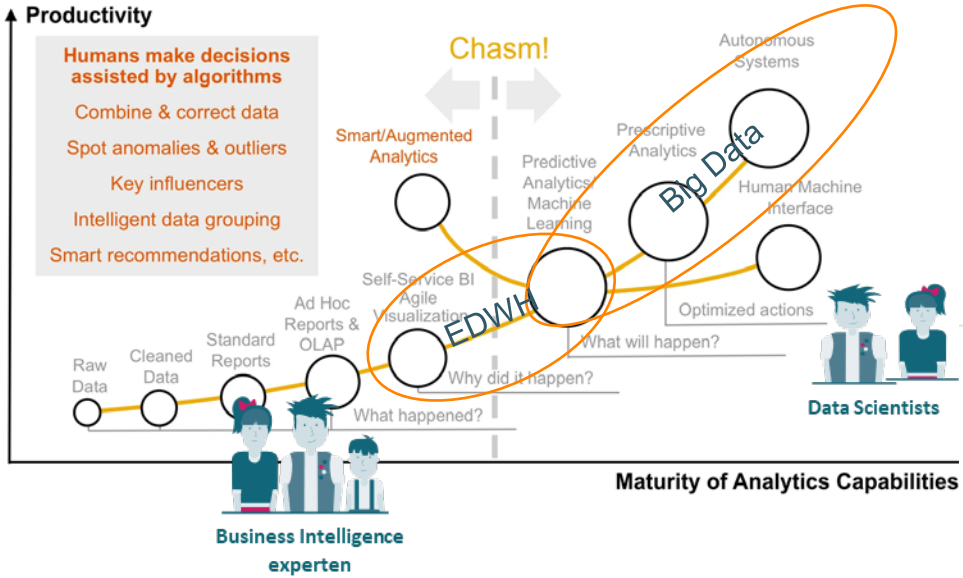
CLOUDERA

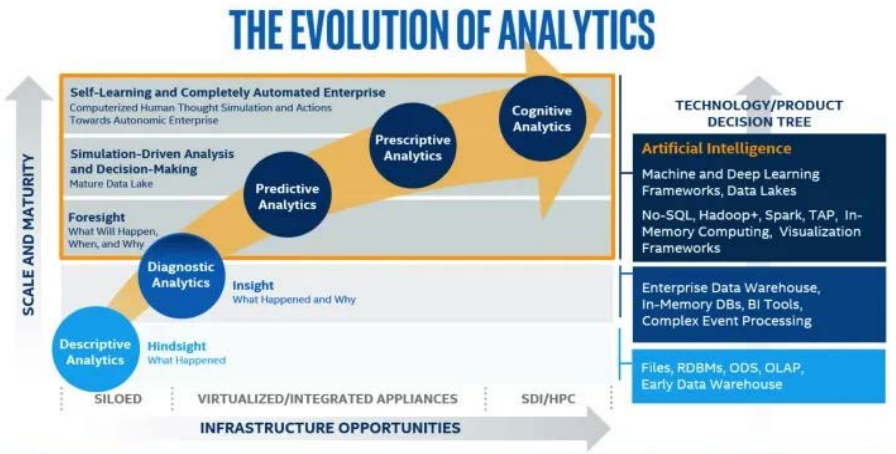
Wie die Wiener Linien ihre operationalen Prozesse mit Big Data unterstützen

Martin Dusek-Lippach (Wiener Linien) &
Jörg Gradwohl (Cloudera)

