

HARMONIZING DATA HORIZONS: A UNIFIED DATABRICKS DATALAKE JOURNEY

A new holistic solution to provide data consumers with a single source of truth, meeting today's scalability and compliance requirements

ABOUT ME



VIJAY KUMAR NAGARAJ
PRODUCT OWNER
UNIFIED DATA LAKE



LINKEDIN PROFILE
SCAN TO CONNECT

ABOUT ME



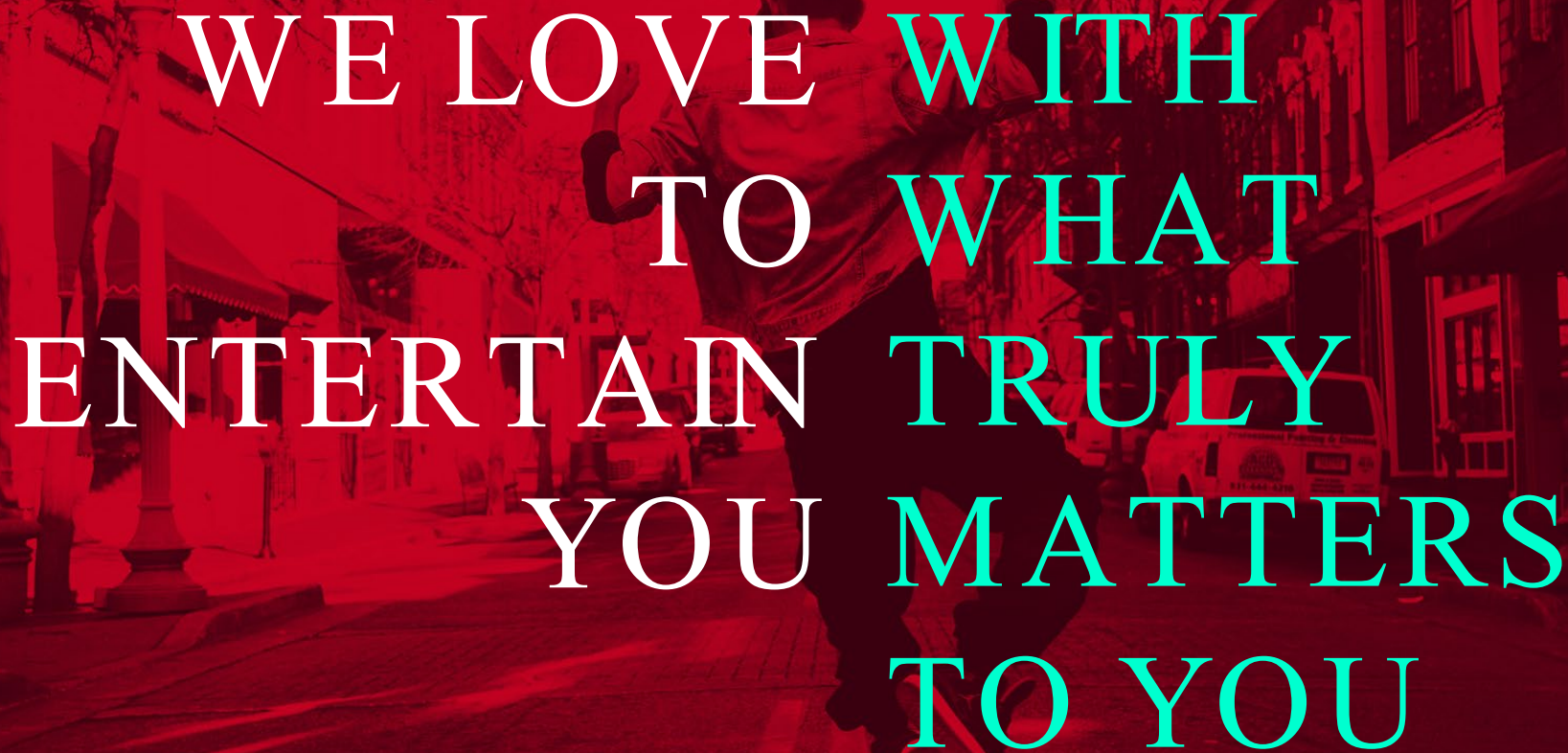
ProSiebenSat.1
Media SE



Along with

b.telligent

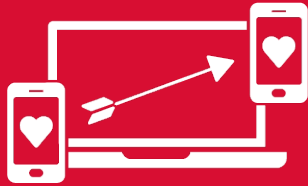
smart data. smart decisions.



