

AWS State, Local, and Education Learning Days

North Carolina



DATA & ANALYTICS TRACK

Democratizing your organization's data analytics experience

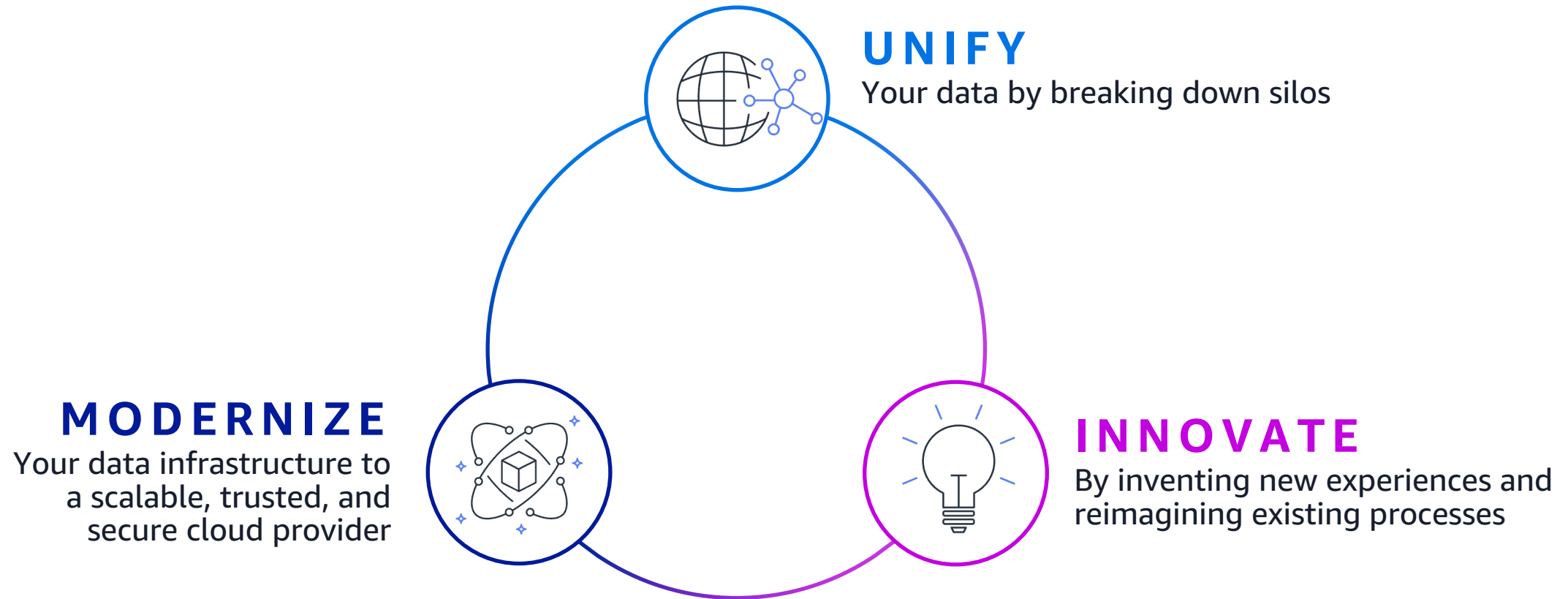
Greg Grieff

Sr. Solutions Architect

AWS

ggrieff@amazon.com

Modern data strategy for better business outcomes



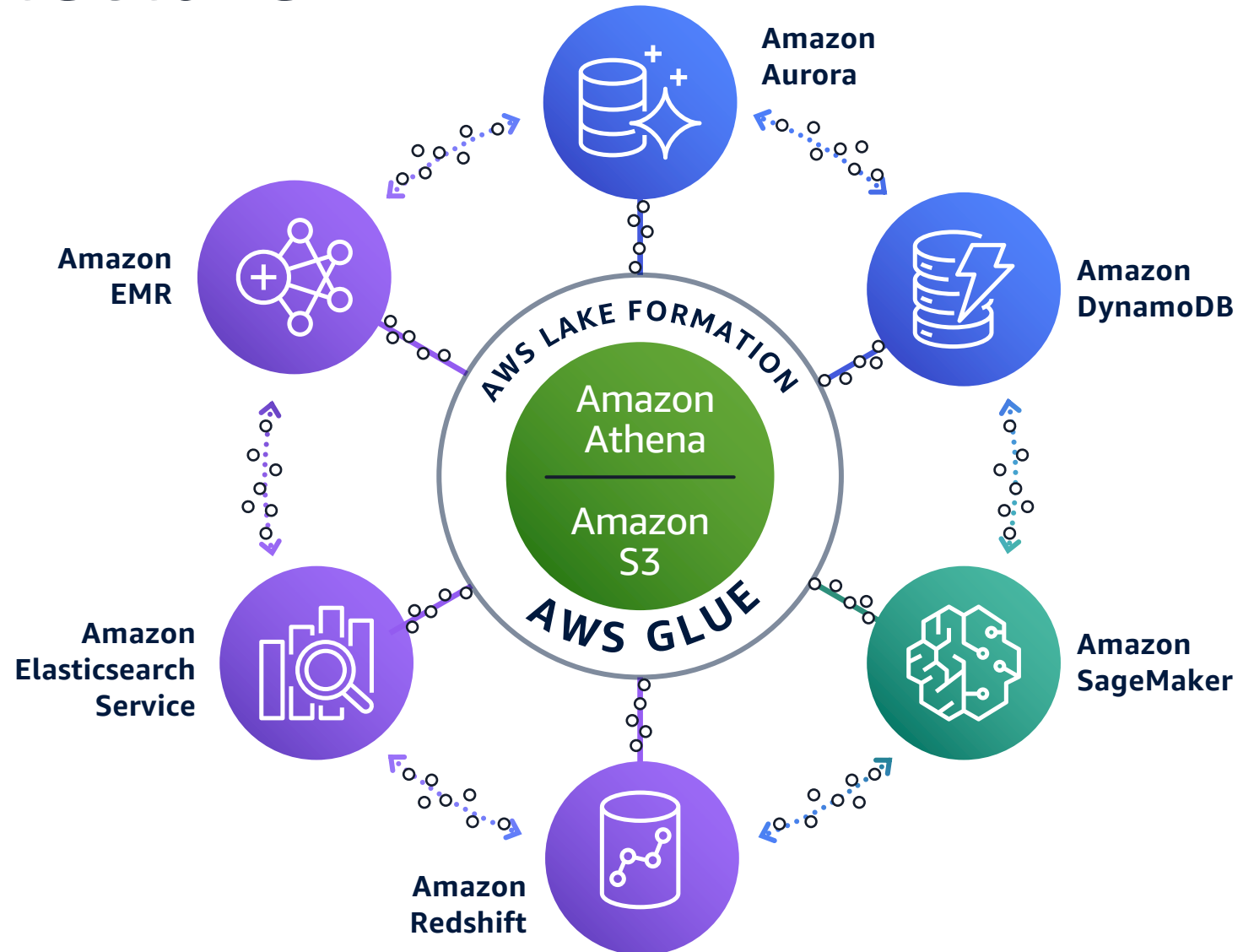
Modern Data Architecture

UNIFY DATA WITH
SCALABLE DATA LAKES

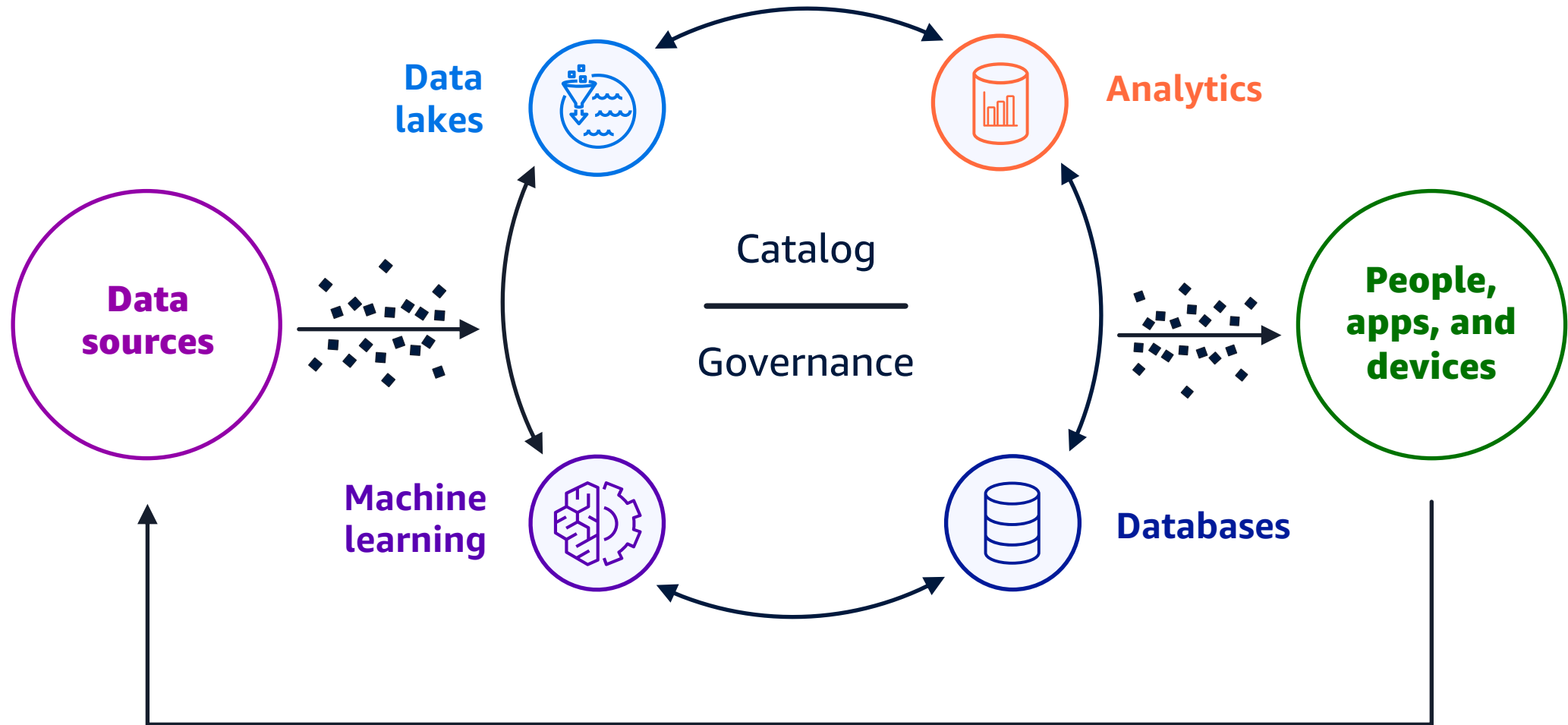
CENTRALIZED DATA CATALOG
SECURITY AND GOVERNANCE

MODERNIZE WITH
PURPOSE-BUILT SERVICES

INNOVATE WITH AI AND