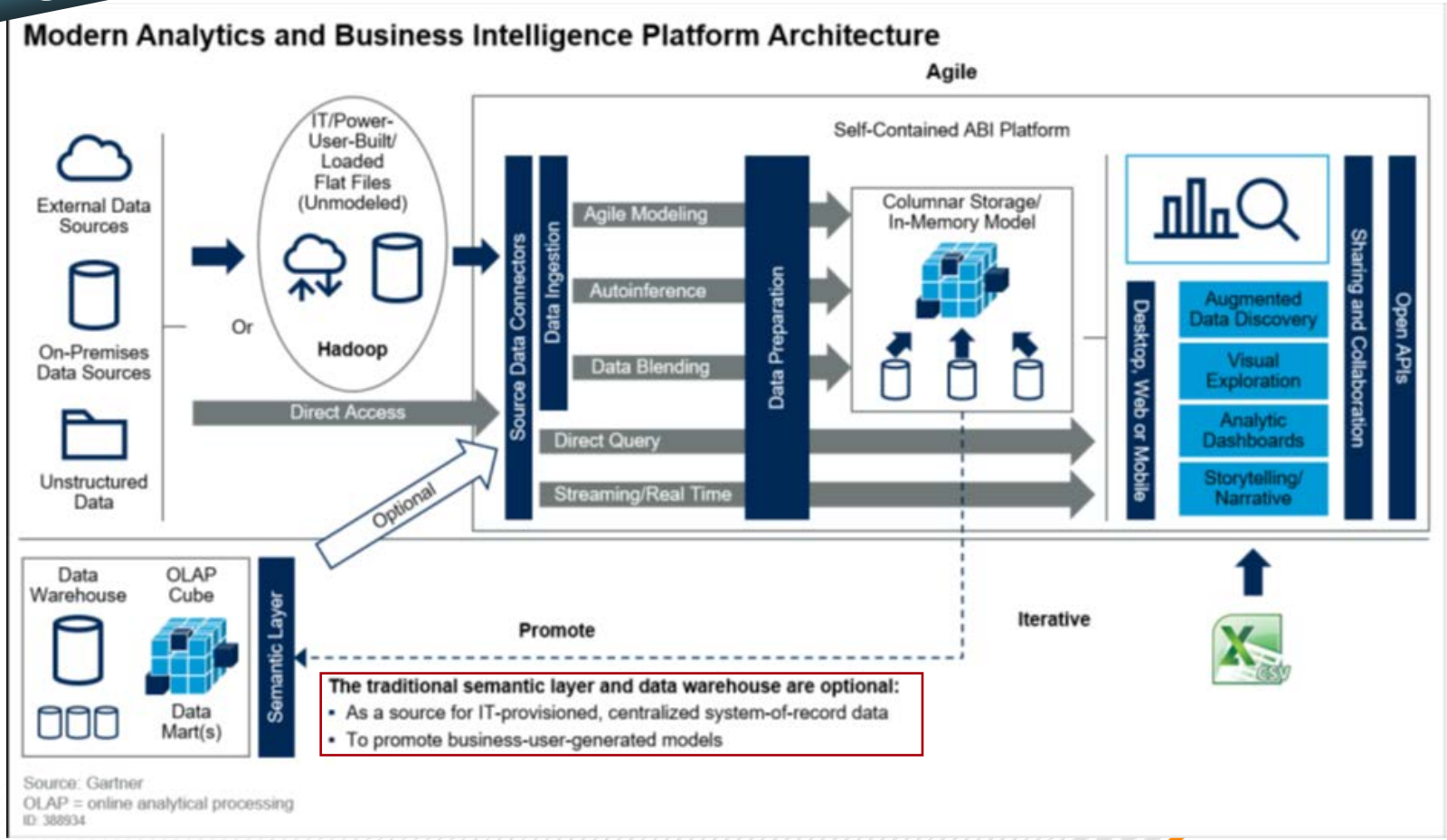
Maturity of Technology in the data space in 5 or more steps