WE LOVE WITH
TO WHAT
ENTERTAIN TRULY
YOU MATTERS
TO YOU

HOME OF LEADING ENTERTAINMENT AND DIGITAL CONSUMER BRANDS IN DACH REGION

DATING & VIDEO



- **Our Dating & Video segment consists of ParshipMeet Group** which grew out of a successful Commerce & Ventures business.
- With ParshipMeet Group, we have created a **leading globally operating player** in the dating industry.
- We are building an ecosystem **across social entertainment, online dating and matchmaking.**

ENTERTAINMENT



- Our Entertainment segment is in **the center of our strategy.**
- We **unite linear and digital entertainment platforms** with our **content, distribution and sales** business.
- **Our guiding principle is “Viewers and Users First”:** We gear our offerings entirely to their needs, providing local and relevant content that we broadcast live and on-demand across all platforms in a targeted way.

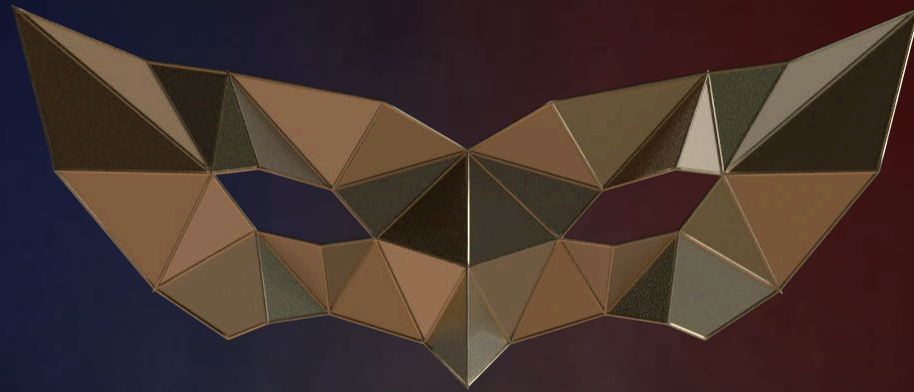
COMMERCE & VENTURES



- In the Commerce & Ventures segment, we are bundling our **growth businesses** that we are building up and making successful **with our media services and the power of our brands.**
- We **invest advertising time** in young digital companies and in exchange receive shares in the companies themselves or in their revenues
- We focus on **consumer brands with long-term structural growth potential** and **synergies** with our **Entertainment business.**

GERMANY'S NEXT
topmodel
by Heidi Klum

**GERMANY'S NEXT
TOPMODEL**



THE MASKED SINGER

The voice of germany

The biggest entertainment brands

Galileo

DAS
DUELL
der WELT
TEAM JOKO GEGEN TEAM KLAAS

ARKAS
E

das GROSSE
BACKEN

1%
QUIZ
WER CLEVER IST
DÜBELT SCHLÄGT

WER
STIEHT
MIR
DIE
SHOW

Sweet & Easy
ENIE BACHT

SCHLAG
DEN
STAR

99
EINE-R
HILFT SIE ALLE

SAT1
frühstücks
TODTWEISER

ran

the taste

THE
MASKED SINGER

GERMANY'S NEXT
topmodel
by Heidi Klum

THE
VOICE
of Germany

THE
VOICE
KIDS

Zervakis &
Opdenhövel.
Live.

tv
total

99
EINE-R
jerks.

INTIMATE.