MACHINE LEARNING



What is data democratization?



Public Safety Data Lake



Records Management System

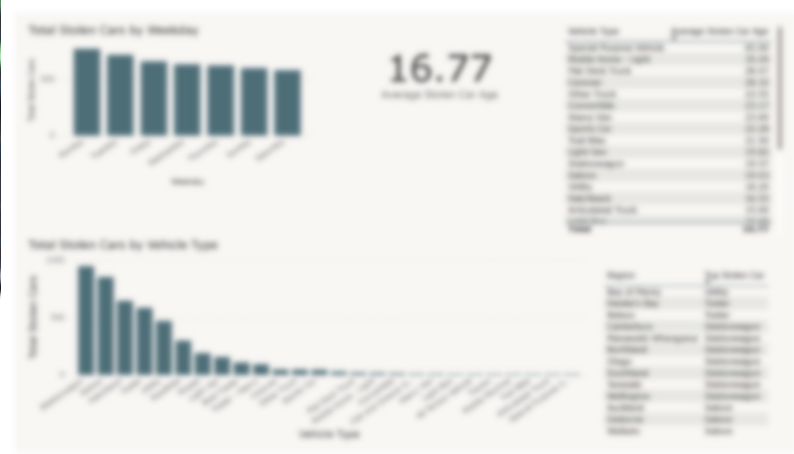
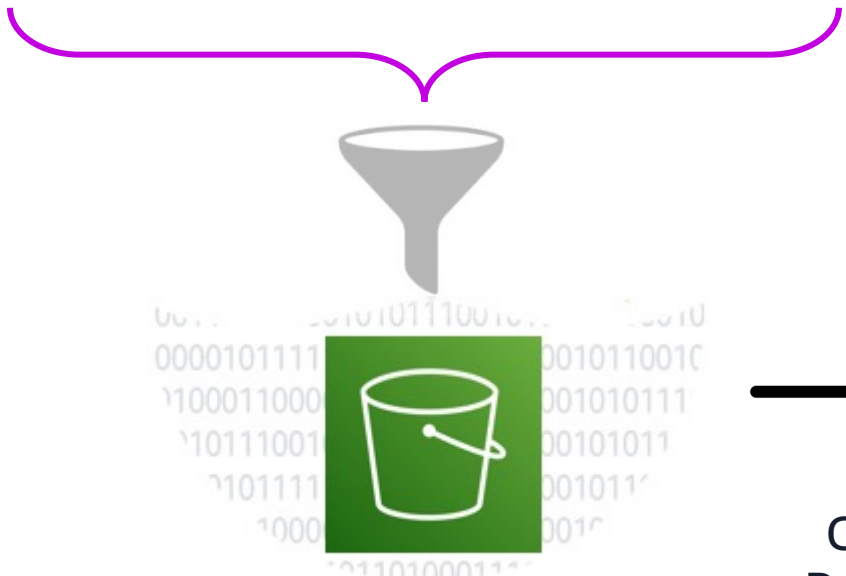
Computer Aided Dispatch

Digital Evidence Archive

Gun Registrations

CCTV Video Capture

License Plate Reader



Curated Data Products,
Dashboards, and Insights

Amazon S3- the best place to build a data lake



**Amazon Simple
Storage Service (S3)**

11 nines (99.999999999%) of durability

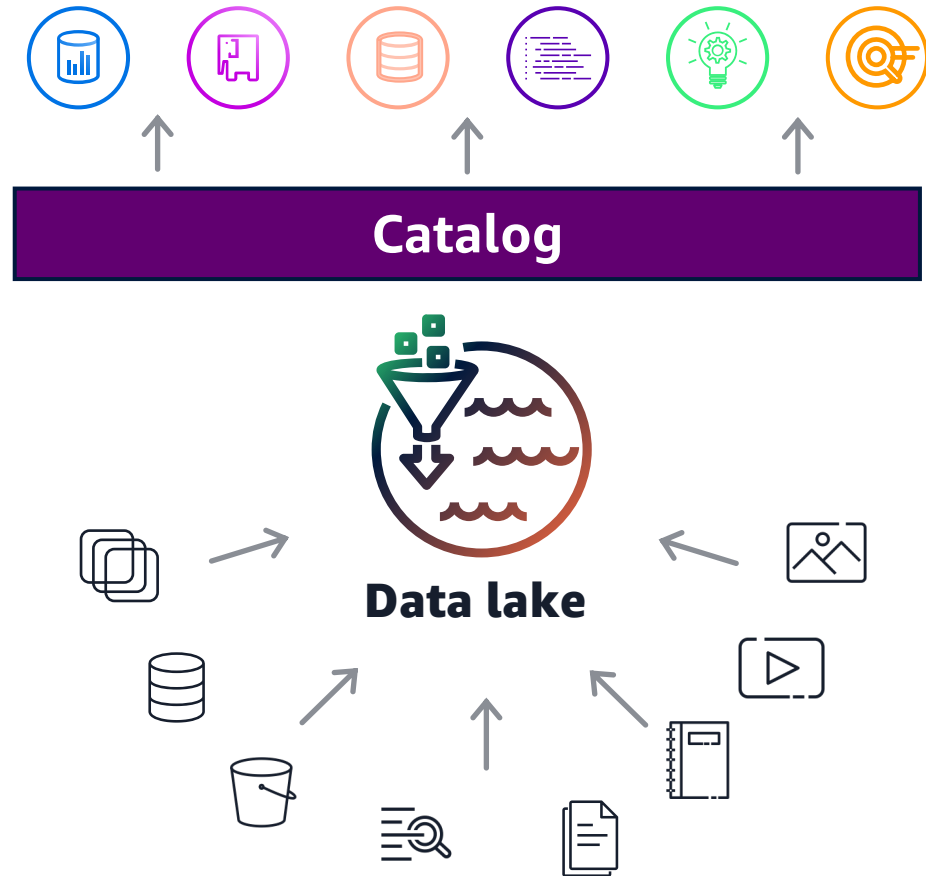
Security, compliance, & audit capabilities