Source: DELL EMC



Gartner

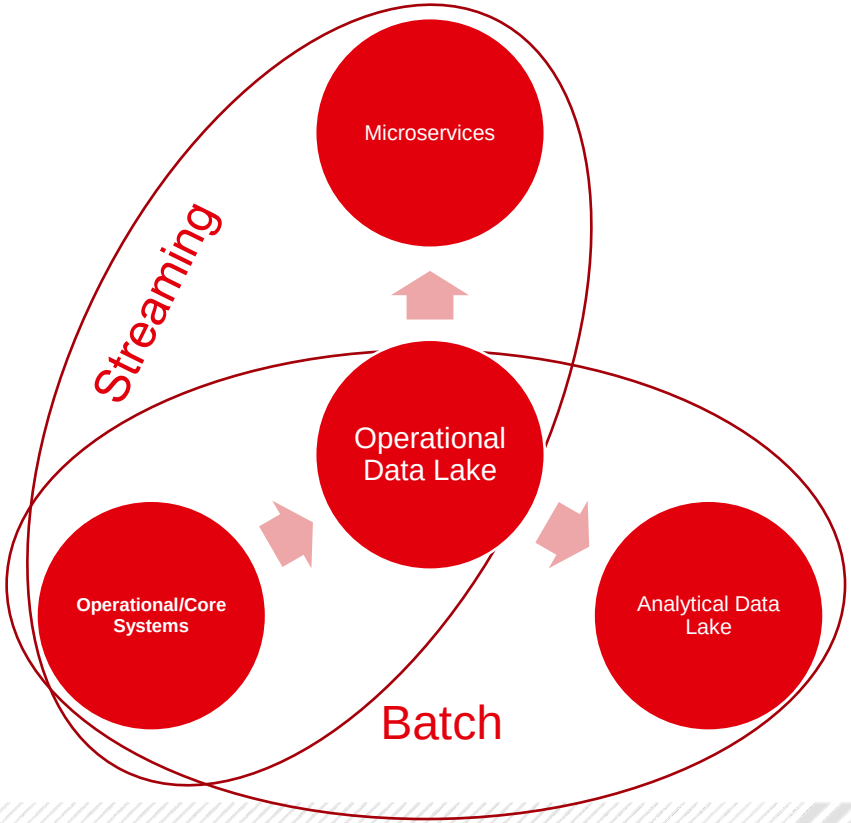


Design Principle:
Operational effort should remain low!

- Cluster capability
- Dynamic resource allocation
- Capable of near realtime
- Autonomous Database
- Connection to Hadoop
- Commodity hardware

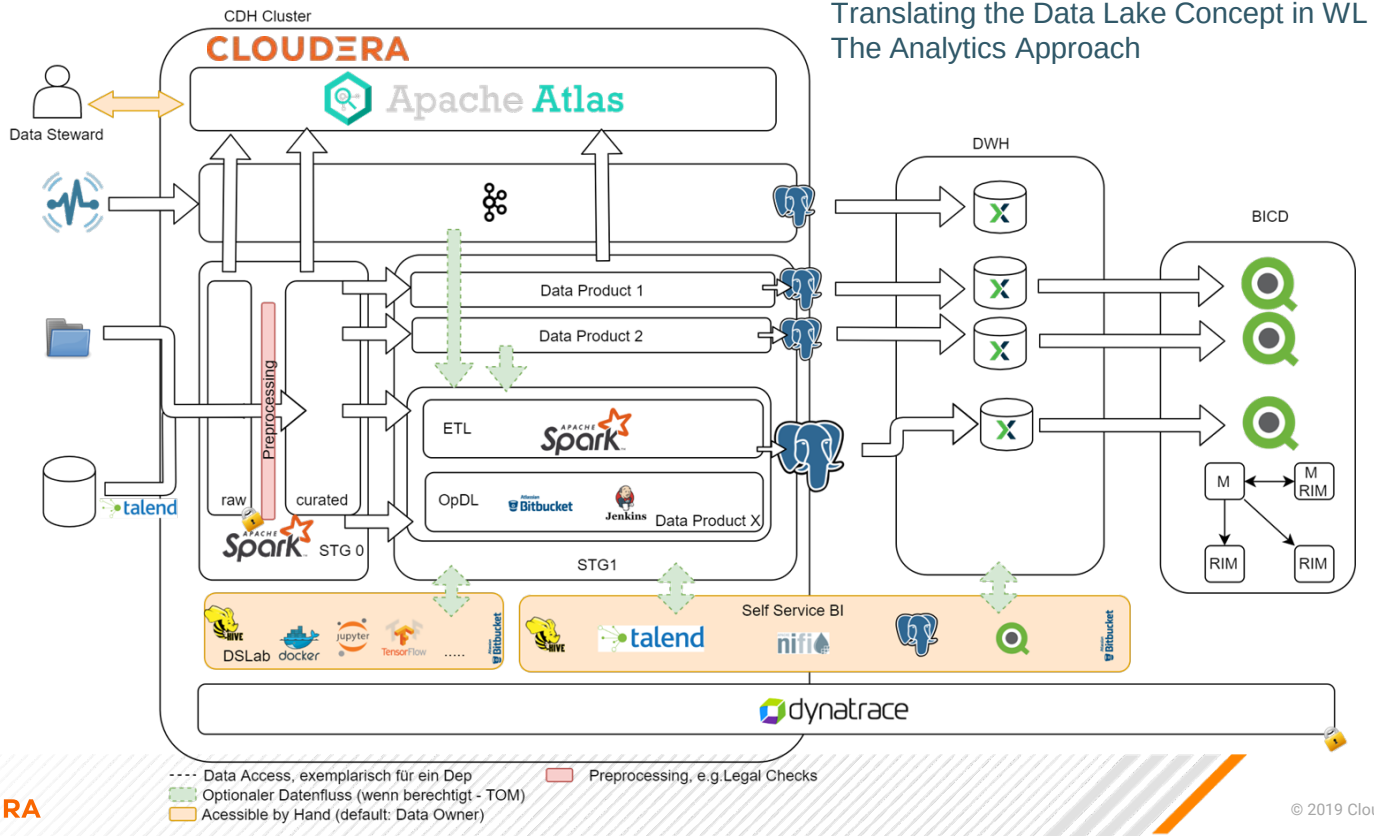
- Open development interfaces like R, Phyton, ..
- High Performance Exploratory Zone
- Connection to the Active Directory
- Implementation should largely be done in ANSI/ISO SQL (view logic and the like)

