



AWS Learning Days

Building a modern data strategy

Jason Moldan

Principal Solutions Architect
Amazon Web Services

Agenda

- Voice of the customer
- AWS modern data strategy
- Customer case studies and reference architectures
- How can AWS help?

Decision-making in the past

- In the past, decision-making revolved around the **enterprise data warehouse**



Data no longer fits

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

Data growth

>10x
every 5 years

Data
platforms needs

To live for
15+
years

To scale
1,000x



IDC, "Data Age 2025"



Accessibility of data



Data scientists



Business users



Analysts



Applications



Machine learning



SQL analytics



Scientific



Real-time, streaming

There are **more people**
accessing data

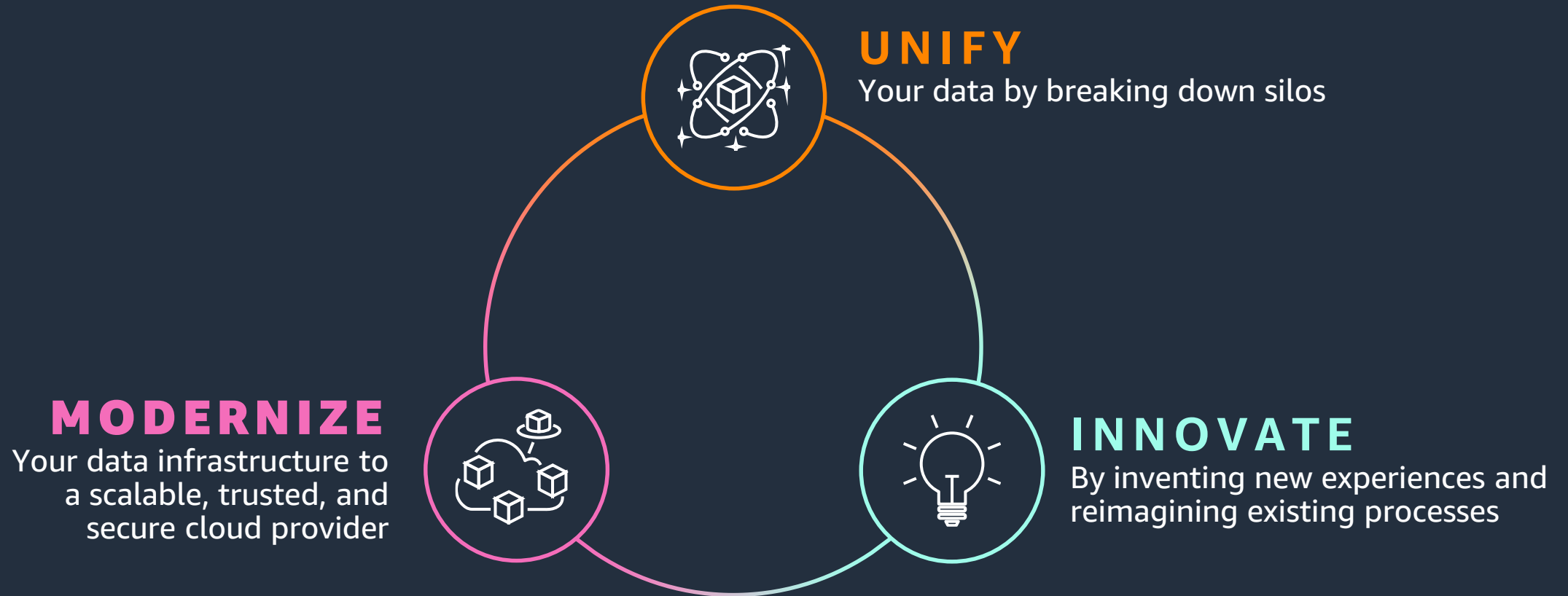
And in **different ways**

More regulatory pressure

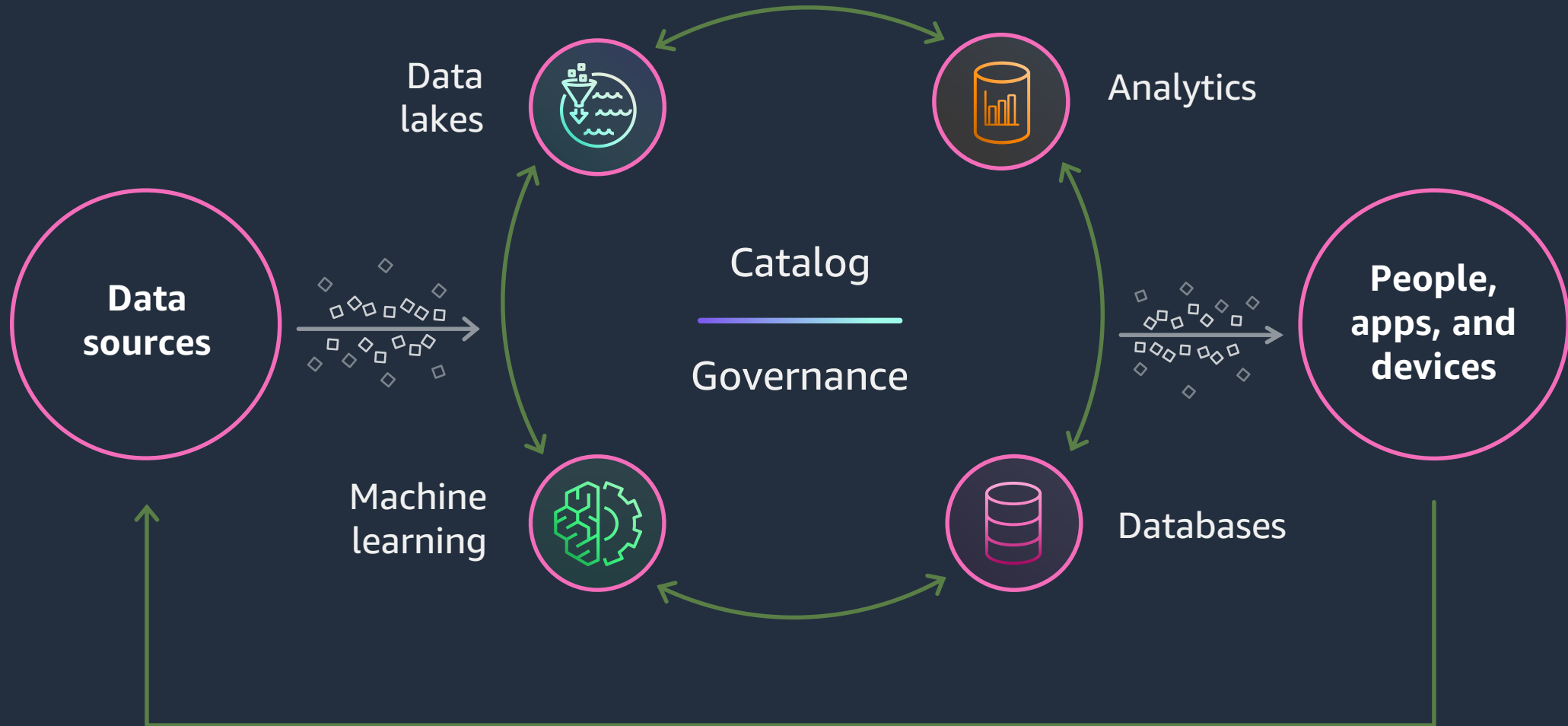


Modern data strategy for better business outcomes

Start anywhere



Modern data strategy



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



Inform Policy



Unifying data to enable 360-degree views



Transforming
human services



Enhancing efficiency



Save time,
money, and
lives



Unify Data with Scalable data lakes

Amazon S3: Data lakes on AWS

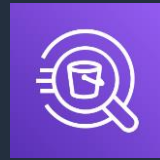
- Store unlimited data in open file formats
- Unmatched durability, availability, and scalability
- Decouple storage from compute
- Choice of analytical and ML engines
- Pay as you go