Fastest performance at the lowest cost

The most ways to bring data in

Virtually unlimited scalability

Amazon Simple Storage Service is the best place to build a data lake



Store all your data in open formats

Cost-effectively scale storage to exabytes

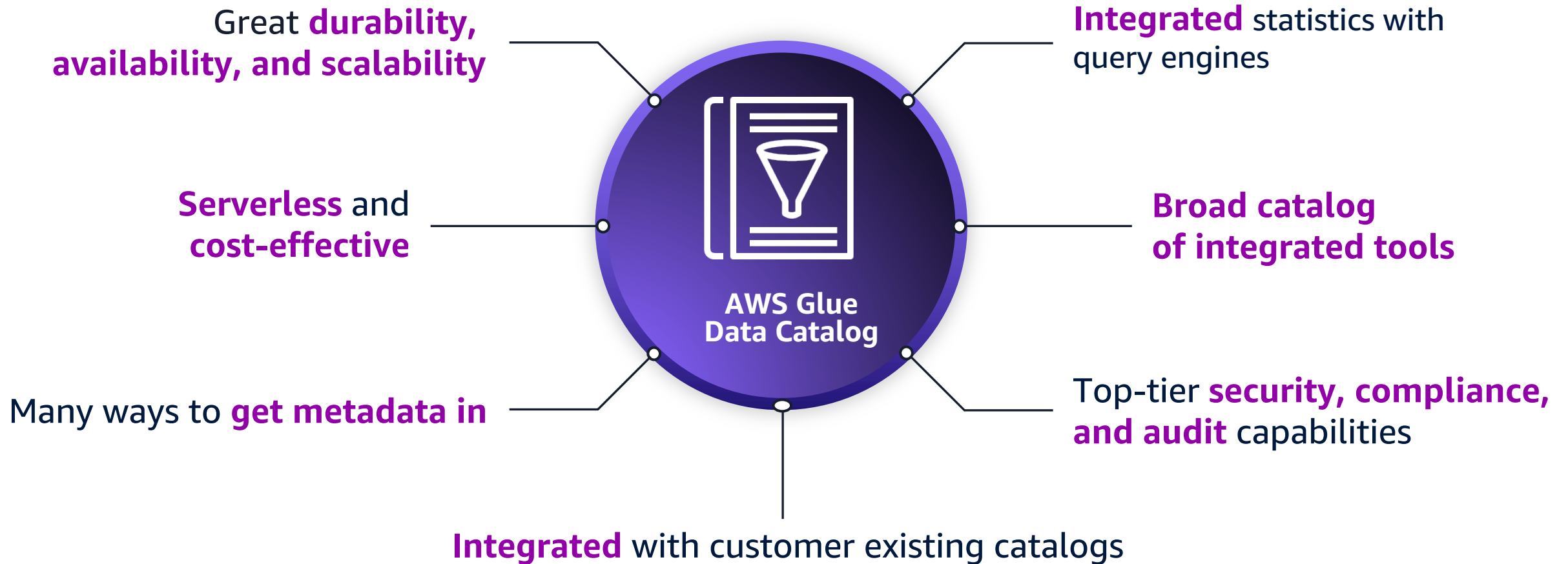
Decouple storage from compute

Choice of analytical and ML engines

Process data in place

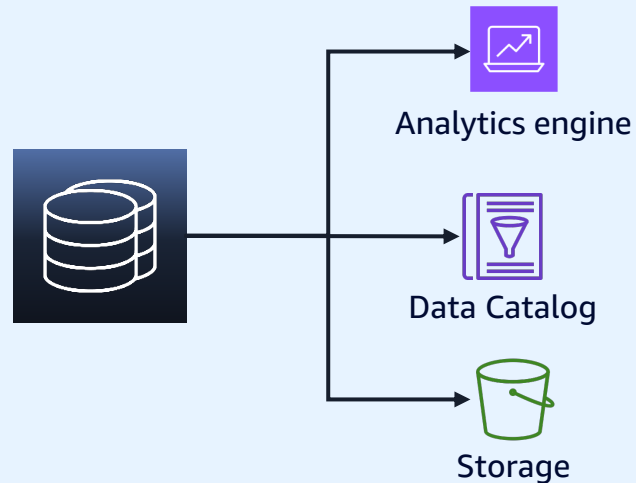
AWS Glue Data Catalog

An open, performant, and scalable technical metastore



AWS Lake Formation: Securing your data lake

Unifying permissions across the data lake stack

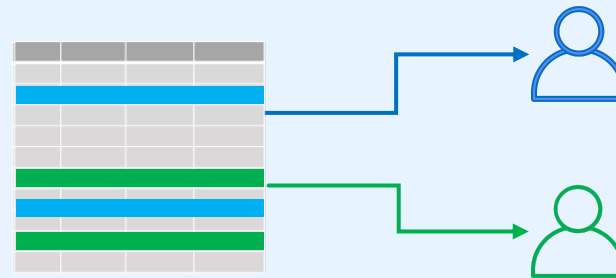


Split storage, metadata, and compute

Each system has different permissions

Syncing permissions is error-prone

Enforcing fine-grained permissions to restrict access

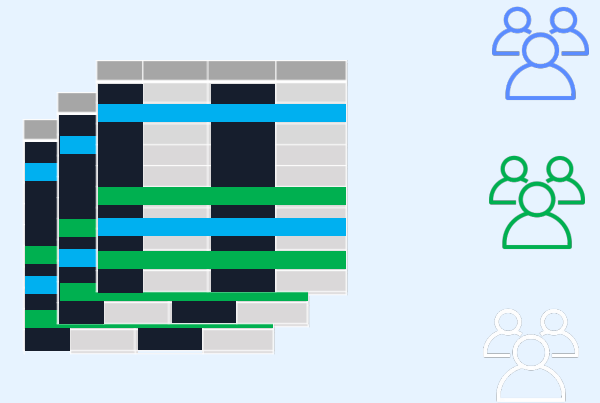


Data lakes contain a lot of data

Users should only access portions

Infrastructure that accesses data needs to only provide access to those with permissions

Scalable permissions to manage data and users



1,000s of DBs and tables

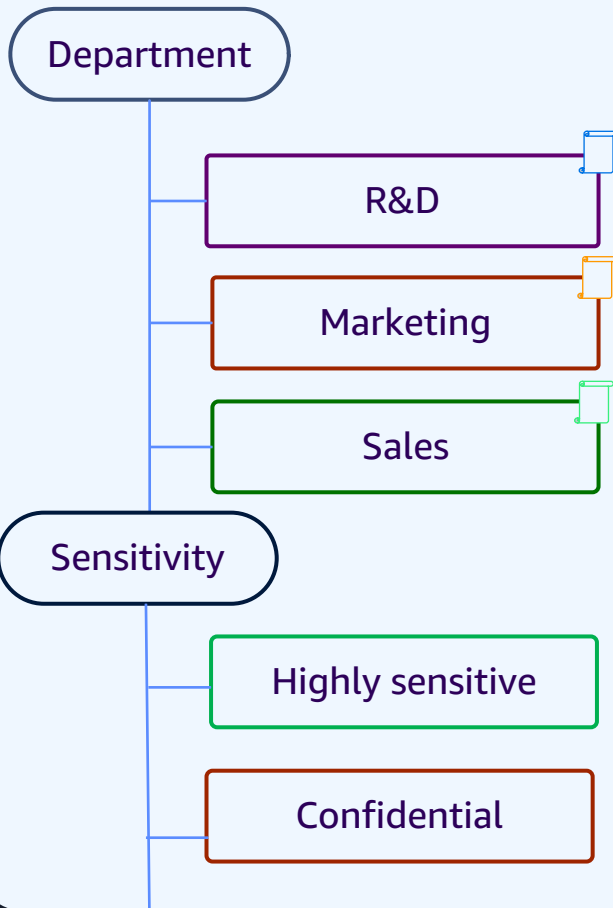
10,000s of users

New data sets added constantly

Data protection rules change constantly

Fine-grained access with tag-based controls

Develop a tag ontology



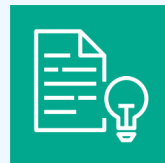
Automate LF-tag values with AI



Glue Data Profiling and Data Quality to dynamically identify tag values

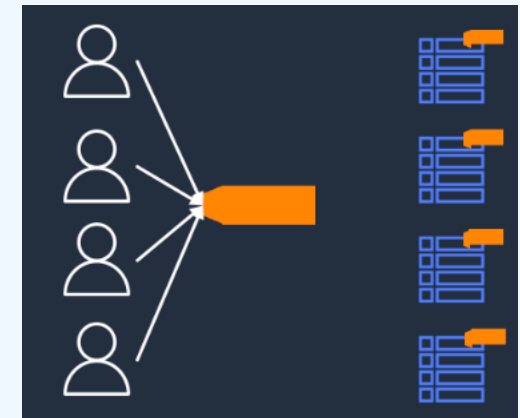


Glue sensitive data detection



Amazon Comprehend DetectPHI

Assign tags people to scale



Scale by applying permission on LF-tags and users get access with integrated engines

