

# **AWS State, Local, and Education Learning Days**

North Carolina



# Building a Modern Data Strategy

**Vishal Lakhotia (he/him)**

Sr. Solutions Architect  
AWS

[lakhov@amazon.com](mailto:lakhov@amazon.com)

**Srinija Desiraju (she/her)**

Solutions Architect  
AWS

[dsrinija@amazon.com](mailto:dsrinija@amazon.com)

# Agenda

Why modern data architecture

Modern data strategy

Reference architectures for common scenarios

Getting started

# Why Modern Data Architecture



**“If we have data, let’s look  
at the data.  
If all we have are opinions,  
let’s go with mine”**

Jim Barksdale  
CEO of Netscape

# Data is just the Building Blocks

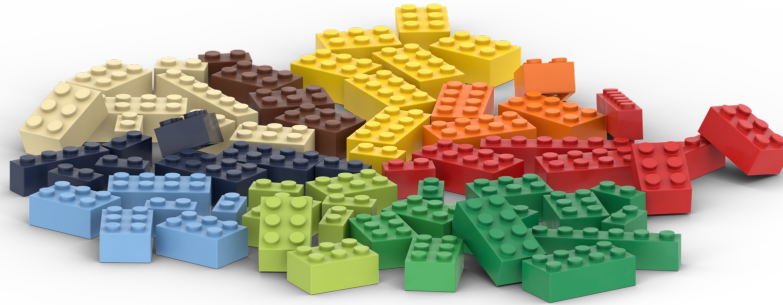
**Data**



**Information**

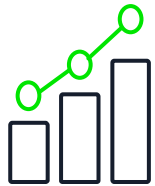


**Insights**



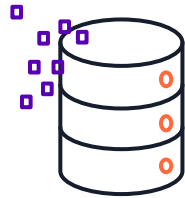
Without structure, tools and processes,  
Data has very little value

# The data challenge



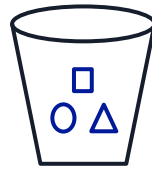
---

Availability of  
electronic data  
is growing  
exponentially



---

Data coming from  
new, disconnected  
sources



---

Increasingly  
diverse in file type  
and volume



---

Used by  
many people (e.g.  
policy makers,  
researchers, etc.)