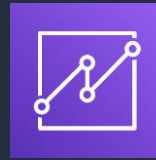


Unified data access, security, and governance

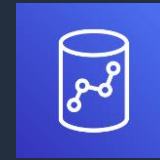
AWS Lake Formation: Unified data governance



Amazon
Athena



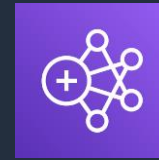
Amazon
QuickSight



Amazon
Redshift

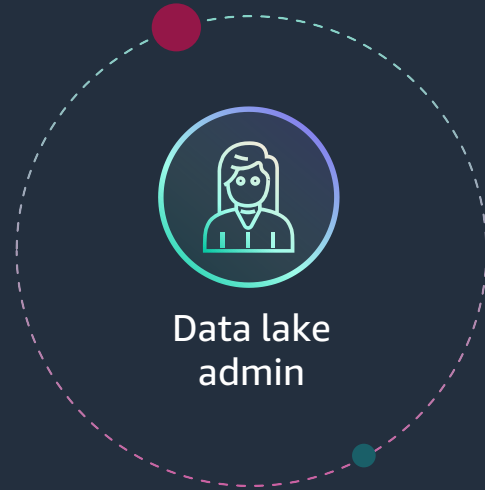


Amazon
SageMaker



Amazon
EMR

Simplified and unified
security management



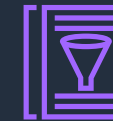
Data lake
admin



AWS Lake
Formation



Access
control



AWS Glue
Data Catalog



Amazon S3 data lake storage

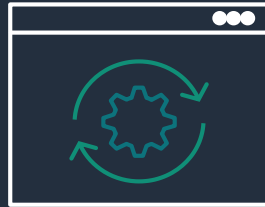


Modernize with Purpose-built data services

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



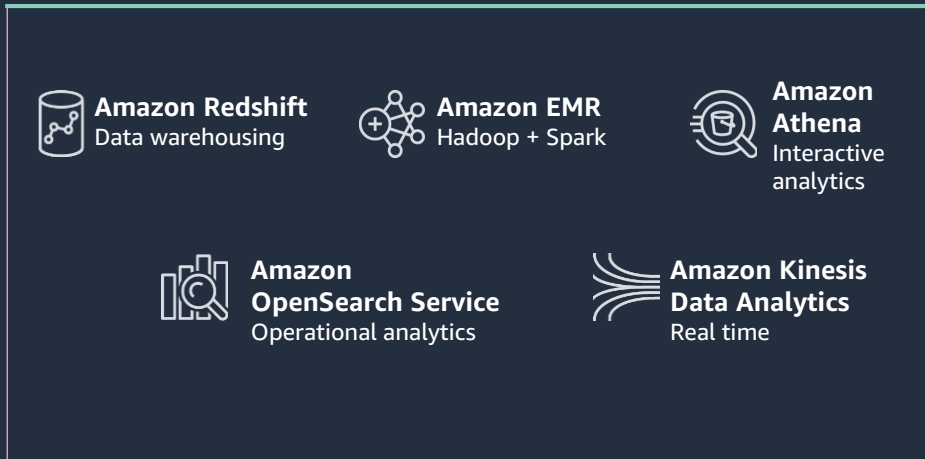
Building modern
applications with
purpose-built
databases