Overview of the Data Architecture @ Wiener Linien

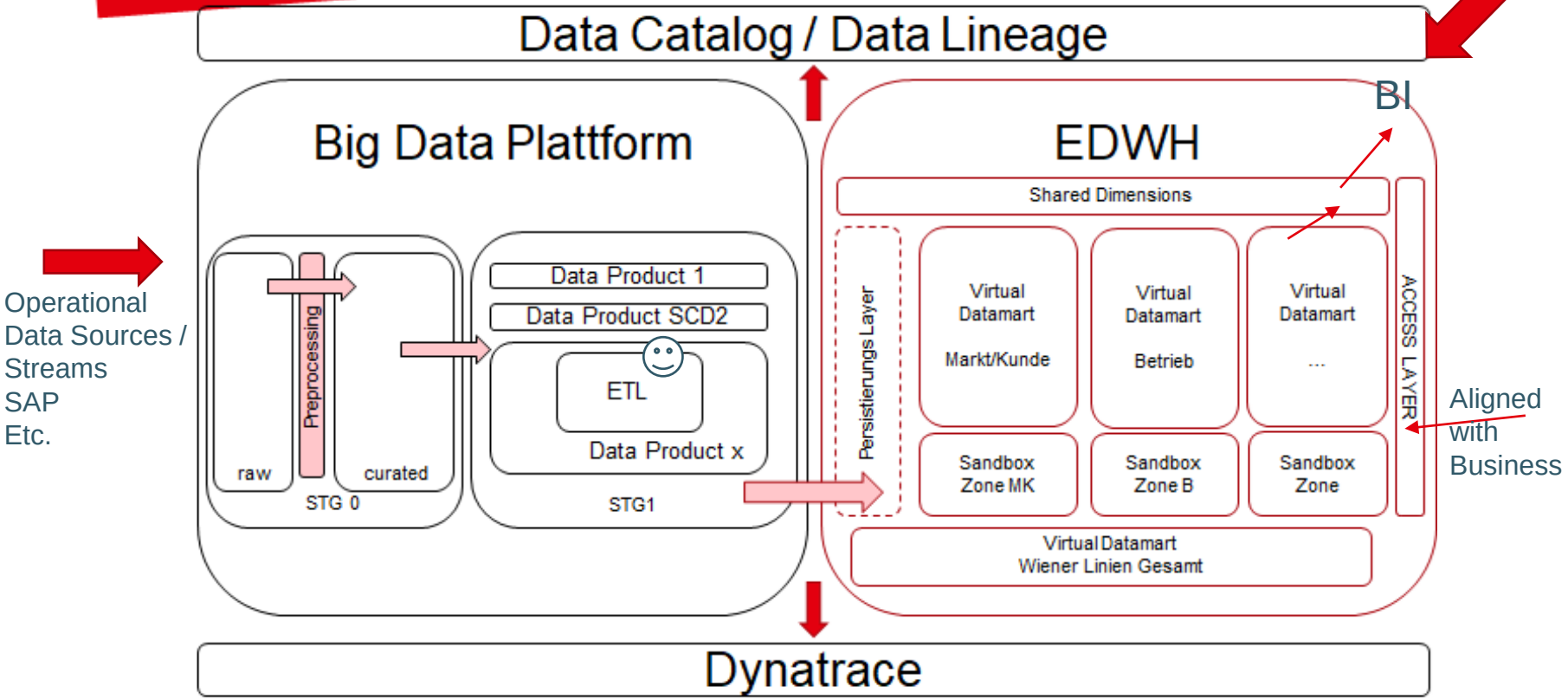


Architecture - Batch

Translating the Data Lake Concept in WL Architecture
The Analytics Approach