Amazon Athena: Explore data with SQL

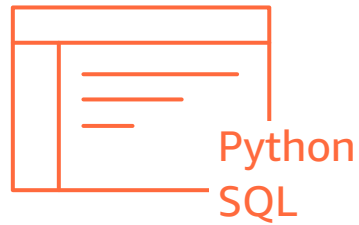


SIMPLE, INSTANT START

Serverless, no setup

Instant start, optimized
runtimes for fast
results

Point to S3 and
start querying

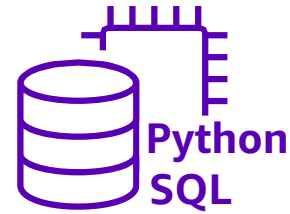


INTERACTIVE, ADVANCED ANALYTICS

Federated queries across
35+ data stores

Use PySpark ecosystem

Simplified notebooks on
console for PySpark



OPEN AND FLEXIBLE

ANSI SQL, Apache Spark

Multiple formats,
compression types, and
complex joins
and data types



COST EFFECTIVE

Pay only for what you use

SQL: Save **30%–90%** on
per-query costs
through compression

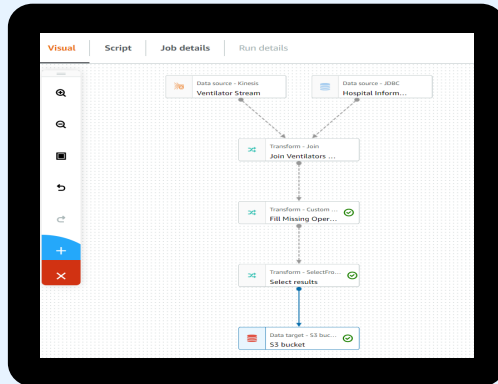
Python: minimize idle
compute charges

Designed for all data users



ETL developer

AWS Glue Studio



Use a rich visual interface

Choose from 250+ built-in transformations

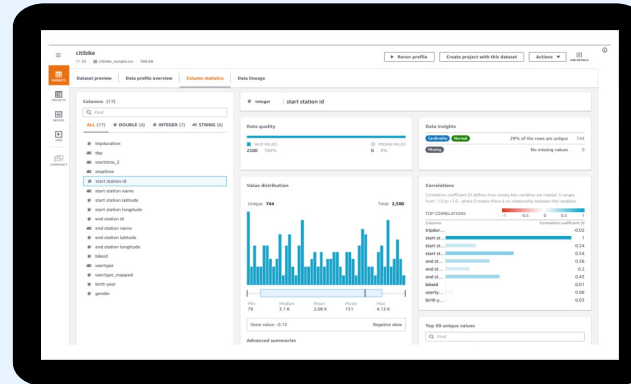
Profile data to understand data patterns and anomalies

Work on large datasets at scale



Business analyst | Data scientist

AWS Glue DataBrew



Run ETL jobs without writing code

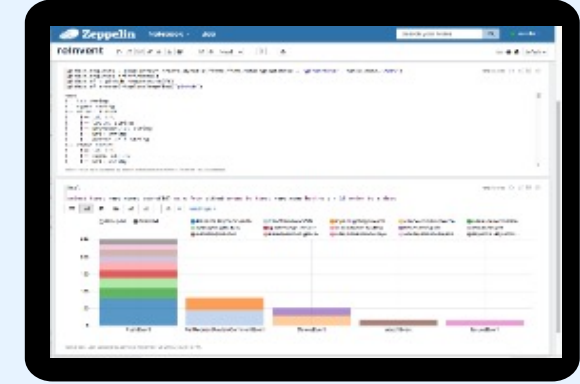
Monitor thousands of jobs in one place

Apply advanced transforms through code snippets



Data engineer

AWS Glue notebooks



Clean and normalize data with a rich visual interface

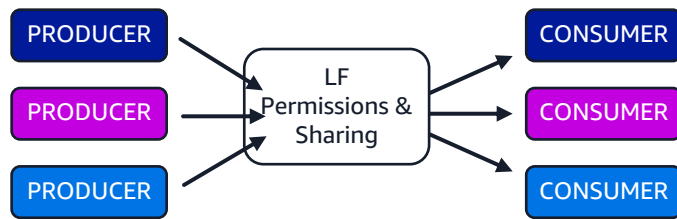
Choose from 250+ built-in transformations to automate tasks

Profile data to understand data patterns and anomalies

Zero copy data sharing models

AWS SUPPORTS DATA SHARING WITHOUT COPYING DATA GLOBALLY

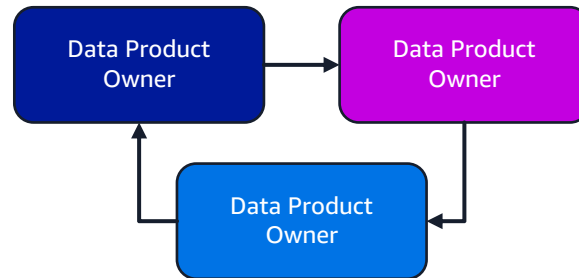
Hub and Spoke



Hub and Spoke Multi-Account

Share data across your organization and across regions

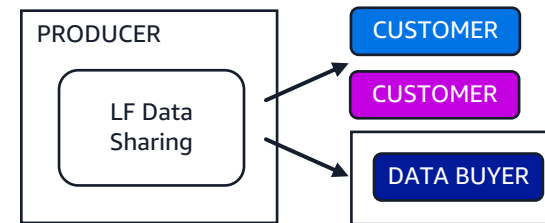
Data Mesh



Share data products via a Data Mesh

Share data with organizational autonomy by leveraging Lake Formation Sharing

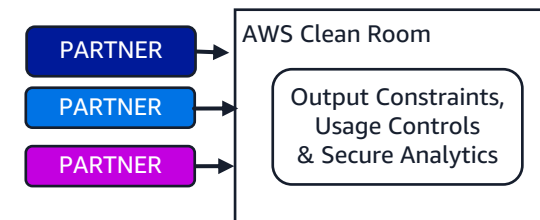
Business to Business



Multi organization Governed data sharing

Share data via Lake Formation or Amazon Data Exchange for Data Monetization

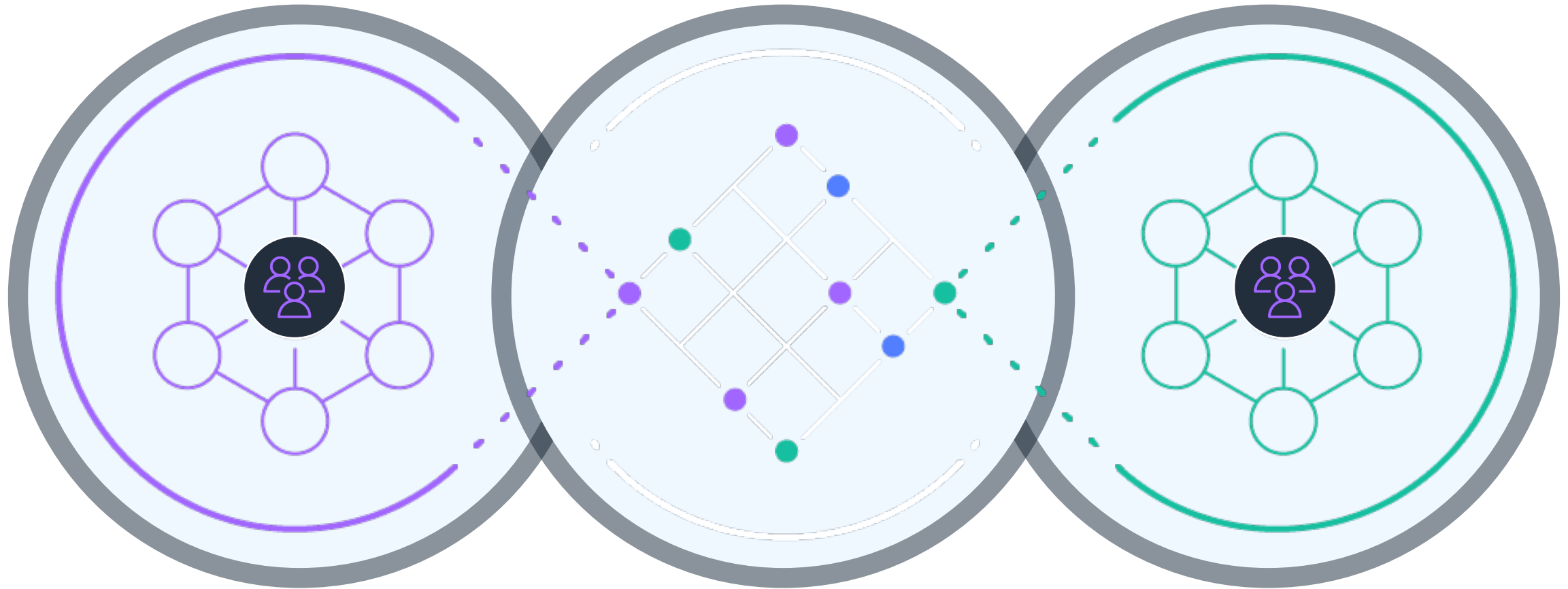
Partner Collaboration



Governed and Secure Clean Room

Data producers and data consumers collaborate

Catalog, discover, share, govern, and analyze across organizational boundaries
via centralized business friendly metadata

