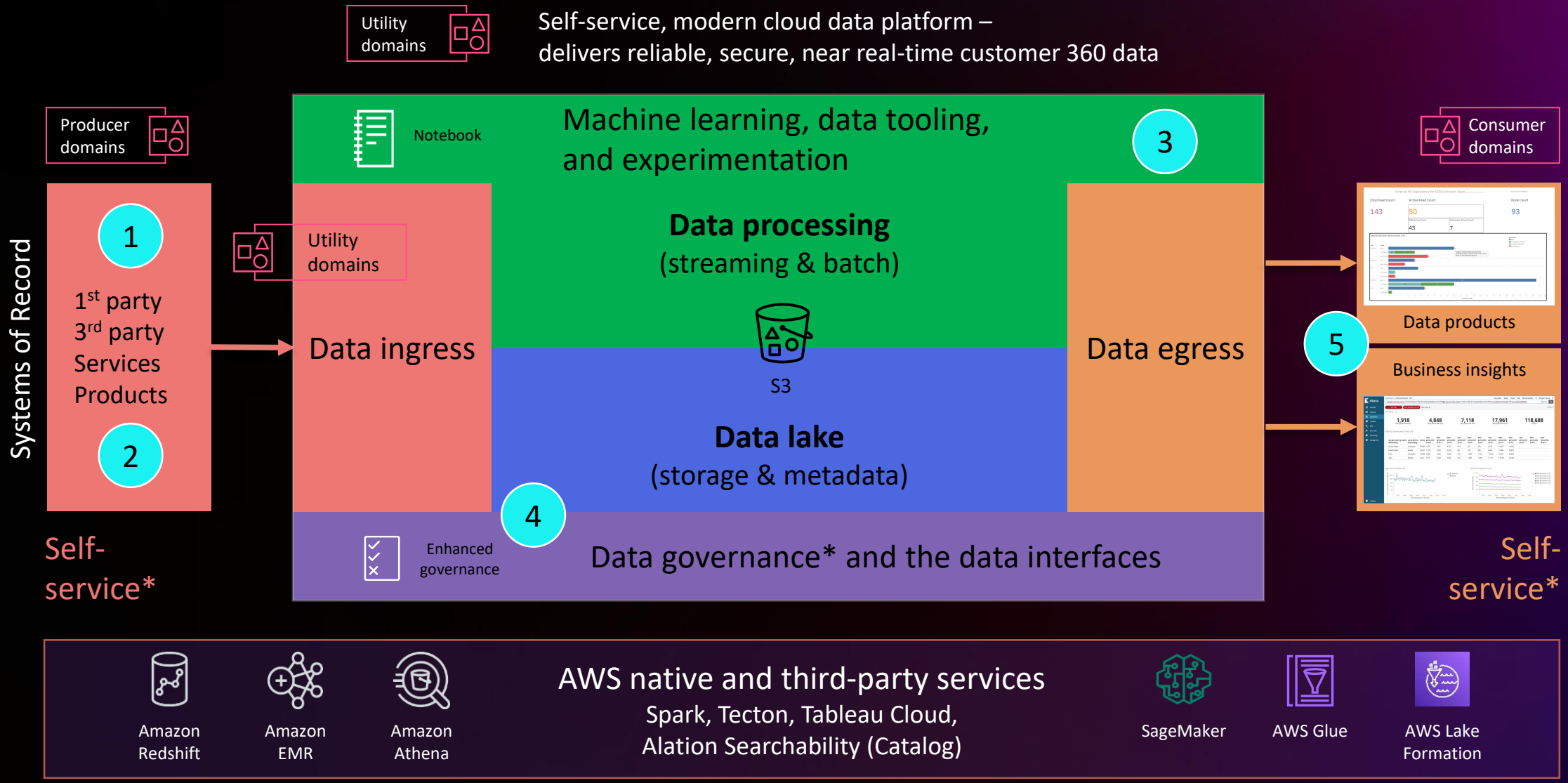
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh



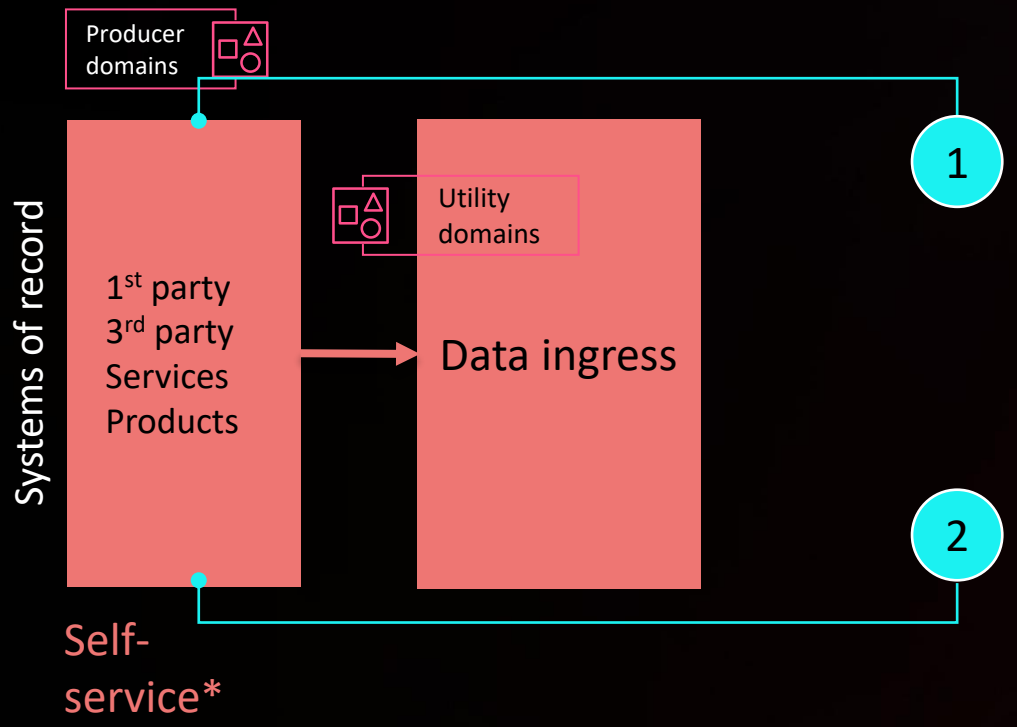
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh



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

Data mesh



1

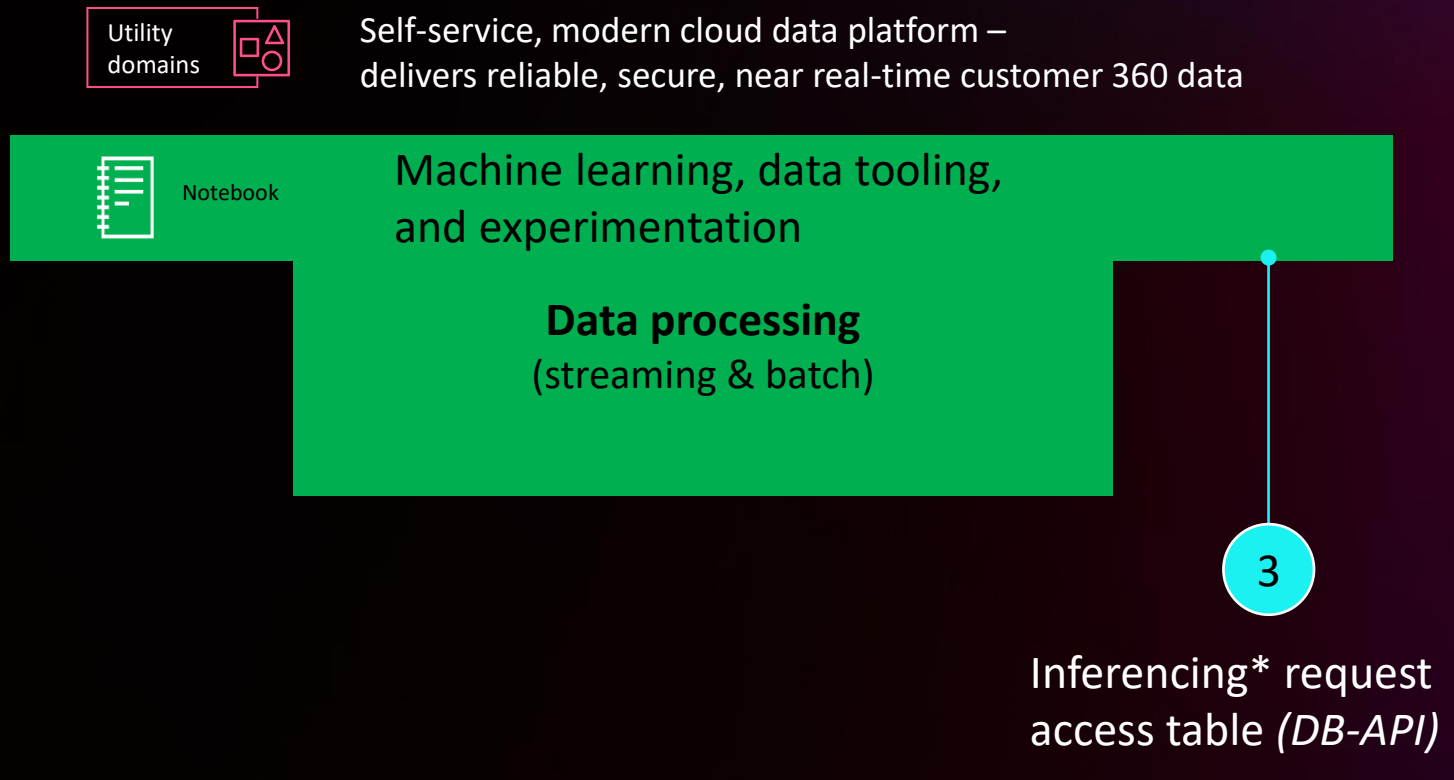
Producers instrument DOMAIN* telemetry, share data via data platform