THE
GREAT
FIGHT NIGHT

SLAVIK
AUF STAATS
NACKEN

IMMER 5 TRYBACKS
TEAMMATES

BAYWATCH
BERLIN
MIT KLEIN HEUFER, BÜHLER, UND ANDEREN

MORD
AUF
EX

MORD
LUST

APOKALYPSE
& FILTERKAFFEE

As much
AS WE WE ALSO
LOVE TO LOVE OUR
ENTERTAIN DATA
YOU

CHALLENGES AND DATA COMPLEXITY



CHALLENGES



Data Silos and Fragmentation



Complex ETL Processes



Inconsistent Data Quality



Data Security and Compliance



Scalability and Performance

DATA COMPLEXITY

Source Systems

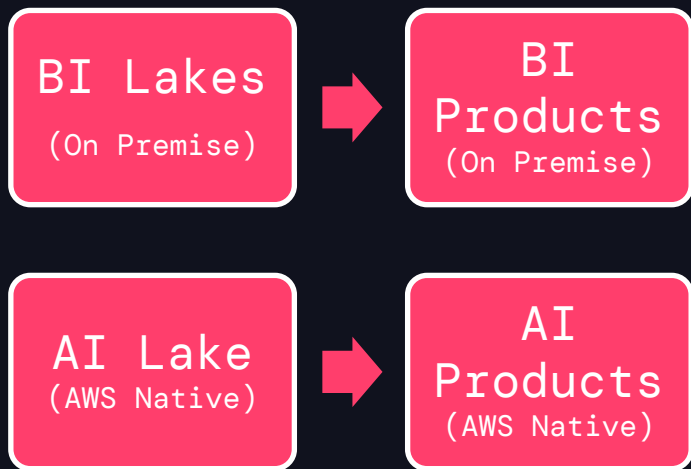


Source File Types

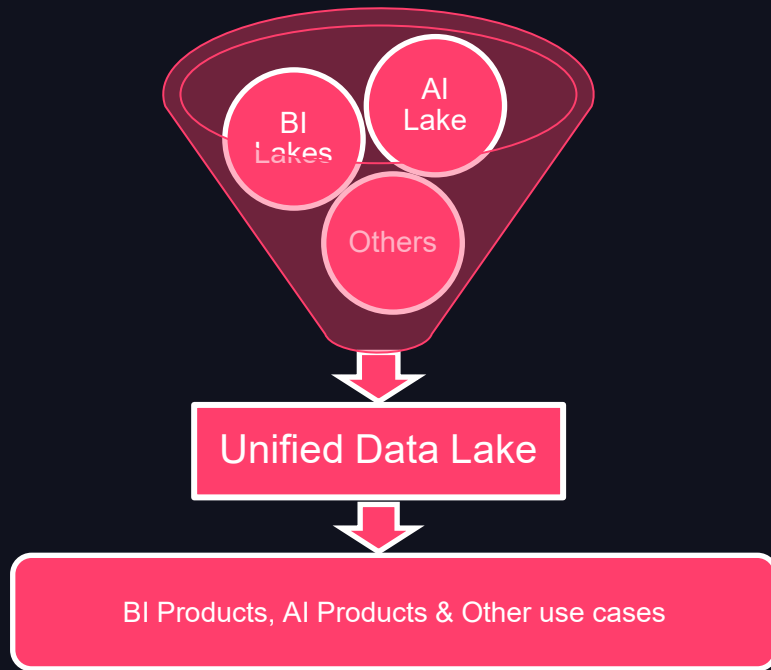


MIGRATION CONTEXT

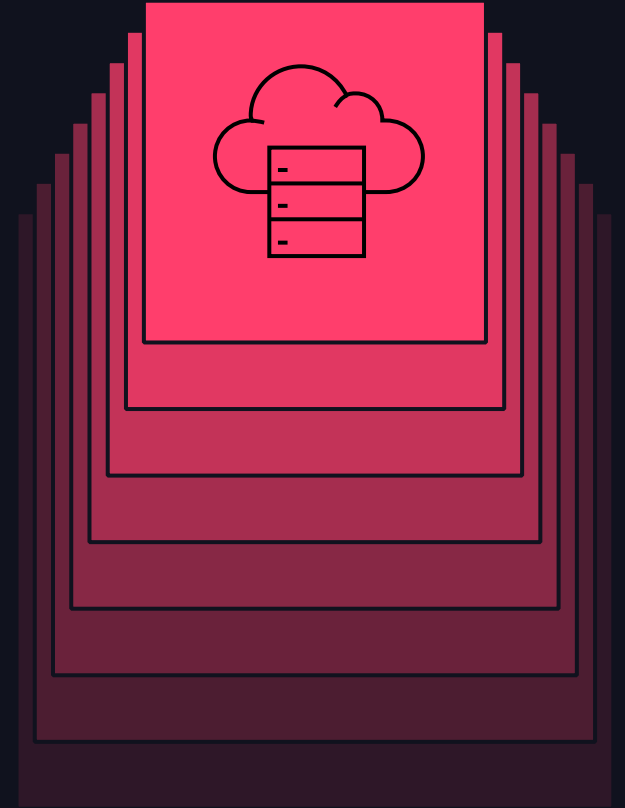
Then



Now



CLOUD MIGRATION STRATEGY AND PLATFORM SELECTION



REBUILD & REARCHITECT

with



databricks



DELTA LAKE

SCALABILITY & PERFORMANCE