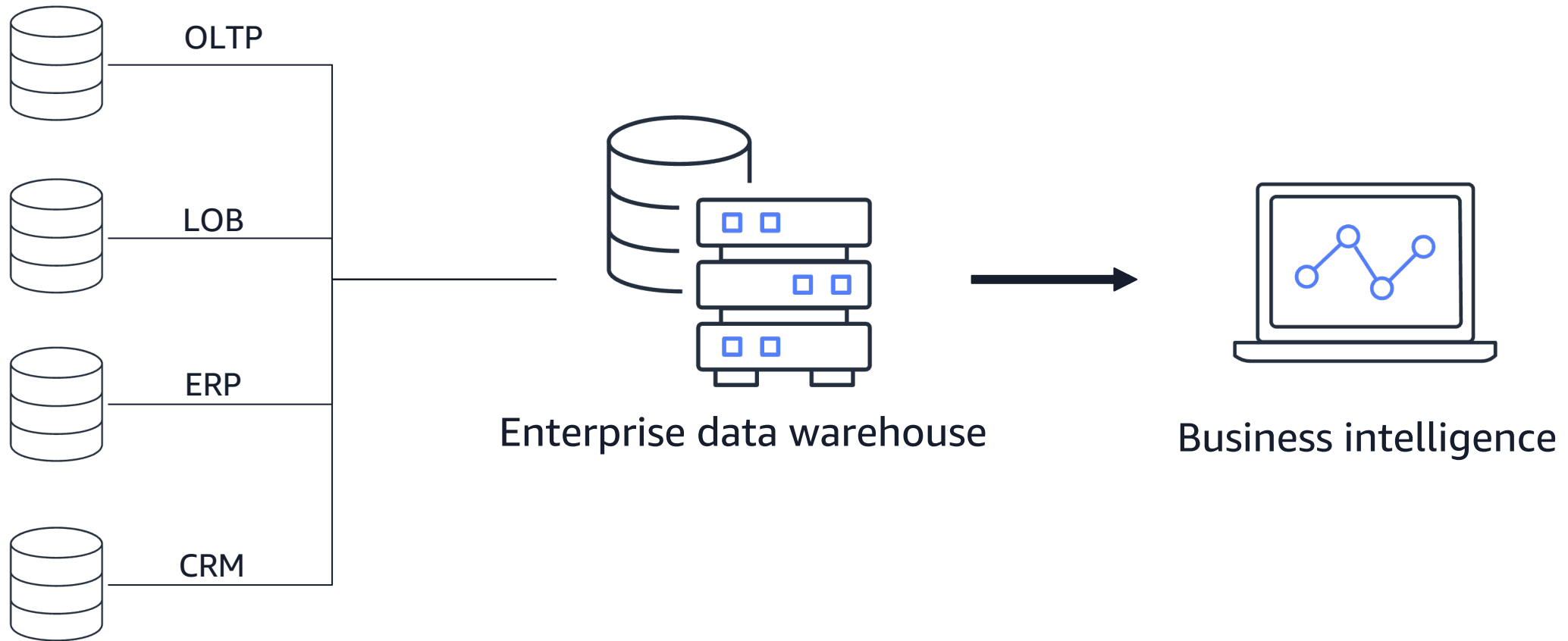


---

Analyzed by  
many applications

# Current state

Currently, decision-making revolve around the **enterprise data warehouse**





# Data no longer scales

There is more data and more diversity of data than people think

---

Data growth

**>10x**  
every 5 years

Data  
platforms need

To live for  
**15+**  
years

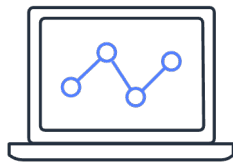
To scale  
**1,000x**



# Accessibility of data



Data scientists



Business users



Analysts



Applications