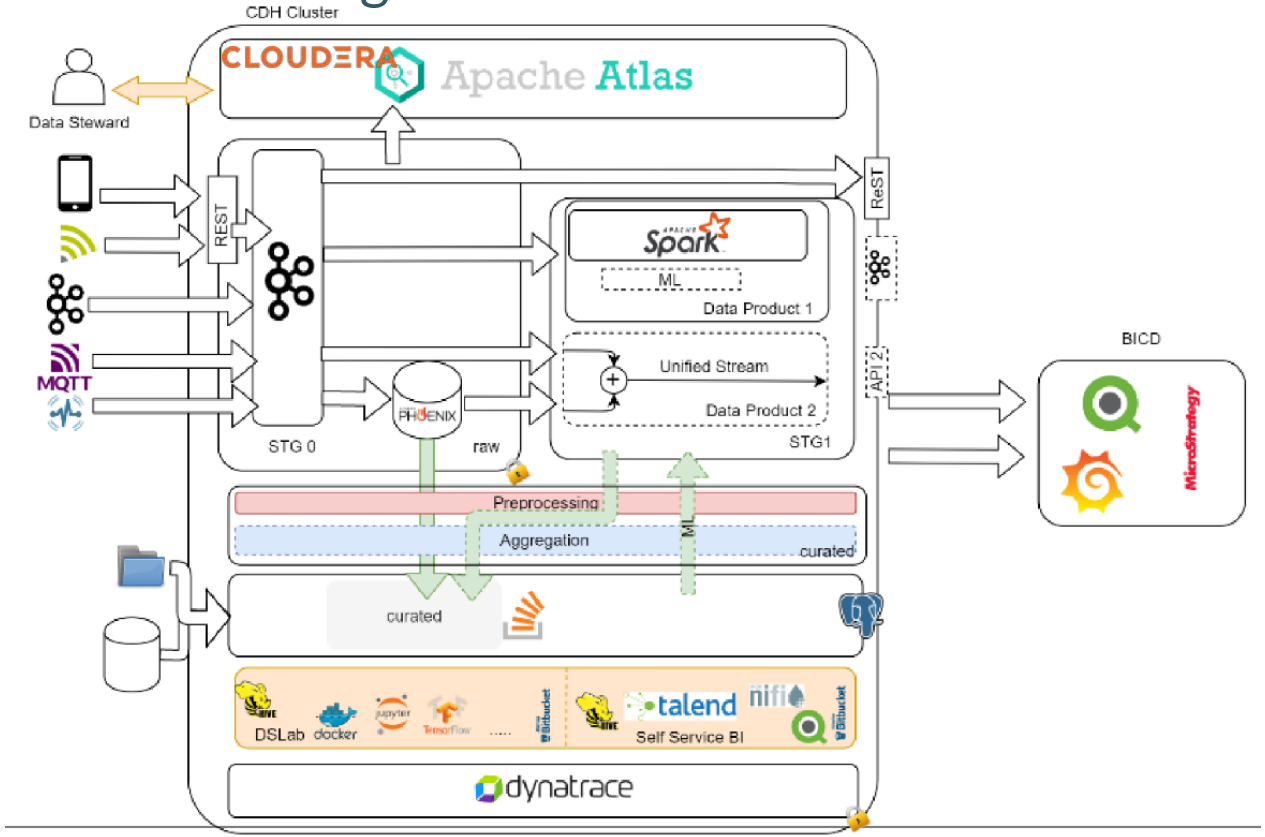
Architecture



Operational
Data Sources /
Streams
SAP
Etc.

Aligned
with
Business

Architecture - Streaming



Status Cloud @ Wiener Linien

Wir sind schon Multi-Cloud...