SECURITY AND ACCESS CONTROL

UNIFIED DATA GOVERNANCE & COMPLIANCE

SUPPORT FOR MULTIPLE USE CASES

ACID (Atomicity, Consistency, Isolation, Durability) transactions and schema enforcement make the lake more reliable and performant

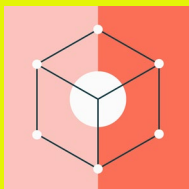
Seamlessly scale to accommodate growing data volumes and user loads.

Cloud-based solutions offer elastic scaling, ensuring consistent performance.

Databricks provides robust security features, including role-based access control, encryption, and network isolation in combination with AWS

Unity Catalog removes complexities in managing cloud native IAM based data & access control

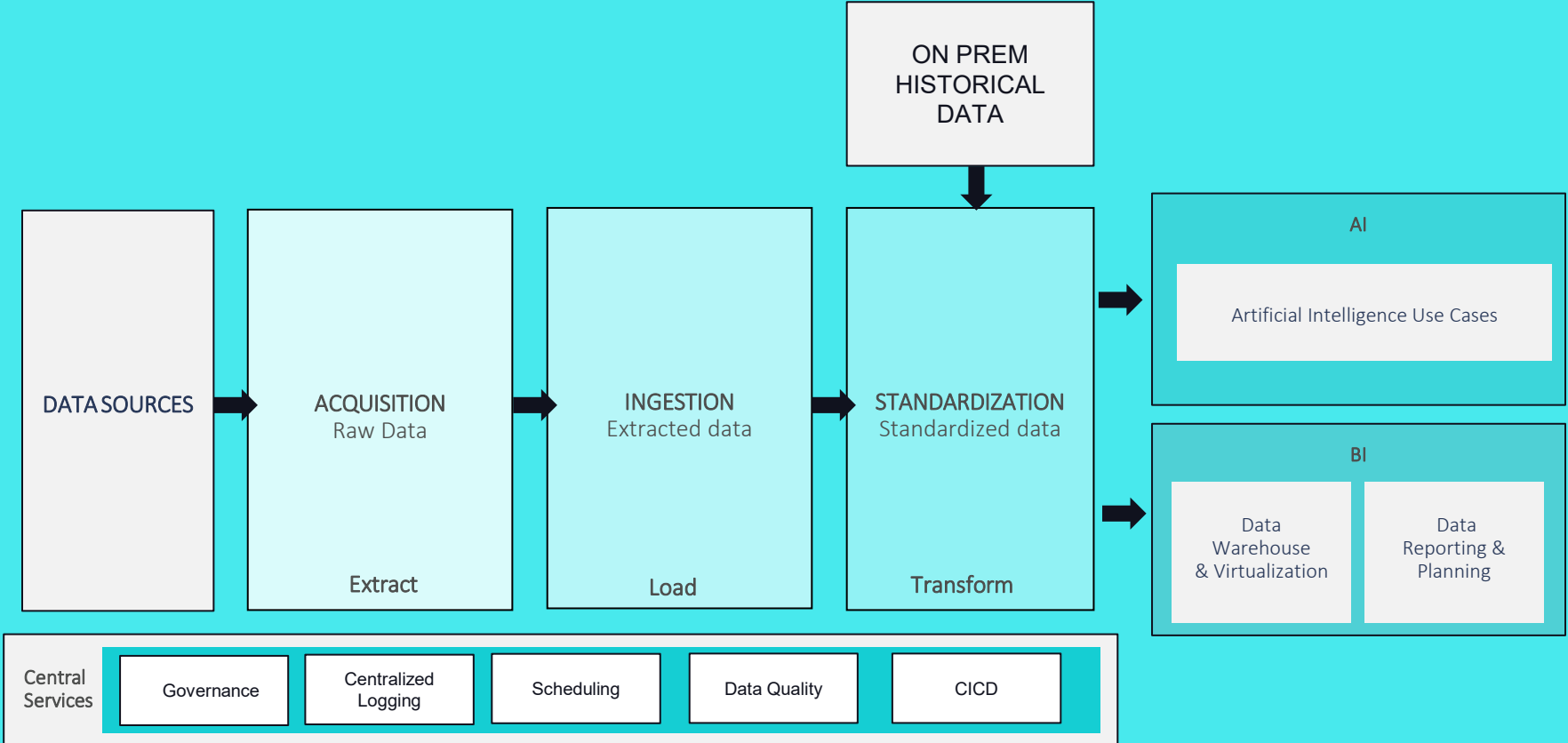
Databricks supports Data Science (GenAI), Data Engineering and BI use cases within an integrated platform