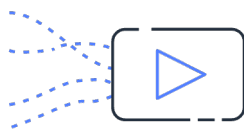
Machine learning



SQL analytics



Scientific

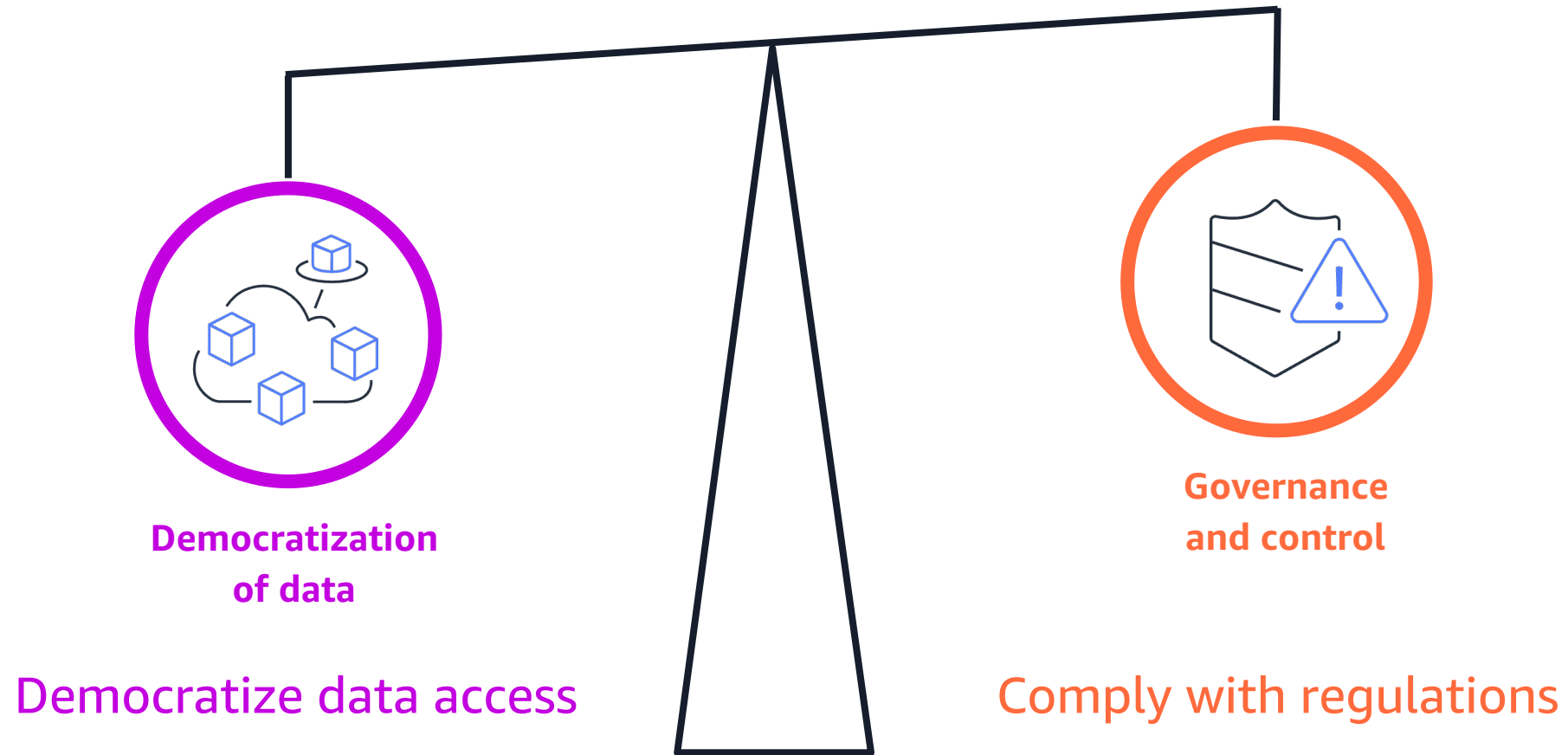


Real-time, streaming

There are **more people** accessing data

And in **different ways**

# More regulatory pressure



# What now? **Let's rethink everything**



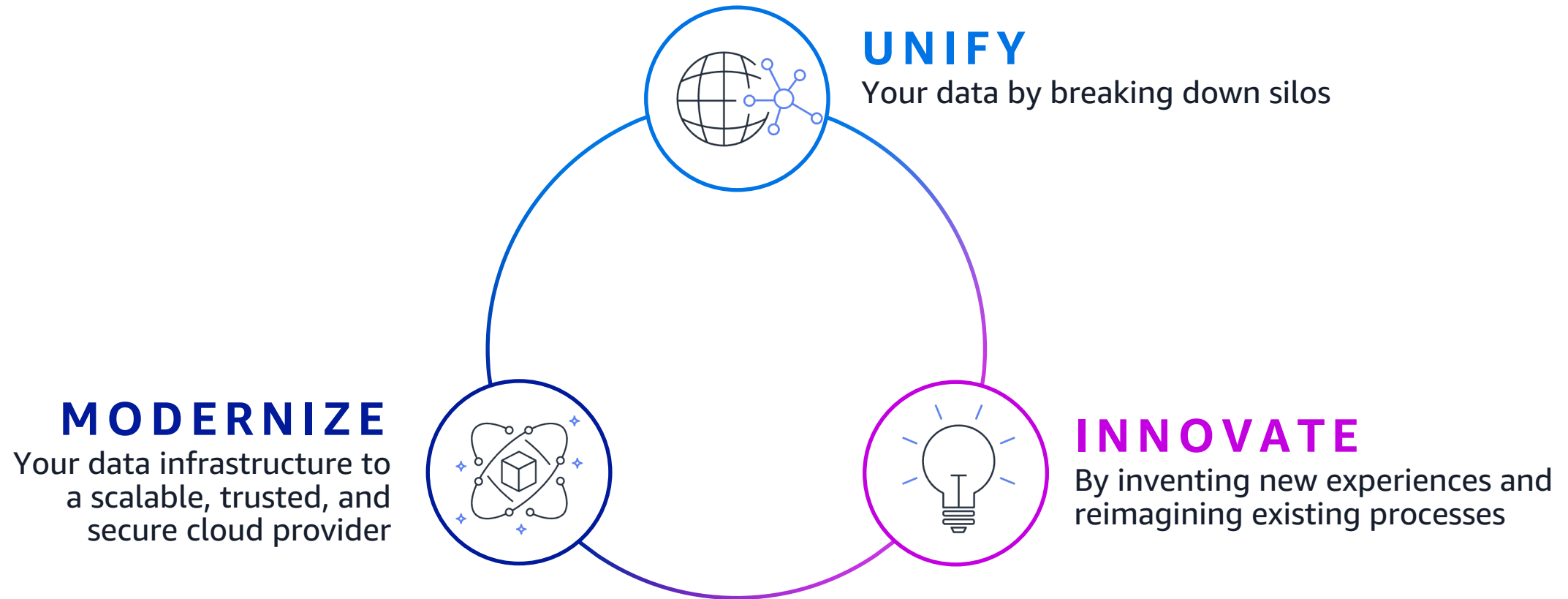
**Raw Data**



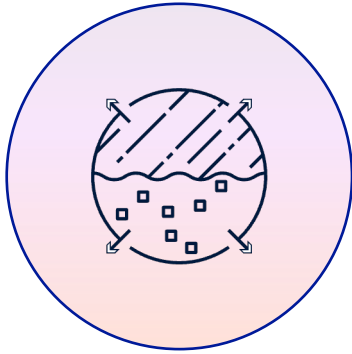
**Insights**

# Modern Data Strategy

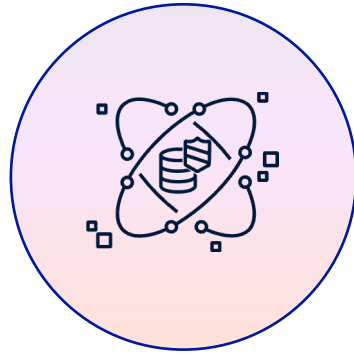
# Modern data strategy for better business outcomes