A family of purpose-built data services

Business intelligence and machine learning



Analytics



Databases



Blockchain



Data lake



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

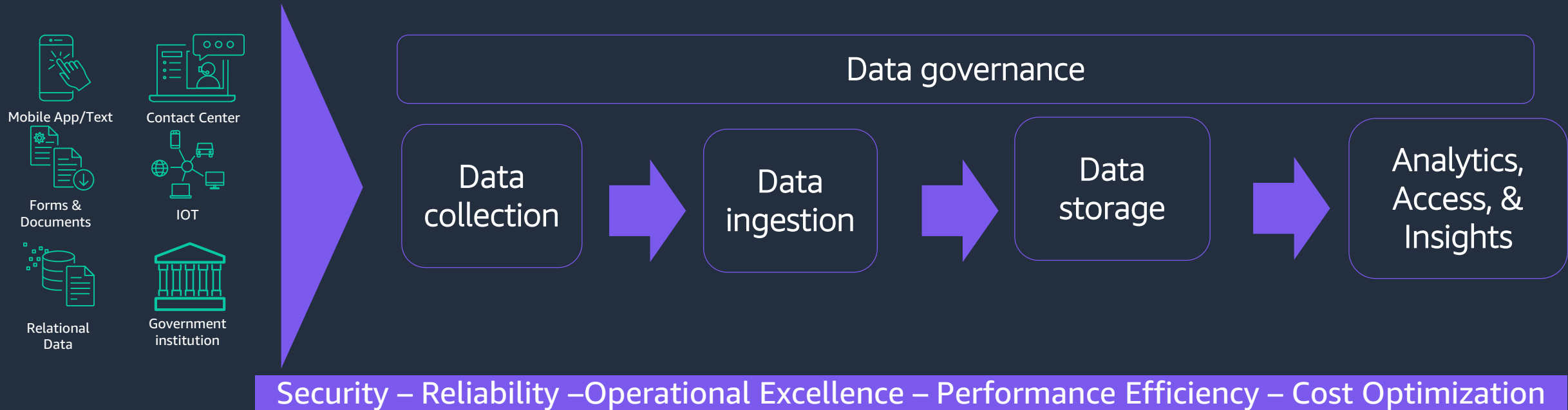
Build new experiences and reimagine old processes with AI/ML



- 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



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

Data validation

- Use serverless computing for event-driven data validation as files come in to reduce cost and processing times



AWS Lambda

Validation



AWS Glue



AWS Glue DataBrew

Data Catalog & ETL

Data from



Providers

Labs



Databases



AWS Transfer for SFTP



AWS DMS

Ingest



Amazon S3

Storage



Amazon EMR
(Apache Spark)