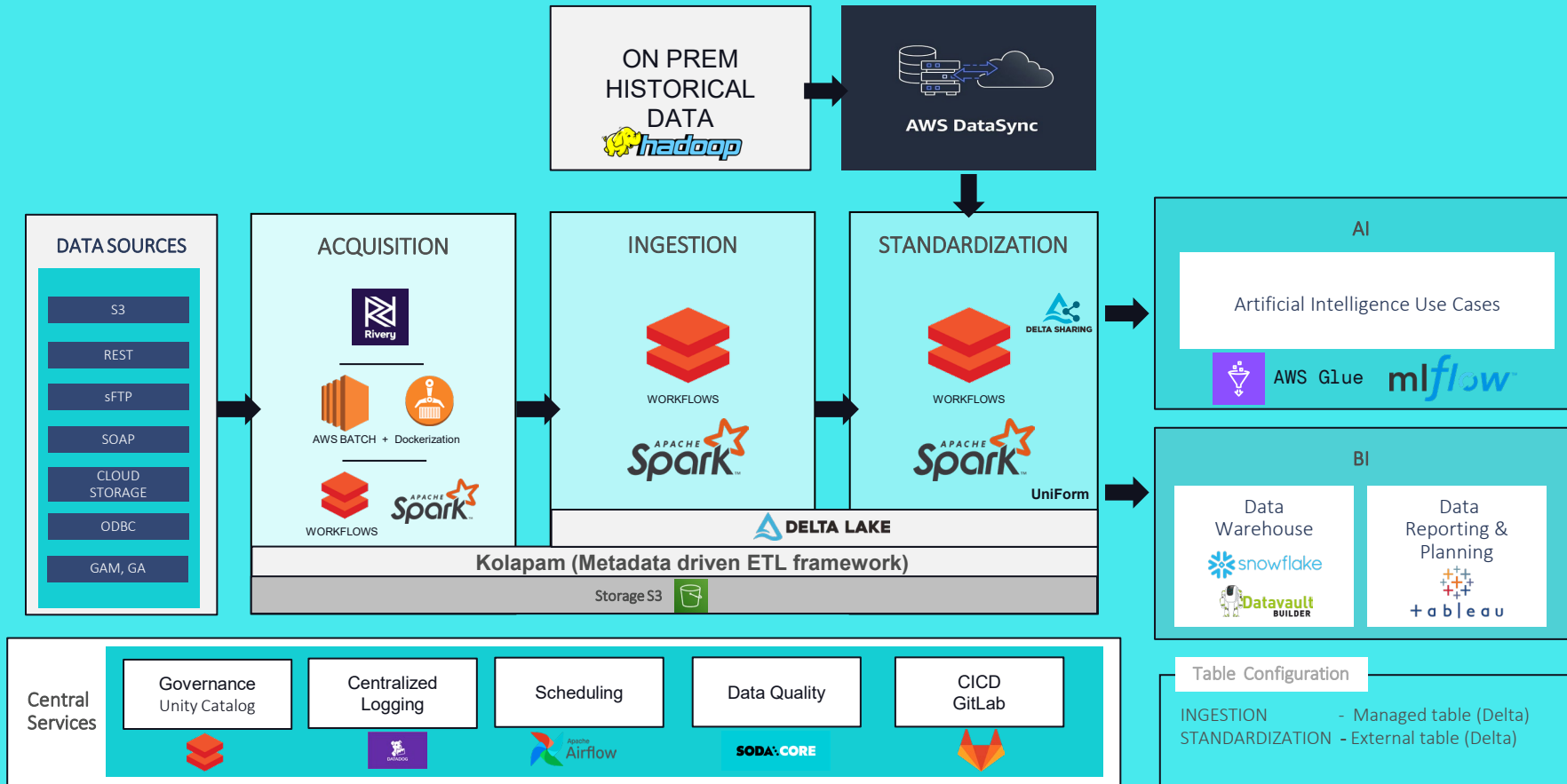
ARCHITECTURE



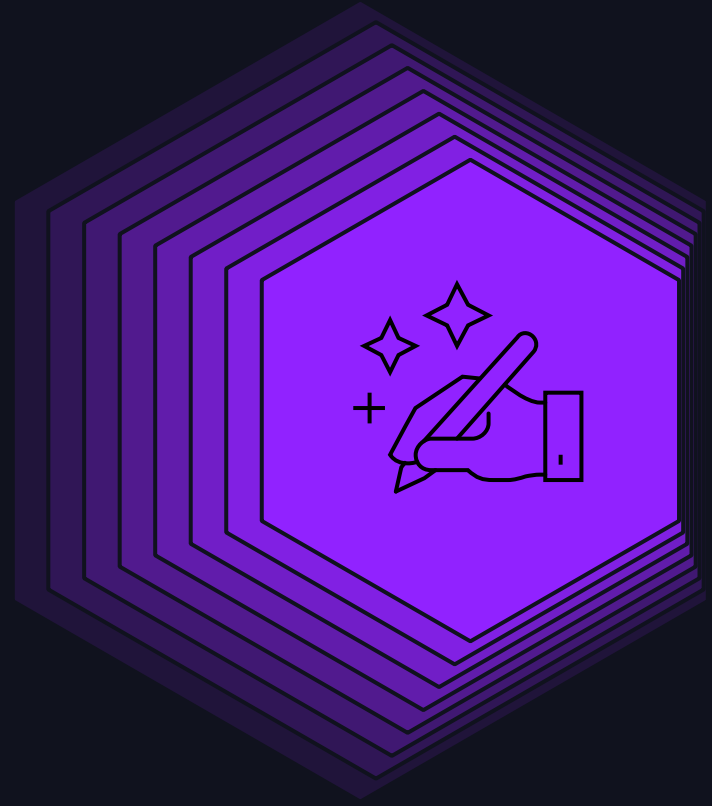