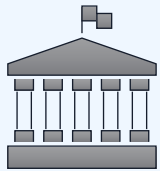


Data producers

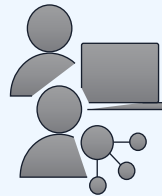
Amazon DataZone

Data consumers

Amazon DataZone



Govern data access
with fine grained controls
across organizational
boundaries



**Connect people and
data** through shared tools
to drive business insights

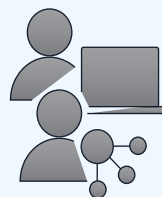


**Automate data
discovery** and cataloging
with ML and GenAI. Provide
data quality and **lineage**

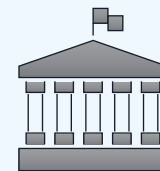
Community College Transfer Pathways



Community College
dual enrollment – SIS and
LMS systems



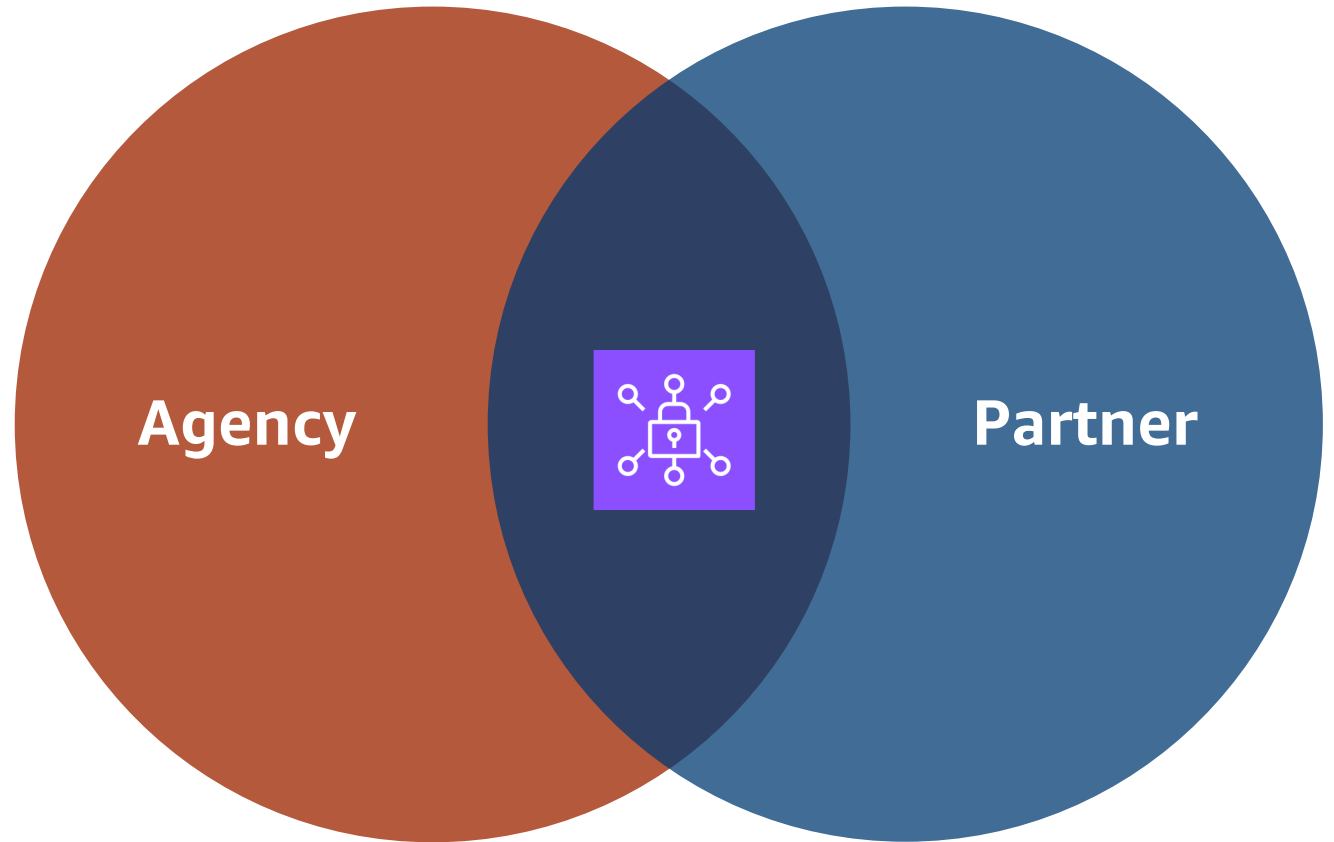
Advisors
can see student progress
across institutional systems