Surveillance models



Amazon Athena



Amazon Redshift



Amazon QuickSight

Insights & Analytics



Analyst
Data Scientist
Regulatory User

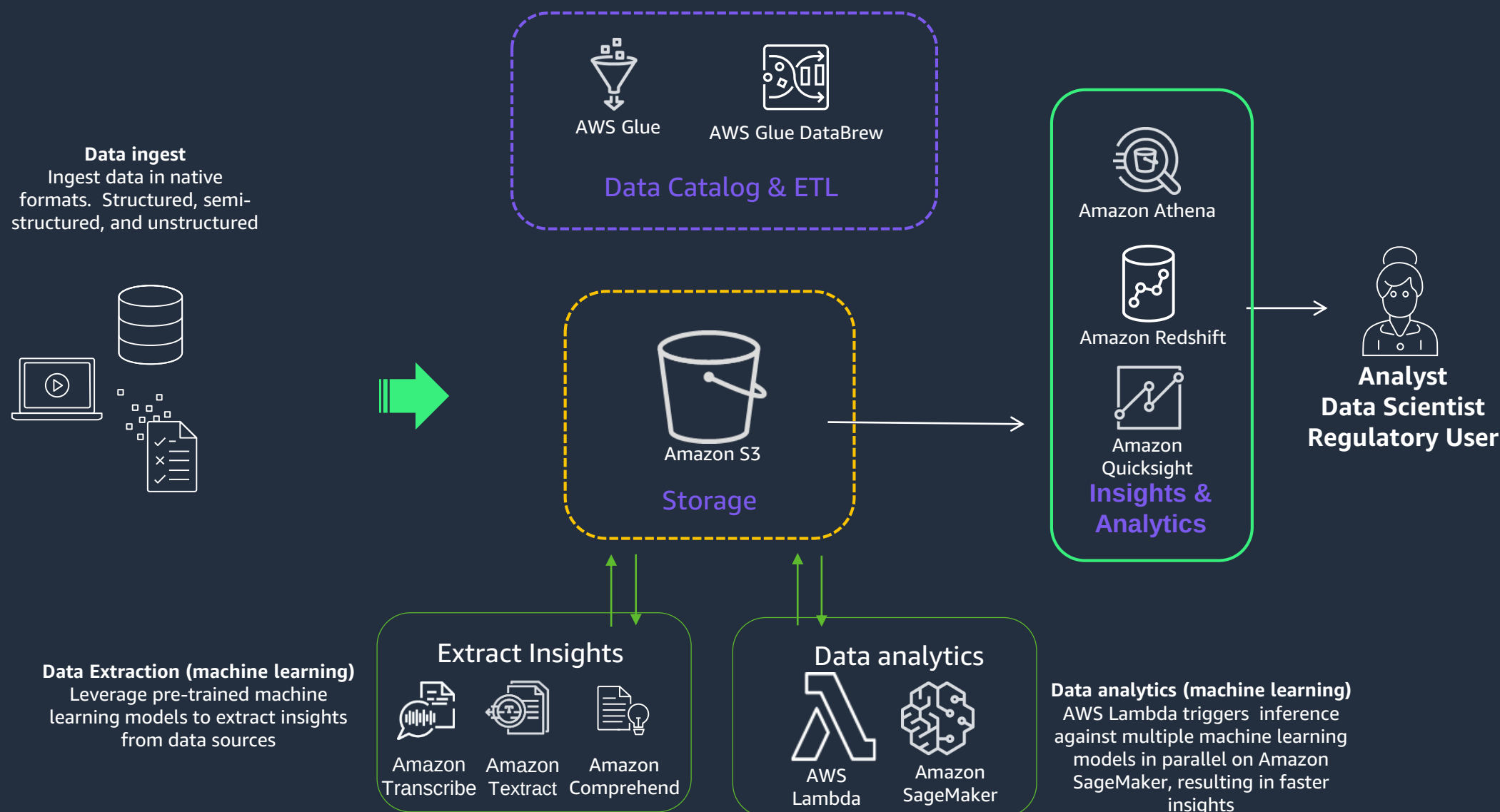
Data analytics

- Amazon S3 acts as a data lake feeding traditional business analytics tools but also enabling data scientists to leverage Machine Learning to provide new, more effective surveillance

Data ingestion

- Leverage fully managed and scalable services such as Amazon S3 and AWS Transfer for SFTP which automatically adjusts to highly variable market activity
- Database migration service (DMS) can be used for batch or CDC data transfer