Azure
M365,
POCs neue
Betriebsapps

Google
Kunden-Apps &
Location

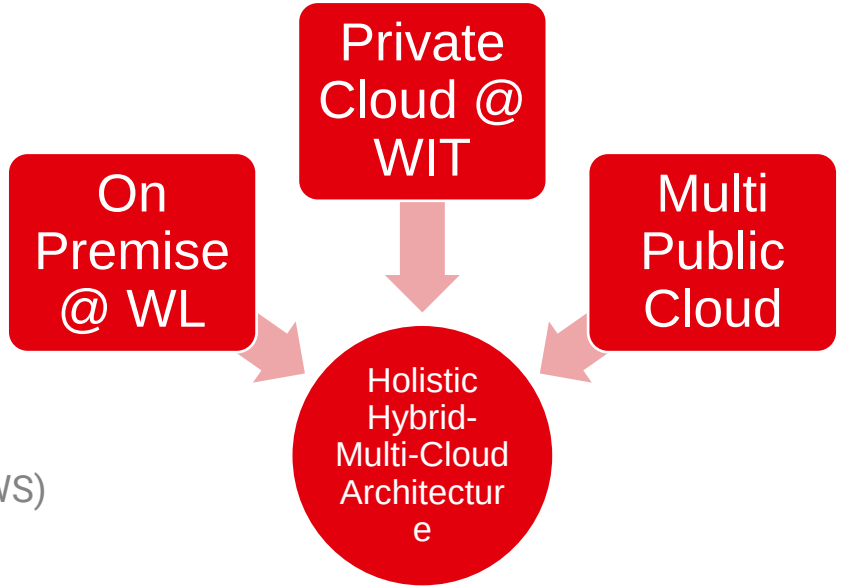
AWS
Siemens
Mindsphere
technische Daten

CLOUDERA

Holistic Hybrid-Multi-Cloud Architecture

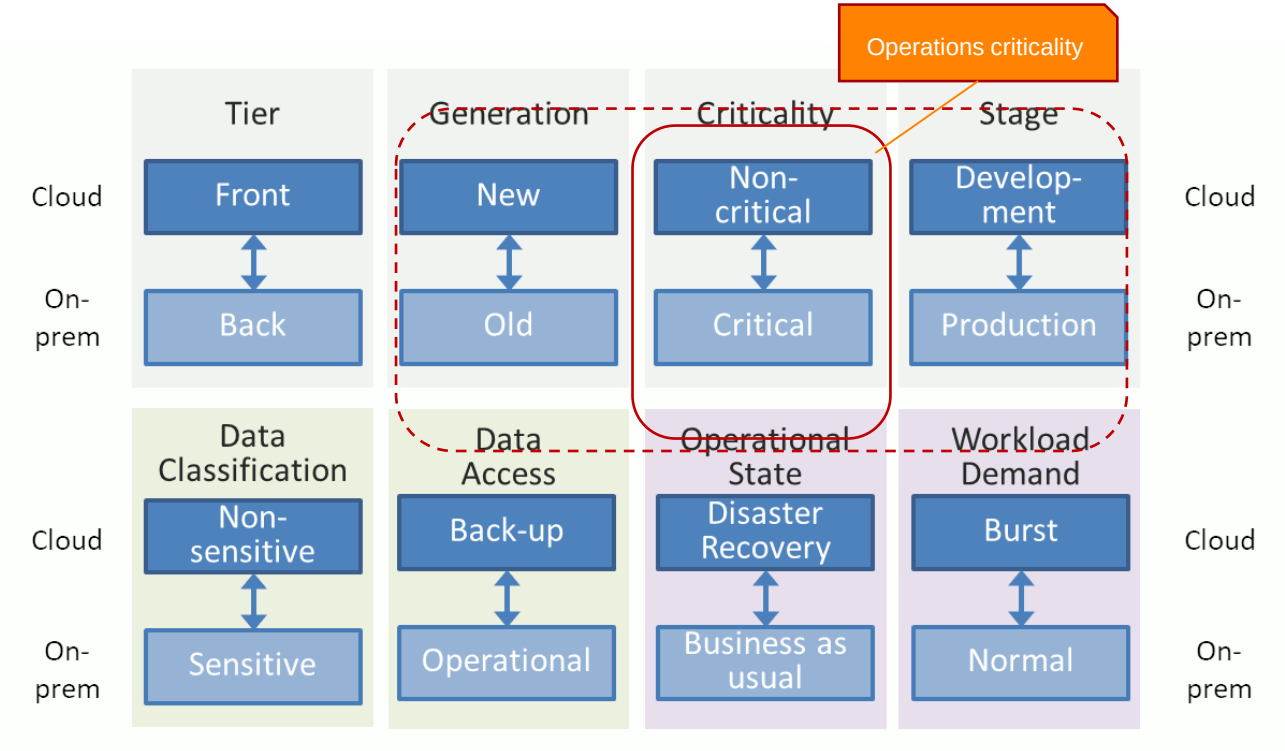
Integration der Systeme on premise & in clouds

- Integration Hybrid Cloud
  - On Premise @ Wiener Linien (WL)
  - Private Cloud @ Wien IT(WIT)
  - Multi Public Cloud
- Zu integrierende IT Systeme
  - SAP (WL, WIT, Azure)
  - Data Lakehouse (WL, WIT, Azure)
  - Operations & Leitsysteme (WL, WIT, Azure, AWS)



Hybrid Cloud: Ways to Slice the Elephant*

considerations



*Source: https://architectelevator.com/cloud/hybrid-cloud/

Data Science & KI Initiatives

Data Excellence / Data Sharing

Langsamfahrstellen

Interdisziplinär Prozesse optimieren!

RBL (Betrieb/ Fahrzeugdaten)

- An welchen Stationen sind wie viele Fahrzeuge pünktlich / verfrüht oder verspätet?