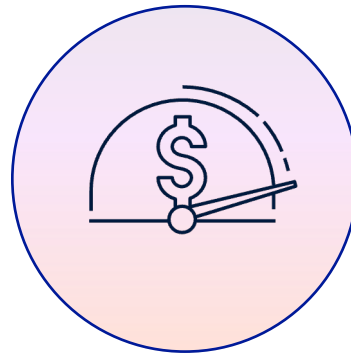
# AWS modern data strategy components



Scalable  
data lakes



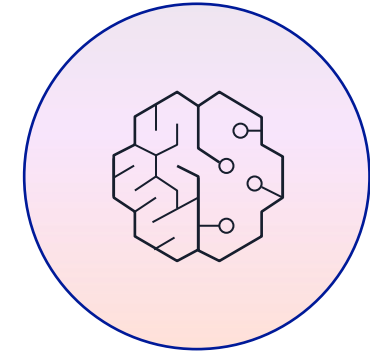
Unified data  
access, security,  
and governance



Purpose-built  
data services for  
performance  
and cost



Serverless  
and easy to  
use



Built-in  
machine  
learning

# Create better business outcomes with data



Make better, faster decisions

---



Improve customer experience

---



Prepare for the future

---



Reduce costs and improve productivity



# Examples



Create better citizen & student experiences & outcomes



Student success  
& community  
relevance



Respond to the unexpected



Support  
research in the  
swine industry



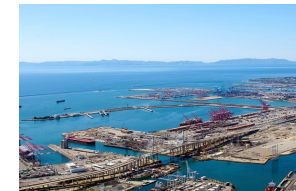
Unifying data to enable 360-degree views



Transform  
human services

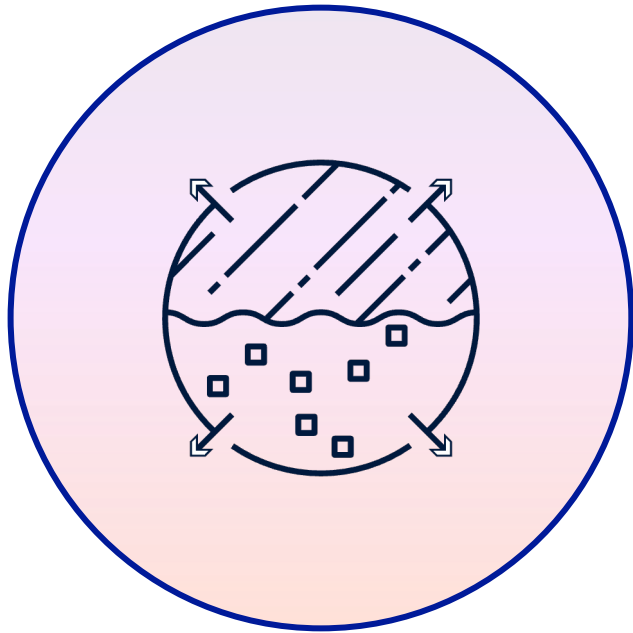


Enhancing efficiency



Create end-to-  
end visibility

Port of Long Beach



# Unify Data with **Scalable** data lakes

# Amazon S3: Data lakes on AWS

Store unlimited data in open file formats

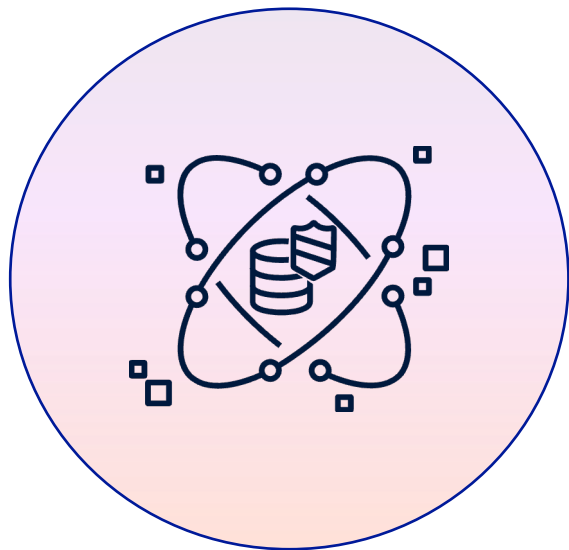
Unmatched durability, availability, and scalability