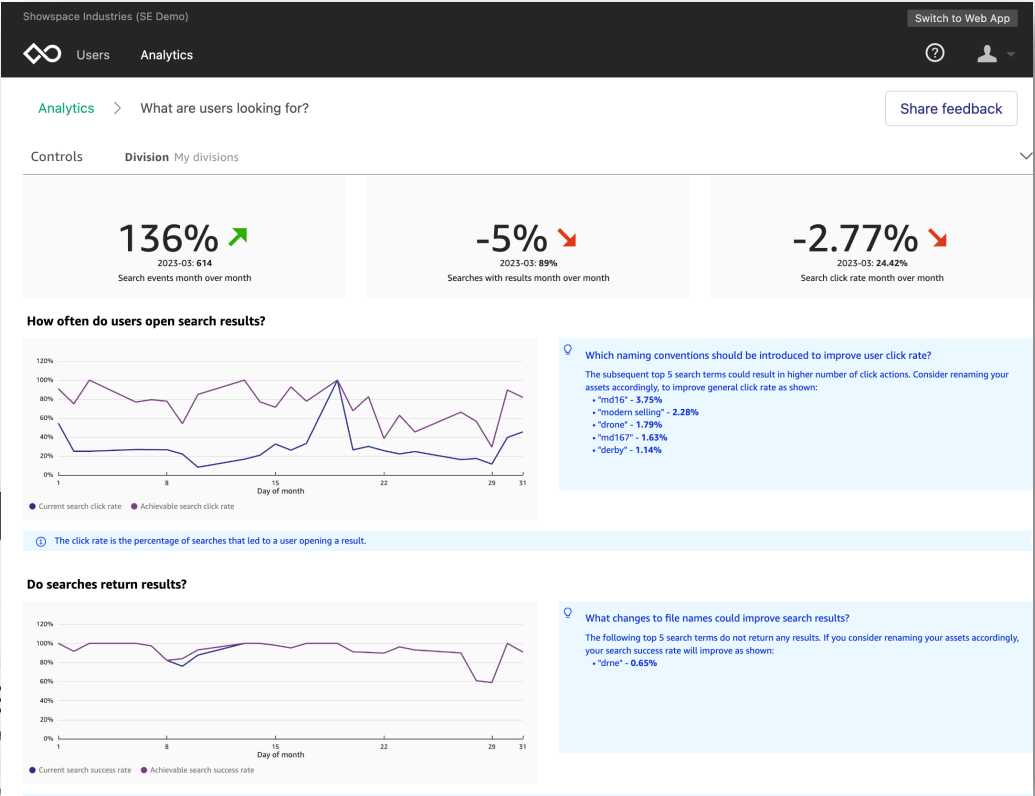
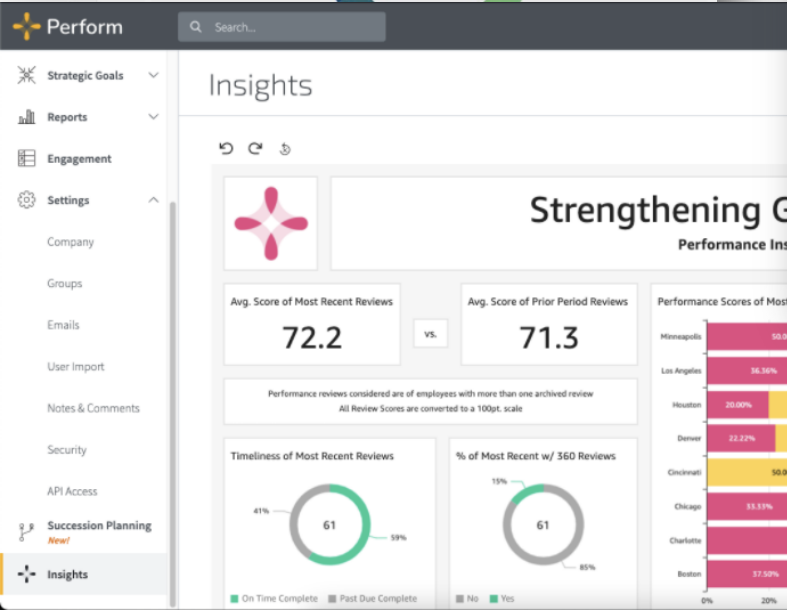
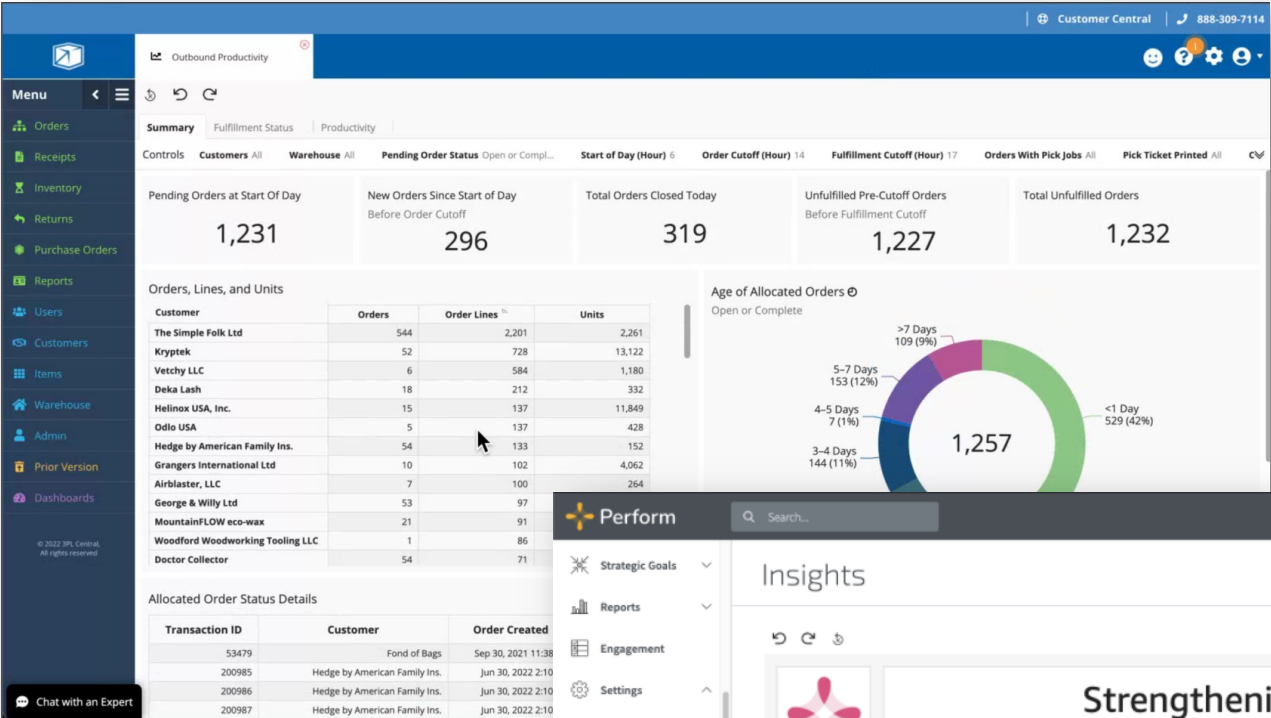
University System
dual enrollment programs
and pathways.

AWS Clean Rooms

Organizations can collaborate with their partners—without sharing underlying data



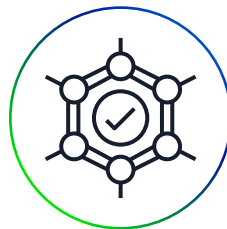
Amazon QuickSight: Data Visualization & Insights





Amazon QuickSight

UNIFIED BUSINESS
INTELLIGENCE AT
HYPERSCALE



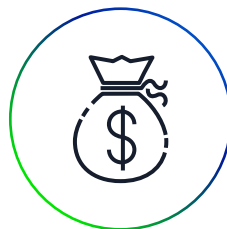
Unified BI for all your analytics needs



Consistent high performance with auto scaling



Augmented analytics with generative AI capabilities



Lower costs by paying for what you need

AI-powered dashboard authoring experience

A NEW DASHBOARD BUILDING EXPERIENCE POWERED BY GENERATIVE BI

Build visuals

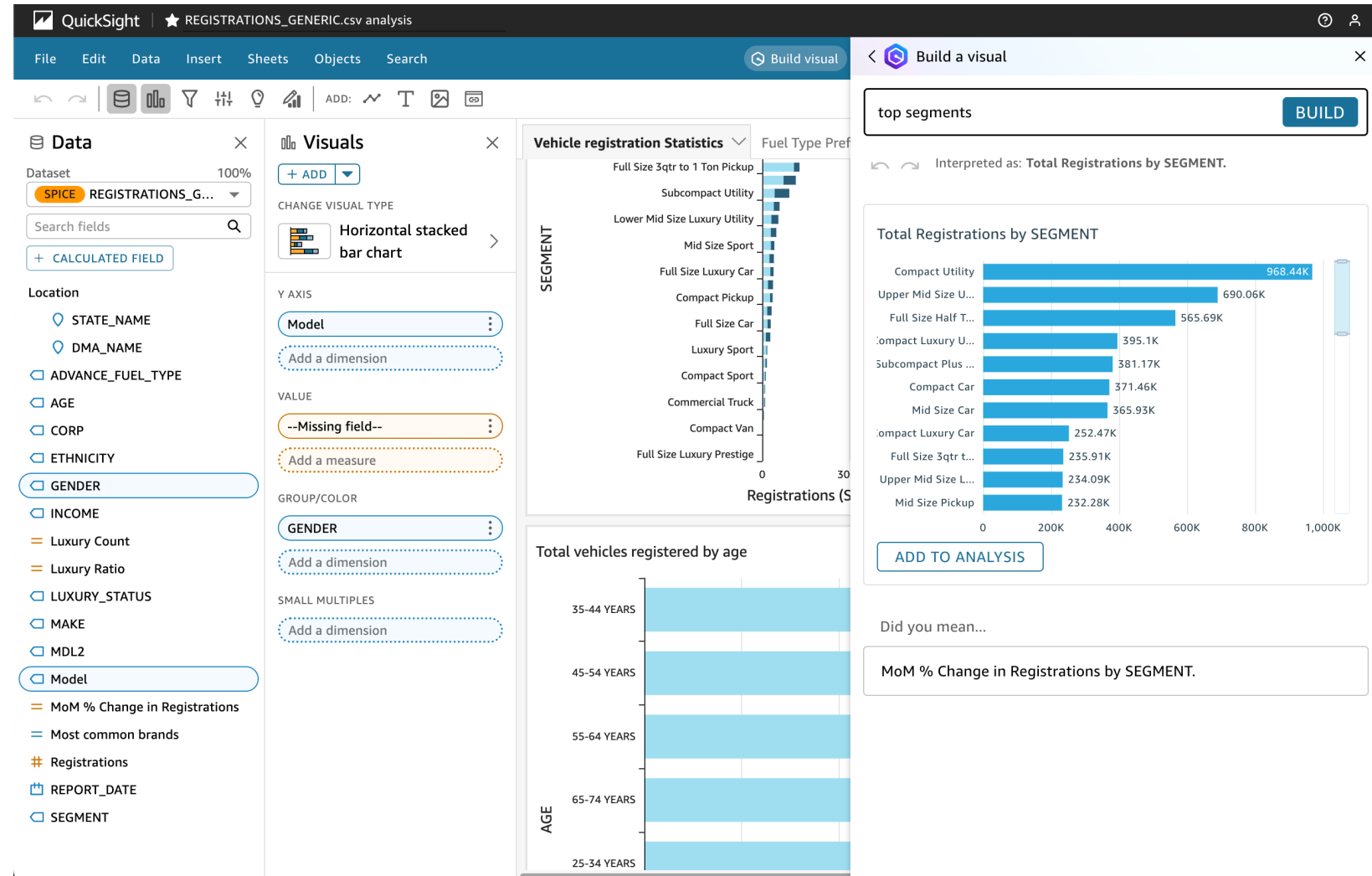
Use natural language to quickly build visuals for dashboards and reports