ISDB (Infrastruktur)

- Wo befinden sich dir LFS?
- Wie heißen die Stationen und Linien?
- Wie hängt er Netzplan zusammen?

HASTUS (Fahrpläne)

- Welche Pläne sind wann und auf welchen Linien gültig?

Mehrwert:

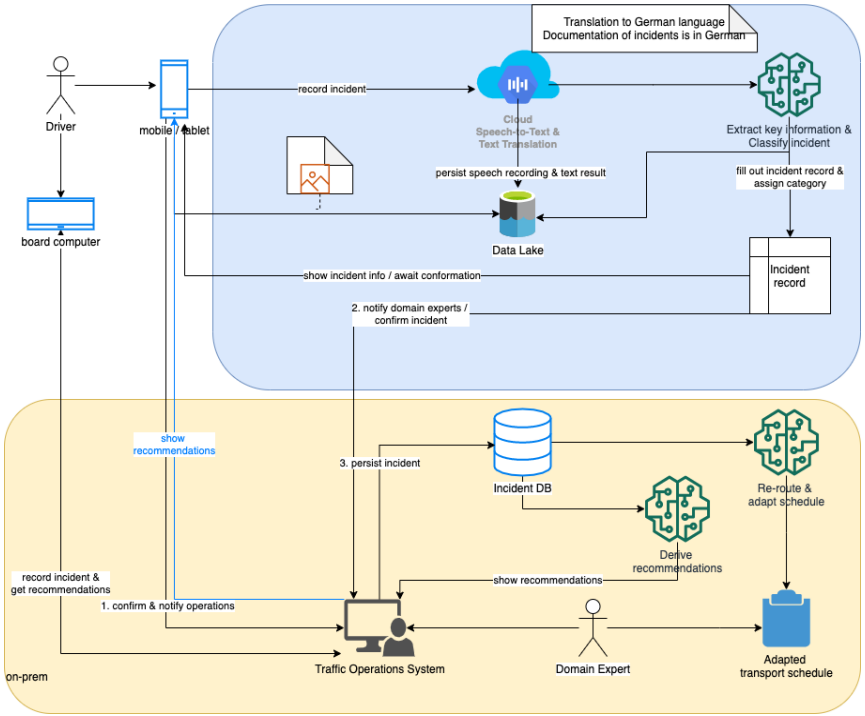
- ✓ **Schweregrad von LFS analysieren**
- ✓ **Bauvorhaben & Wartung der Gleise priorisieren > Effizienzsteigerung im Betrieb, Investitionssicherheit**

Disziplinen:

- ✓ **Data Science**
- ✓ **Betrieb - Straßenbahn**
- ✓ **Infrastruktur – Schiene**
- ✓ **Infrastruktur – Datenbank**
- ✓ **Business Intelligence**