Decouple storage from compute

Choice of analytical and ML engines

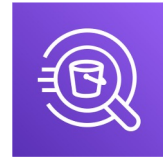
Pay as you go





# Unified data access, security, and governance

# AWS Lake Formation: Unified data governance



Amazon  
Athena



Amazon  
QuickSight



Amazon  
Redshift

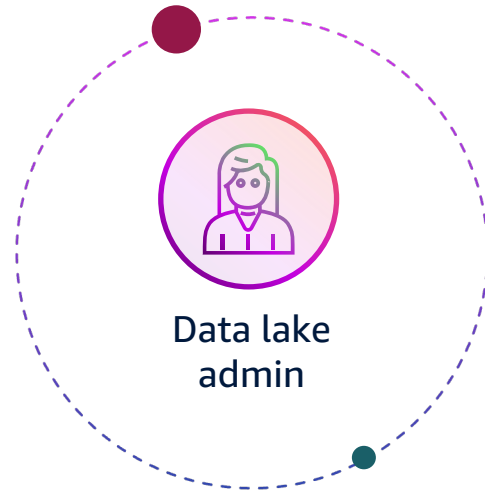


Amazon  
SageMaker

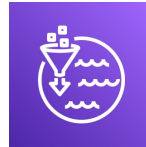


Amazon  
EMR

Simplified and unified  
security management



Data lake  
admin



AWS Lake  
Formation



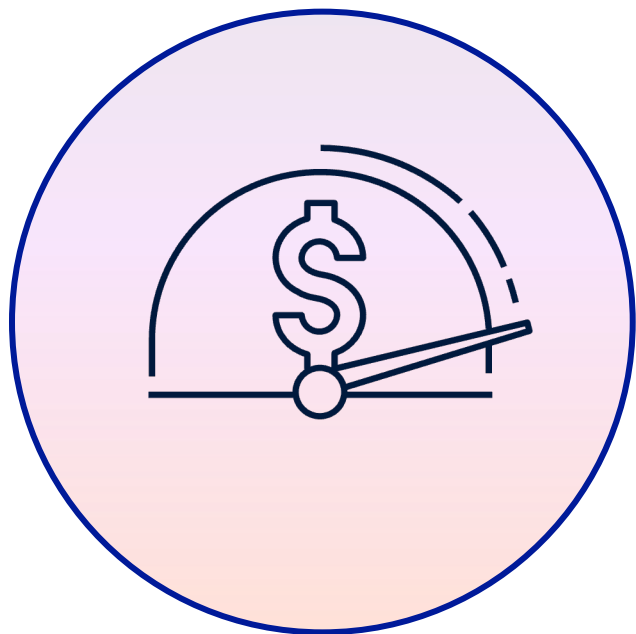
Access  
control



AWS Glue  
Data Catalog



Amazon S3 data lake storage



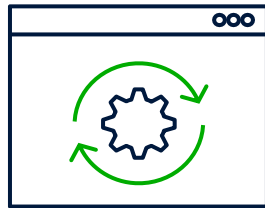
# Modernize with Purpose-built data services

# To get more value from their data, customers are...



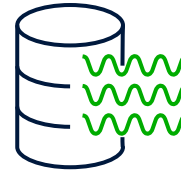
---

Breaking free from  
legacy databases



---

Moving to fully  
managed database  
and analytics services



---

Modernizing your  
data warehouse














---






Building modern  
applications with  
purpose-built  
databases

# A family of purpose-built data services









## Business intelligence and machine learning

 <b>Amazon QuickSight</b> Visualizations	 <b>AWS Data Exchange</b> Data exchange	 <b>Amazon SageMaker</b> ML	 <b>Amazon Comprehend</b> NLP	 <b>Amazon Transcribe</b> Speech-to-text	 <b>Amazon Textract</b> Extract text
 <b>Amazon Personalize</b> Recommendations	 <b>Amazon Forecast</b> Forecasts	 <b>Amazon Translate</b> Translation	 <b>Amazon CodeGuru</b> Code reviews	 <b>Amazon Kendra</b> Enterprise search	Plus many more

## Analytics

 <b>Amazon Redshift</b> Data warehousing	 <b>Amazon EMR</b> Hadoop + Spark	 <b>Amazon Athena</b> Interactive analytics
 <b>Amazon OpenSearch Service</b> Operational analytics	 <b>Amazon Kinesis Data Analytics</b> Real time	




## Databases

 <b>Amazon Aurora</b> MySQL, PostgreSQL	 <b>Amazon DynamoDB</b> Key value, Document	 <b>Amazon Neptune</b> Graph
 <b>Amazon RDS</b> MySQL, PostgreSQL, MariaDB, Oracle, SQL Server, DB2	 <b>Amazon DocumentDB</b> Document	 <b>Amazon Timestream</b> Time series
	 <b>Amazon Keyspaces (for Apache Cassandra)</b> Wide column	 <b>Amazon ElastiCache</b> Redis, Memcached