Create calculations

Build calculations using natural language without looking up or learning specific syntax

Refine visuals

Quickly update visuals by describing desired formats using natural language



AI-assisted storytelling

IMPACTFUL DATA STORYTELLING TO DRIVE ACTIONS

Interpret data for others

Help others derive meaning from data and reach conclusions to drive decisions

Generate stories using AI

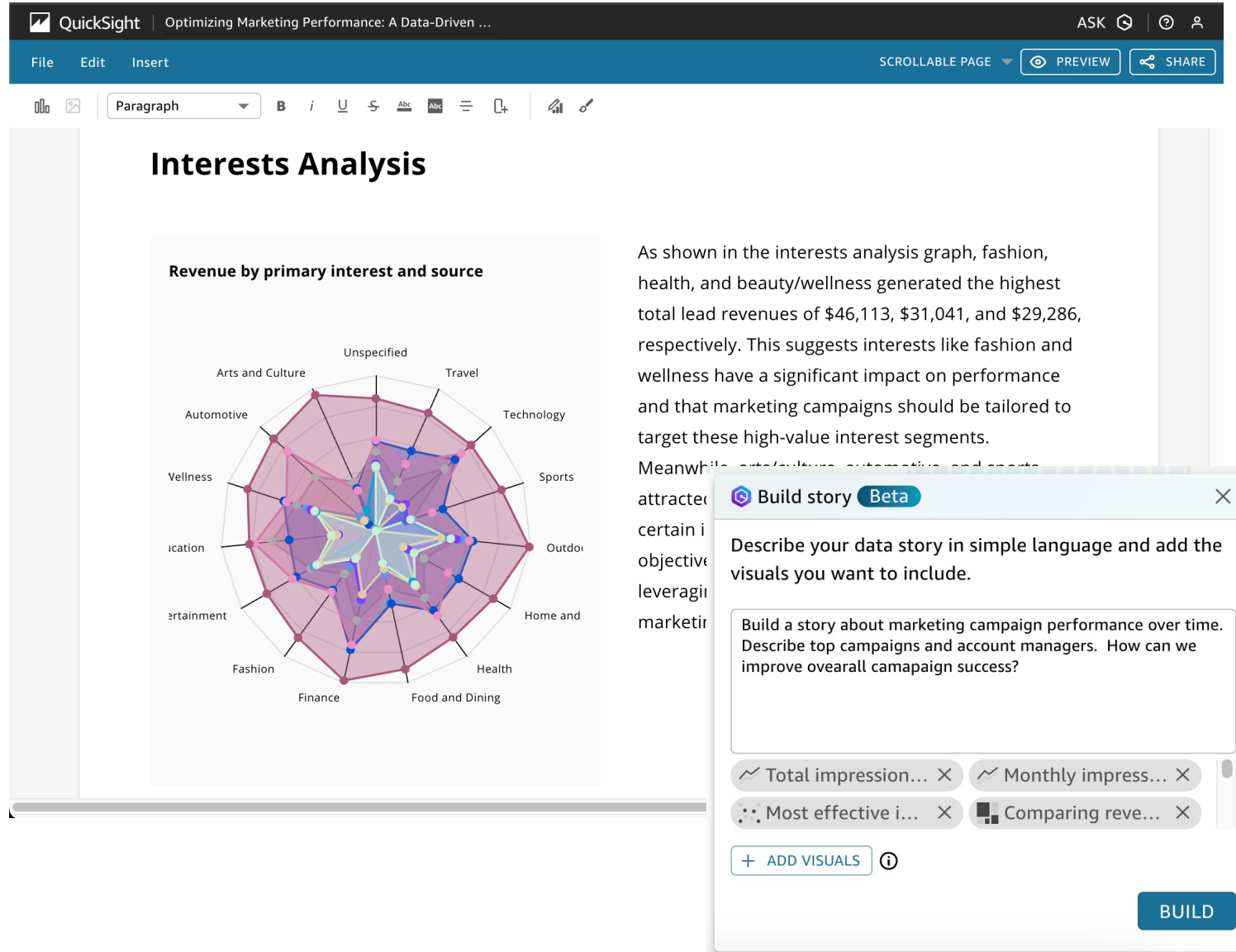
Produce cohesive, powerful, and insightful narratives by analyzing only a few words

Create refined content

Control AI verbosity, customize narrative text, and apply stunning visual themes to bring content

Share up-to-date governed data

Quickly update and disseminate data at any time





Amazon Q



Amazon QuickSight

Amazon SageMaker Studio

brings tools for every step of the ML lifecycle under one unified visual user interface



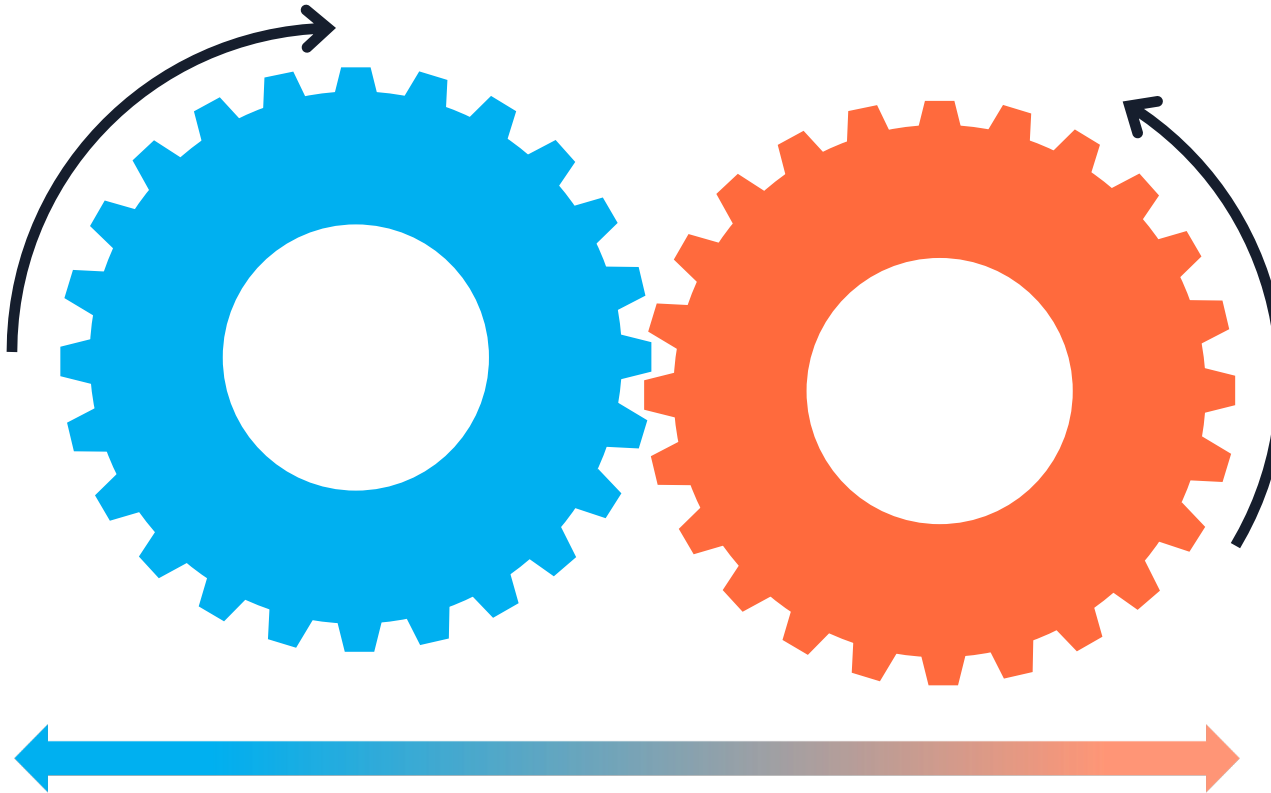
How SageMaker Canvas empowers teams

Accelerate data science teams

Accelerate tedious tasks using no-code machine learning to get to the desired outcomes faster