UNIFIED DATA LAKE ARCHITECTURE



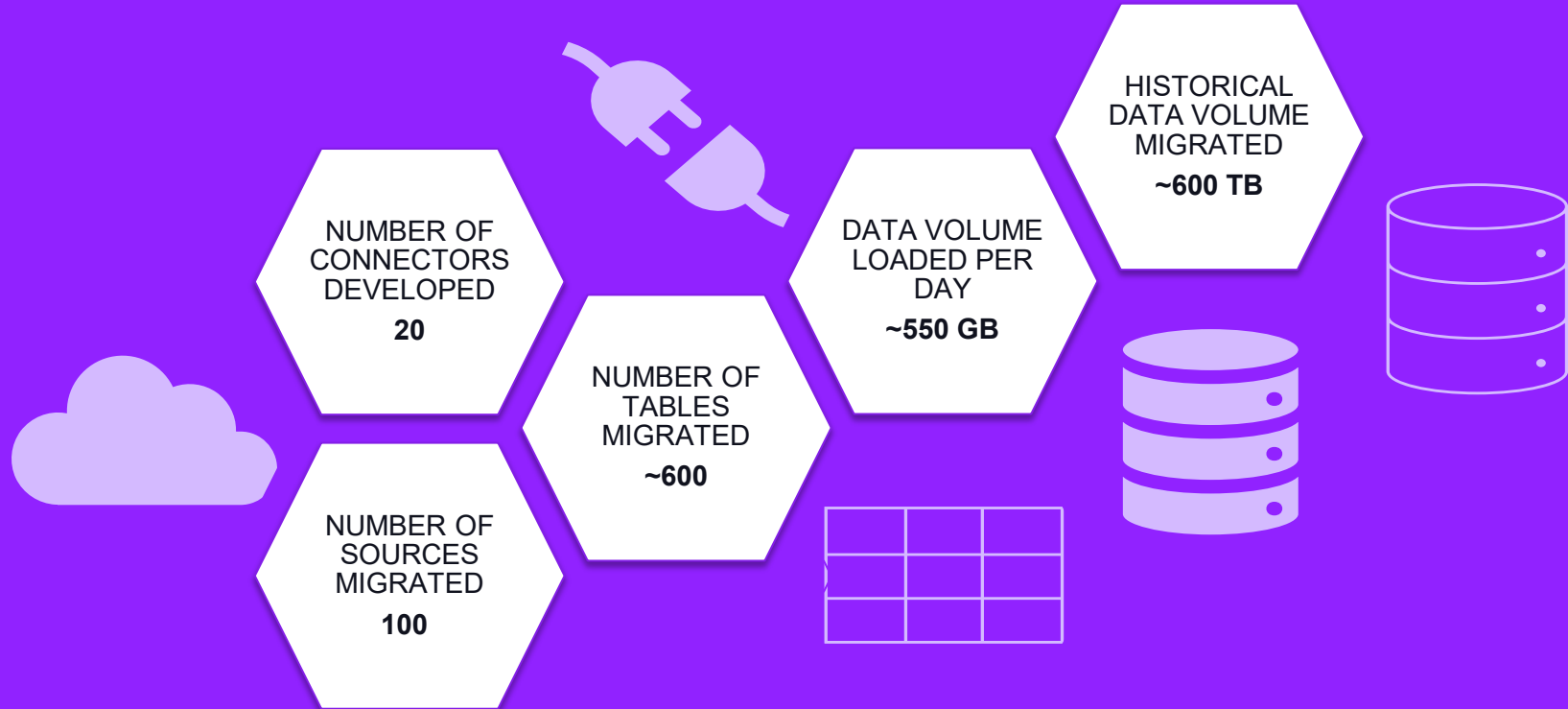
UNIFIED DATA LAKE ARCHITECTURE WITH SERVICES