## Blockchain

**Amazon Managed Blockchain**

## Data lake

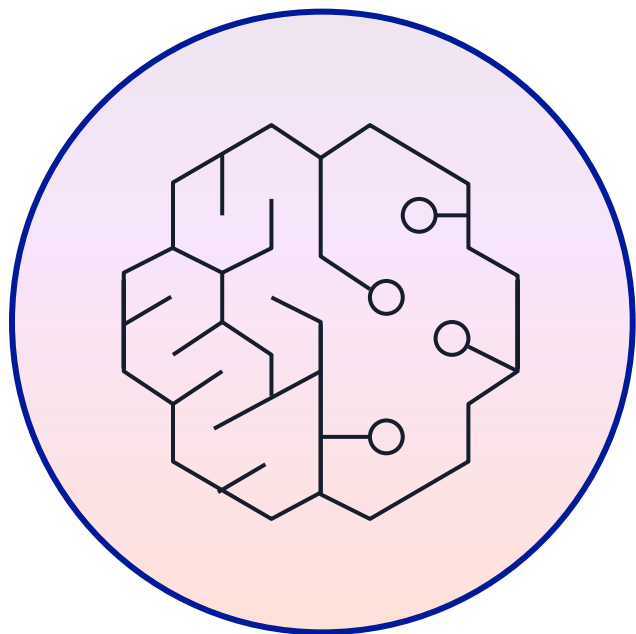
 <b>Amazon S3/ Amazon S3 Glacier</b>	 <b>AWS Lake Formation</b> Data lakes	 <b>AWS Glue</b> ETL and Data Catalog
---	---	---

## Data movement

AWS DMS | AWS Transfer family | AWS Glue | AWS Snowball | Amazon Kinesis Data Firehose | Amazon Kinesis Data Streams | Amazon MSK







# Innovate with AI & machine learning

# Build new experiences and reimagine old processes with AI/ML

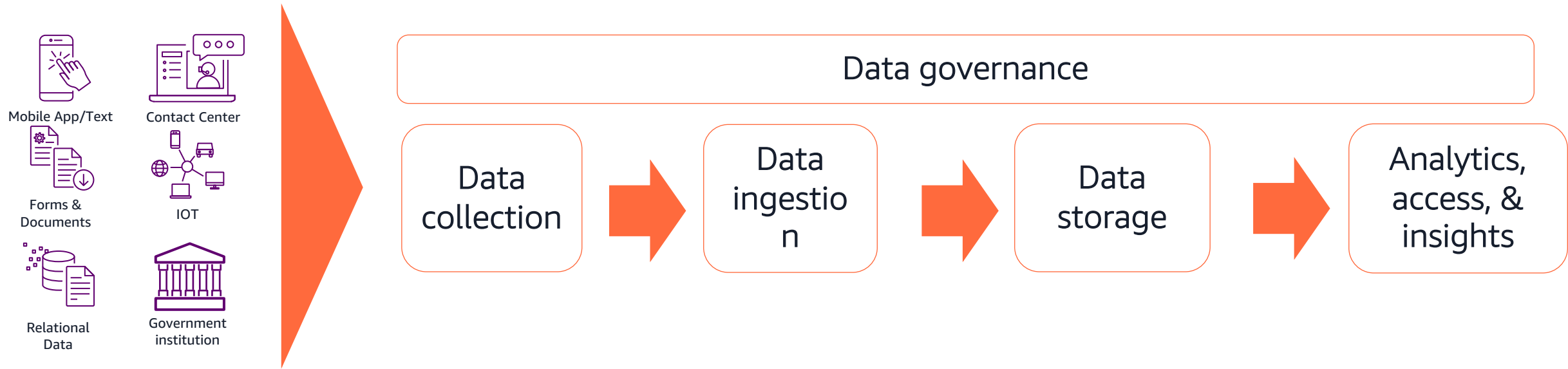


- Make accurate predictions, get deeper insights from your data, and improve customer experience
- Create ML predictions without any ML experience or writing any code
- Build applications with our pre-trained models
- Train and apply your own models
- Use your own algorithms by working directly with ML-optimized AWS infrastructure
- 100,000+ customers use AWS AI and ML services to make predictions from their data

# Putting it all together



# Key components of modern data architecture



Security – Reliability – Operational Excellence – Performance Efficiency – Cost Optimization – Sustainability

Key considerations:

1

Ability to handle the increasing volume, velocity, and variety of data

2

Each component should be independently scalable

3

Make data easily accessible and sharable

# Reference Architectures

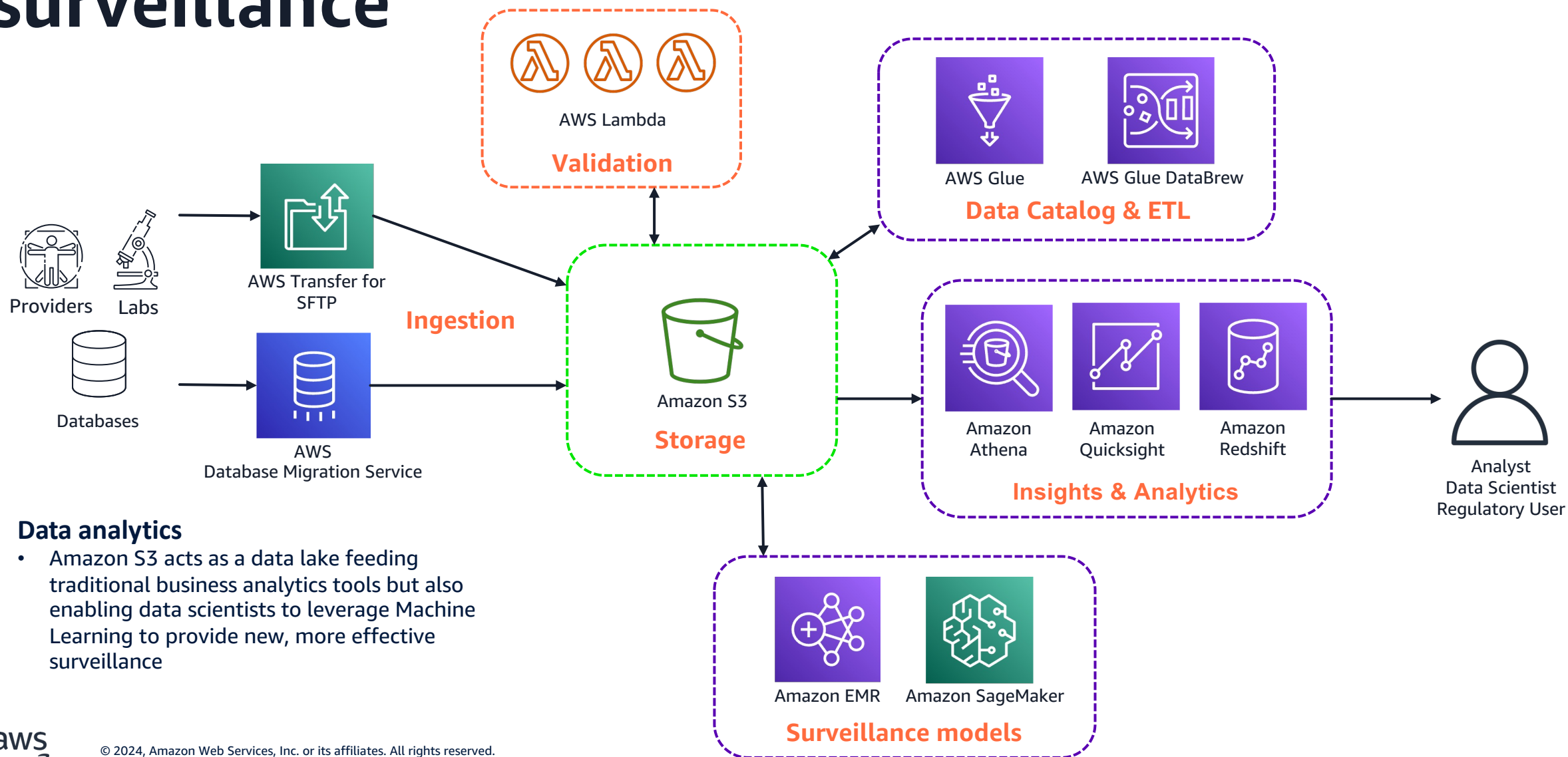


# Public Health Organization

- 1 Pandemic brings 1000% increase in disease surveillance data
- 2 Legacy management systems
- 3 Limited capabilities to consolidate data sets from multiple systems
- 4 Difficulty mandating data formats from various partner organizations