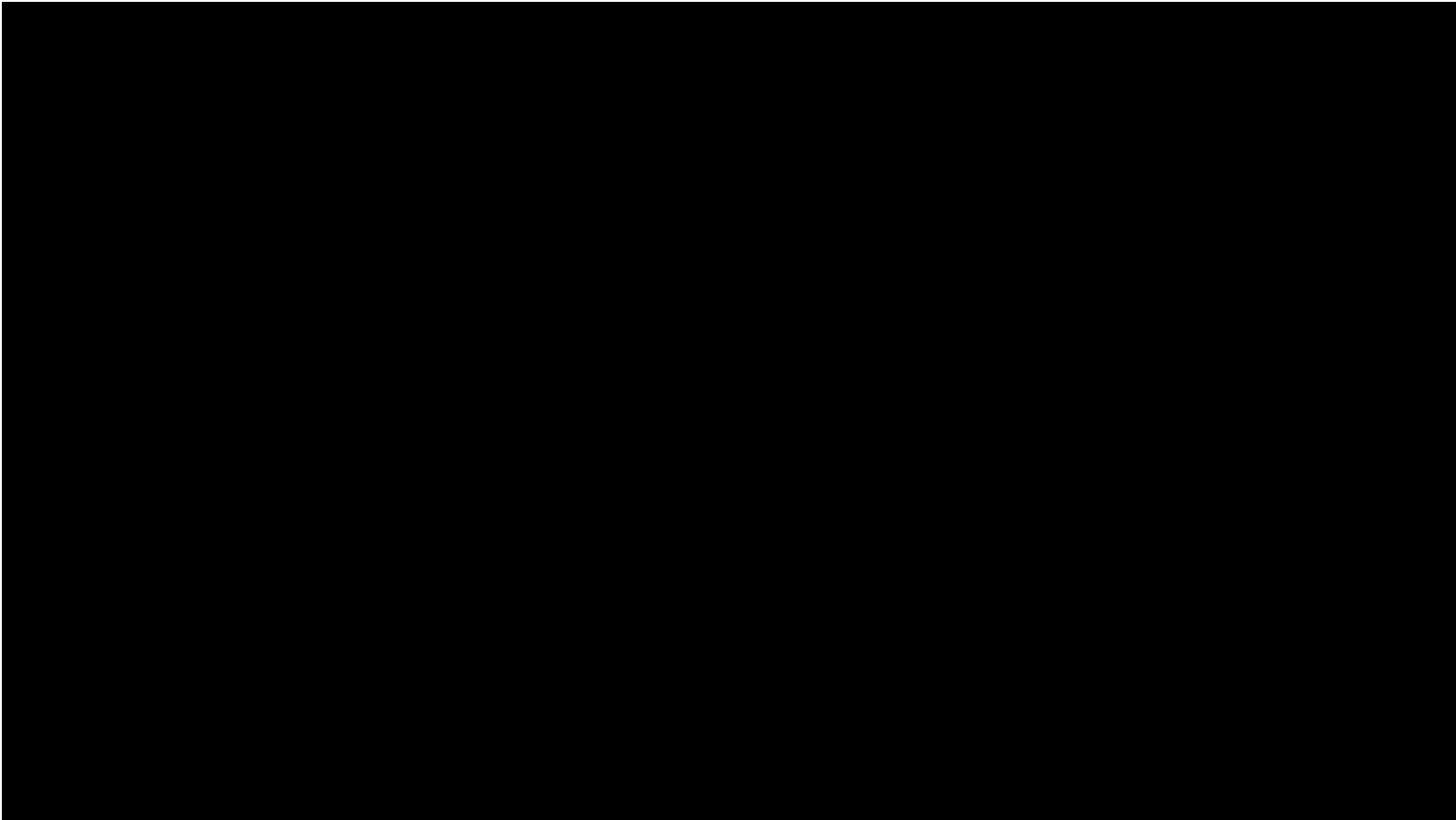
Enable business teams

Give business teams the ability to do ML without any code, scaling the number of people who can create ML-powered insights, forecasts, and predictions



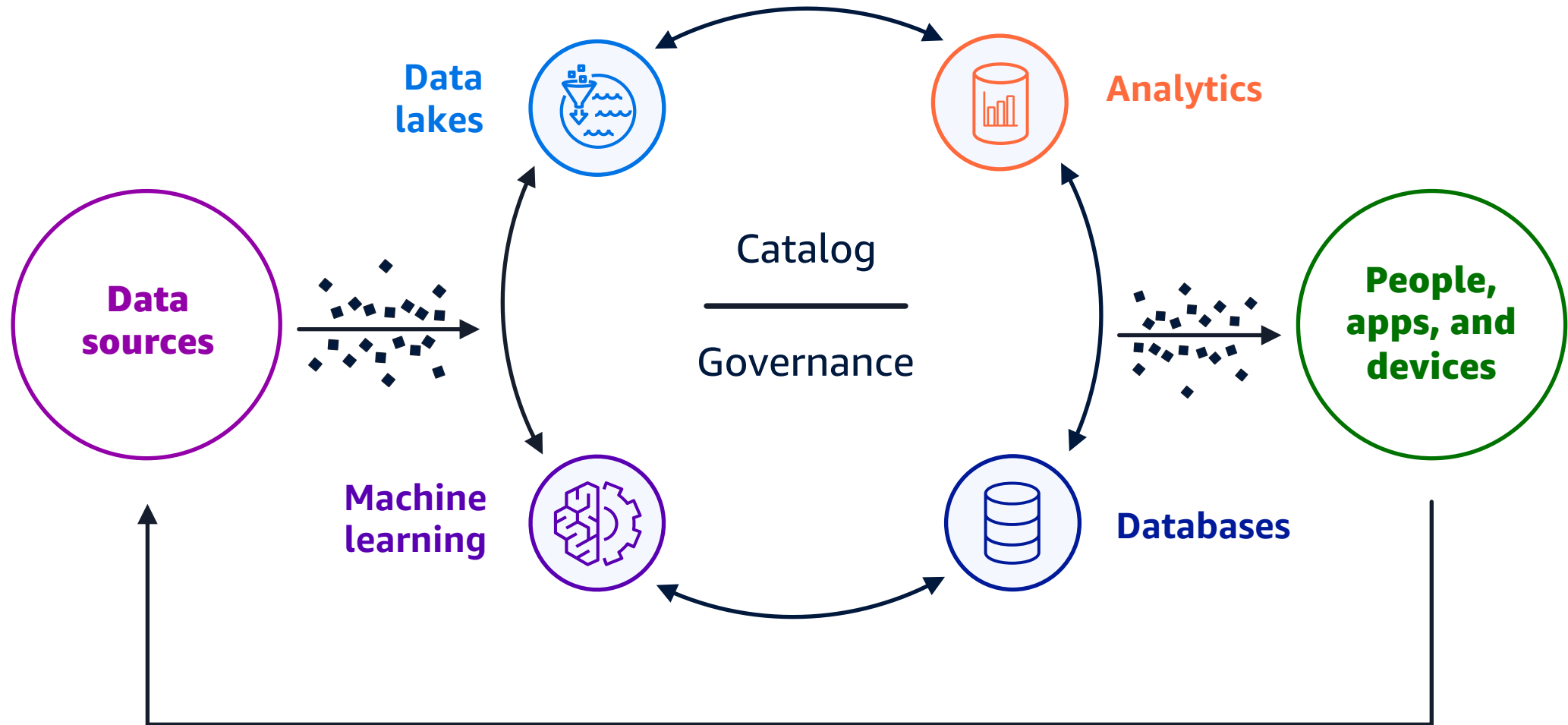
Collaborate together

SageMaker Canvas has several points of collaboration, making it easy for business teams to use models built by data scientists, for data scientists to update models built by business analysts



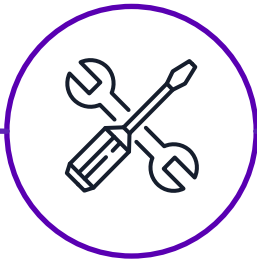
UMD GenAI Example?

Data democratization flywheel



Get started

BUILD WITH US



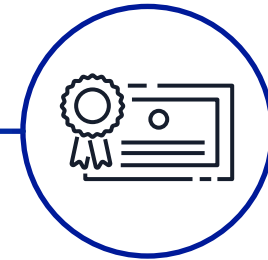
ML Solutions Lab
AWS Professional Services
AWS Immersion Day
Data-Driven Everything
Migration Assistance Program

BUILD WITH PARTNERS



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

UPSKILL YOUR TEAMS



AWS Training and Certification
ML Embark Program

Thank you!

Greg Grief

ggrieff@amazon.com

Please complete the
survey for this session



Track: Data & Analytics

Session: Democratizing your
organization's data analytics
experience