Sample reference architecture for

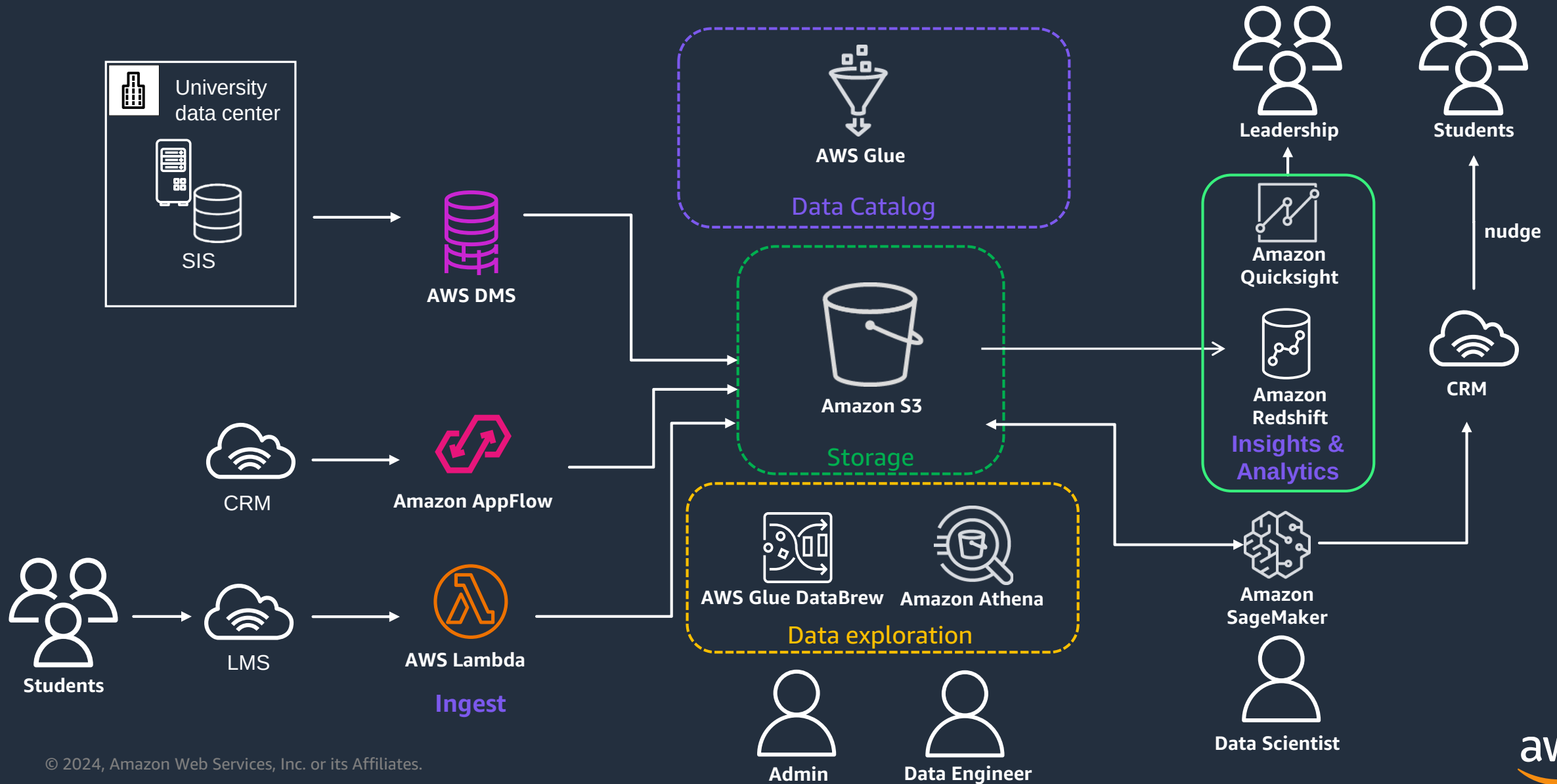


Improving student outcomes - Retention

- 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!

Jason Moldan

jmoldan@amazon.com



Step 1: Data & Analytics Track

Step 2: Building a modern data strategy

Resources & References

[Minnesota Department of Health– Data Lakes Video Case Study](#)

[Minnesota Department of Health – Public Sector Summit](#)

[A Healthier future for North Carolina - re:Invent](#)

[UDSA SNAP Case study](#)

[Utah Mainframe Modernization](#)

[Georgia Department of Community Health MMIS migration](#)

[Rhode Island Department of Labor - Citizen engagement](#)

[Utah Department of Health COVID response](#)

[Piedmont Healthcare – Epic Systems migration](#)

[Great Lakes Health Connect](#)

[State of Maryland transforming social service – MDThink](#)

[Driving results in higher education with data](#)

Extend the capabilities of your data warehouse with a modern data strategy on AWS

A lake house approach on AWS makes it easy to access data across applications, repositories, and lines of business to gain deeper insights