2

Producers register DOMAIN* data in catalog

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

Data mesh

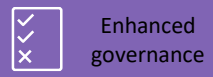


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh

Data governance council enforces policies

4



Data governance* and the data interfaces

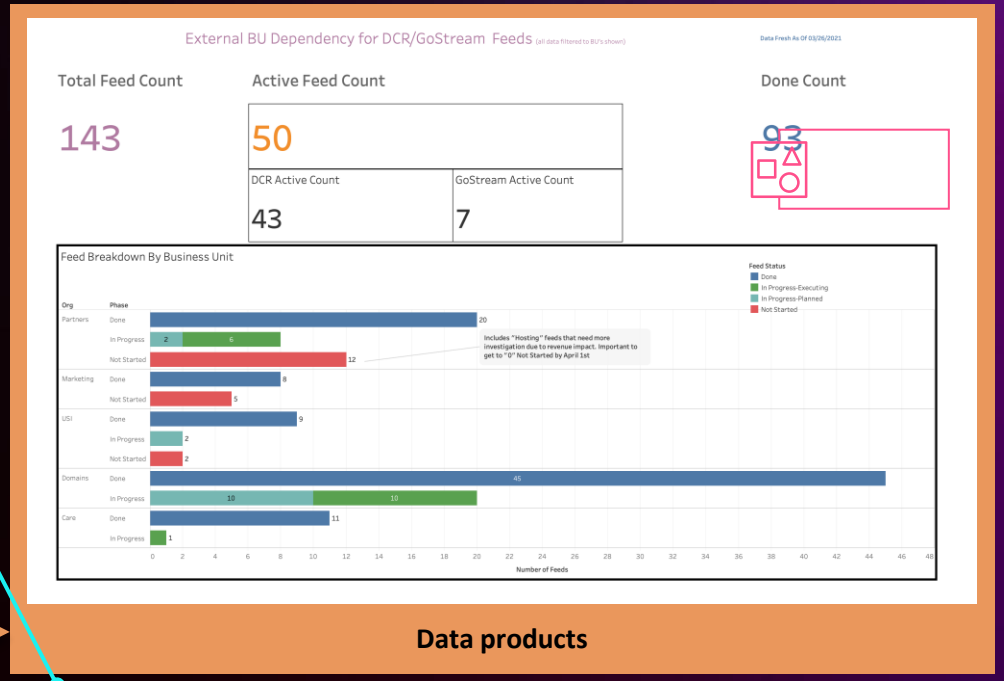
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh

Consumers use shared DOMAIN* data

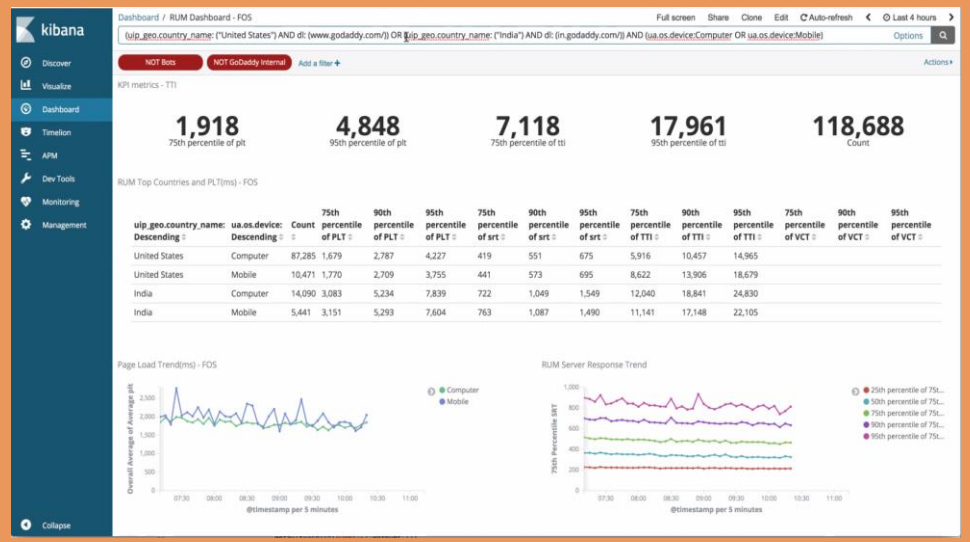
5

Data egress



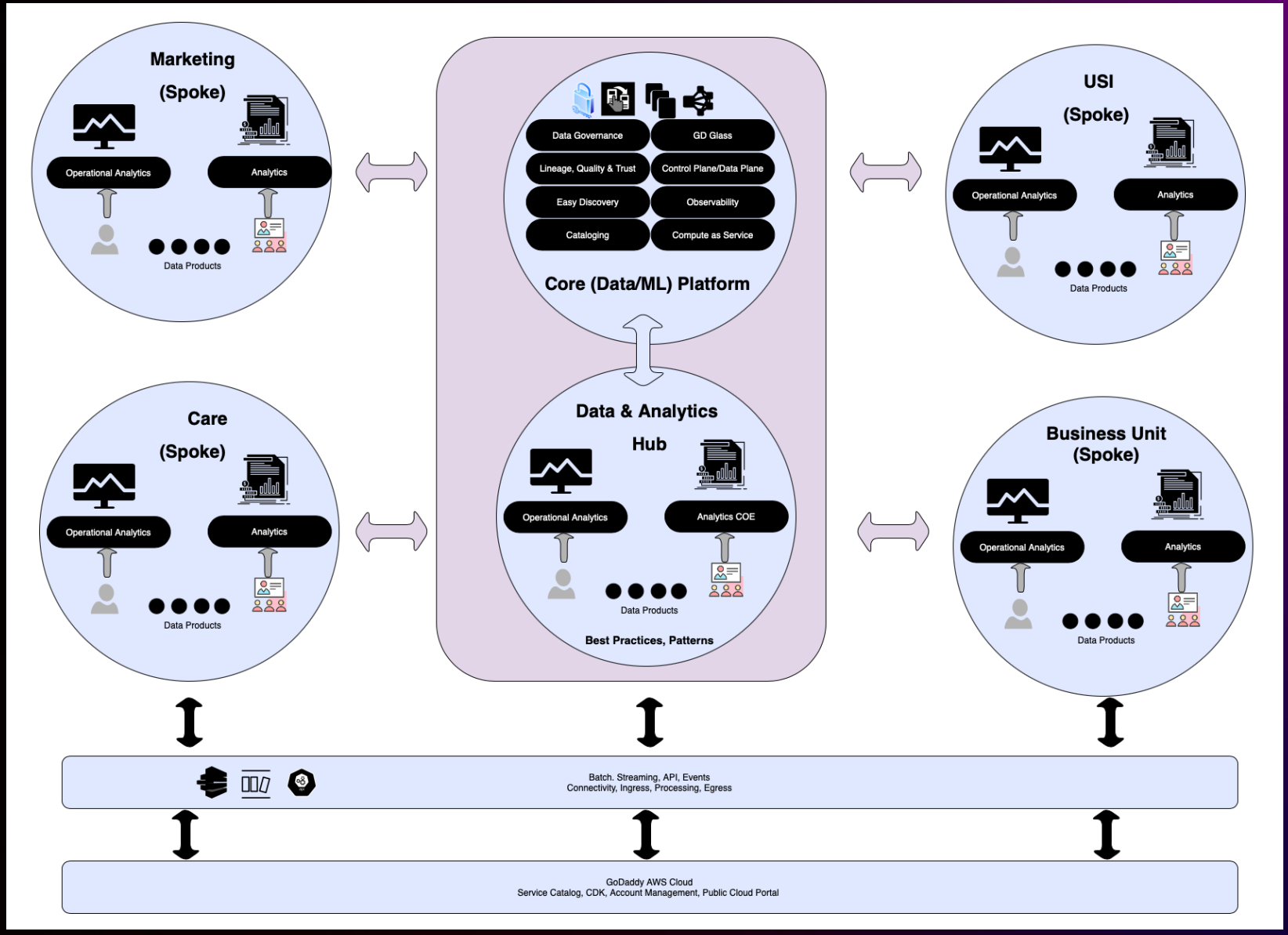
Data products

Business insights



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

GoDaddy conceptual/ domain architecture



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

Business outcomes

- Created hierarchical views of data products at different levels so that business users can analyze information to make quicker business decisions
- Automated access management framework to enable self-served access to data within and across lines of business
- Accelerated the data platform adoption to 10+ LOBs and 300+ teams globally, with more to come in the future
- Enabled data scientists to find and access data needed to generate ML models across LOBs
- Achieved 10s of millions of dollars in cost savings from data reuse and better management of purchased datasets

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

Thank you!

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

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