# Sample reference architecture for disease surveillance





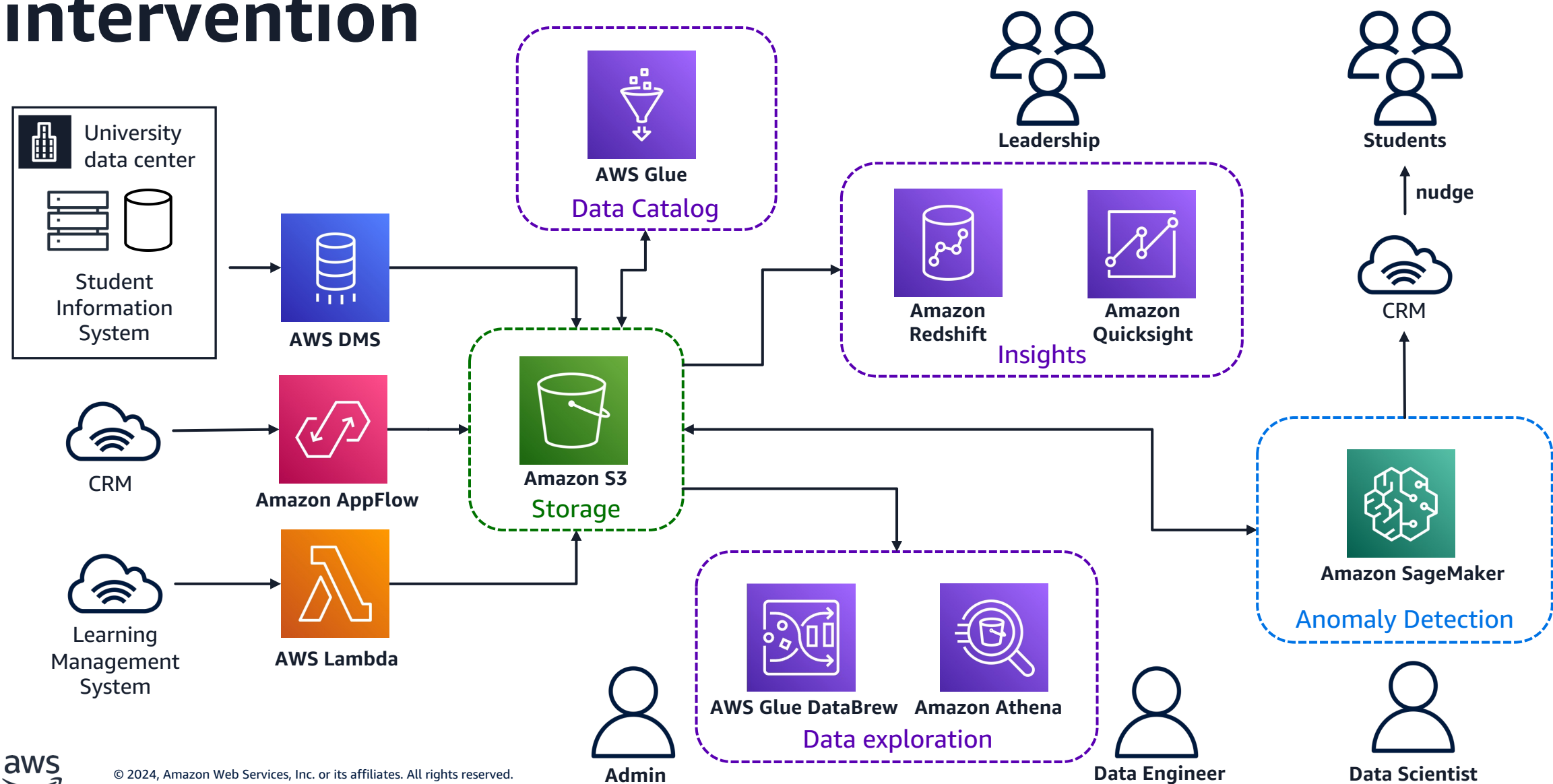
# Improving student outcomes - Retention

- 1 Identify at-risk students from behaviors
- 2 Aggregated student touchpoint data from the **SIS, LMS, and CRM**
- 3 Feed insights into communication platform for **early intervention and nudging**





# Sample reference architecture for student intervention



# Moving the needle on retention

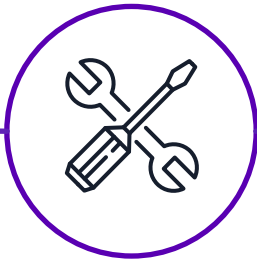
MARYVILLE UNIVERSITY

- 1 IT staff participated in data lake and modern data architecture **skills development**
- 2 Aggregated student touchpoint data from the **SIS, LMS, and CRM** into a data lake in 6 weeks
- 3 Automated processing and machine learning to **identify at-risk students** from behaviors
- 4 Fed insights into communication platform for **early intervention and nudging**



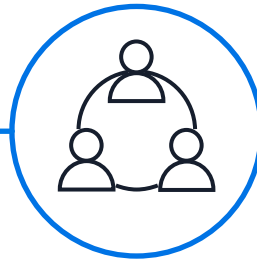
# Get started

## BUILD WITH US



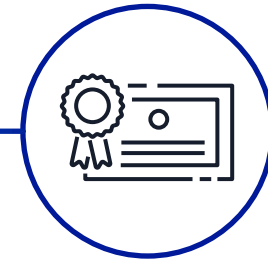
ML Solutions Lab  
AWS Professional Services  
AWS Immersion Day  
Data-Driven Everything  
Migration Assistance Program

## BUILD WITH PARTNERS



AWS Partner Network—  
100,000+ partners  
AWS Marketplace (ISVs)

## UPSKILL YOUR TEAMS



AWS Training and Certification  
ML Embark Program

# Thank you!

**Vishal Lakhotia**

lakhov@amazon.com

**Srinija Desiraju**

dsrinija@amazon.com



© 2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.



The image part with relationship ID rld3 was not found in the file.

Track: Data & Analytics

Session: Building a modern data strategy