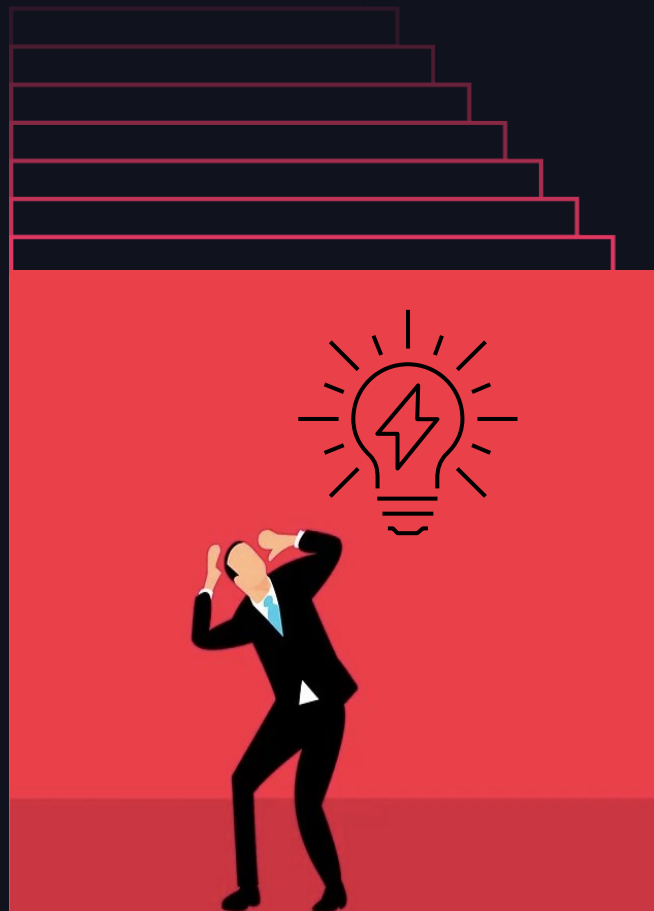
REAL TIME METRICS



REAL TIME METRICS



LESSONS LEARNT



UNITY CATALOG

EARLY ADOPTERS

IAC

CICD

DB CLUSTERS

SHARED METASTORE

Unity Catalog implementation from the very beginning made governance easier

Identify early adopters and work closely to reduce time to a usable MVP

Infrastructure as Code setup for Databricks and AWS resources using Terraform

CICD setup for faster deployments

Use Databricks multi node clusters for all compute intensive data loads for high scalability

Use Databricks fine grained access control for cross workspace user access to catalogs, schemas and tables with shared metastore



DATA+AI SUMMIT