Vorfallsmanagement & Passagierfluss-Lenkung bei Vorfall

Optimierung & Änderung von betrieblichen Geschäftsmodellen

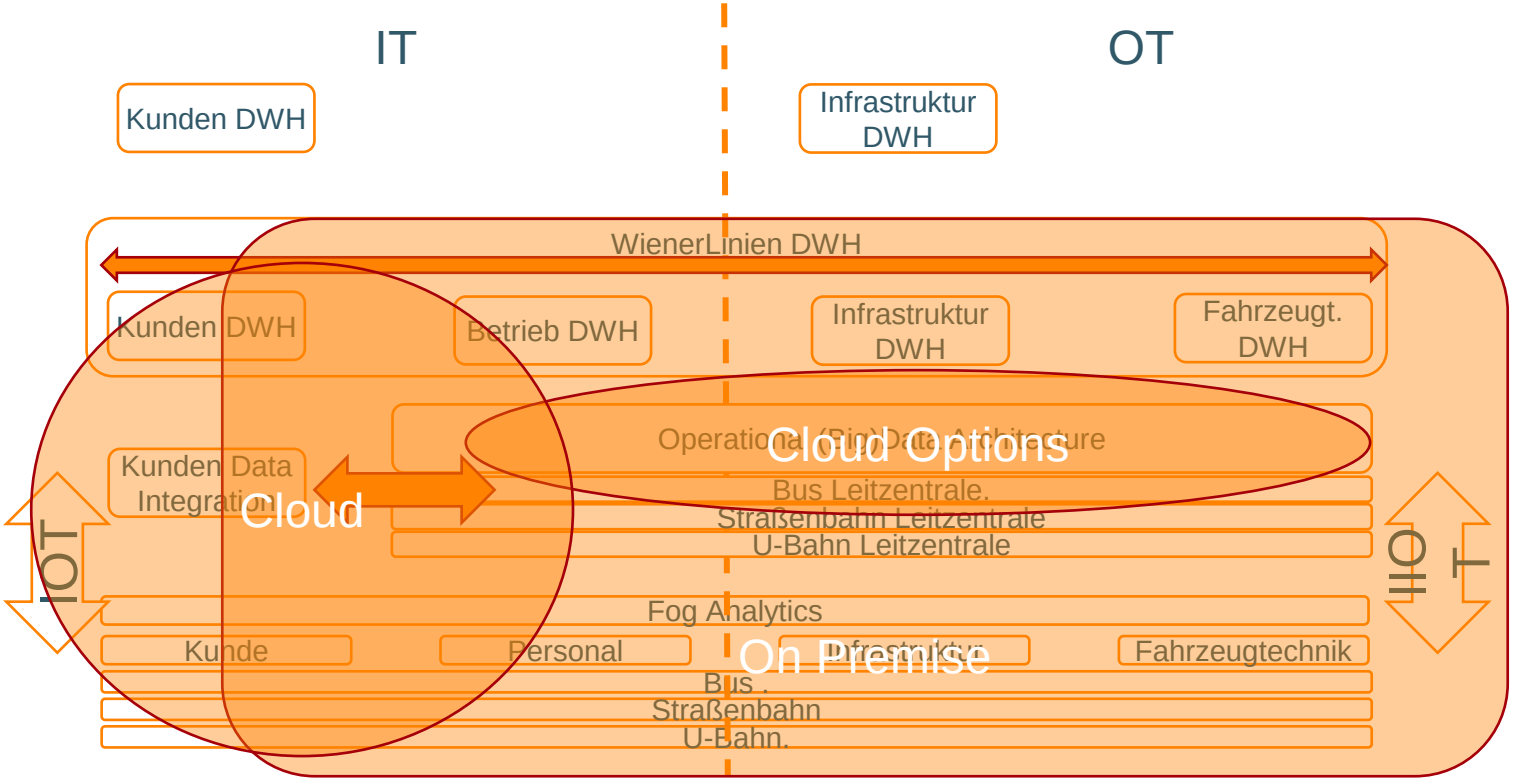


Effiziente Registrierung von Vorfällen mithilfe mobiler Lösungen für Sprachausgabe und NLP. Zentrale Benachrichtigung und Verwaltung von Vorfällen, Unterstützung der Verkehrsumleitung und Zeitplananpassung.

Disziplinen:

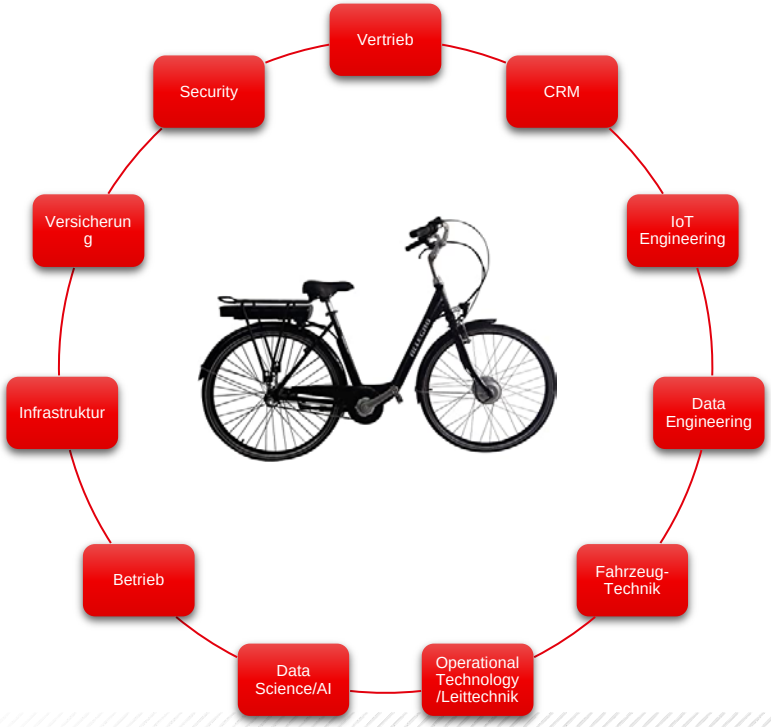
- Betrieb
- IT (Applikationen, Mobile Devices, Hybrid Cloud)
- Data Science KI / Data Engineering
- Operational Technology / Leitsysteme
- IoT

High Level Data Architecture based on Functions



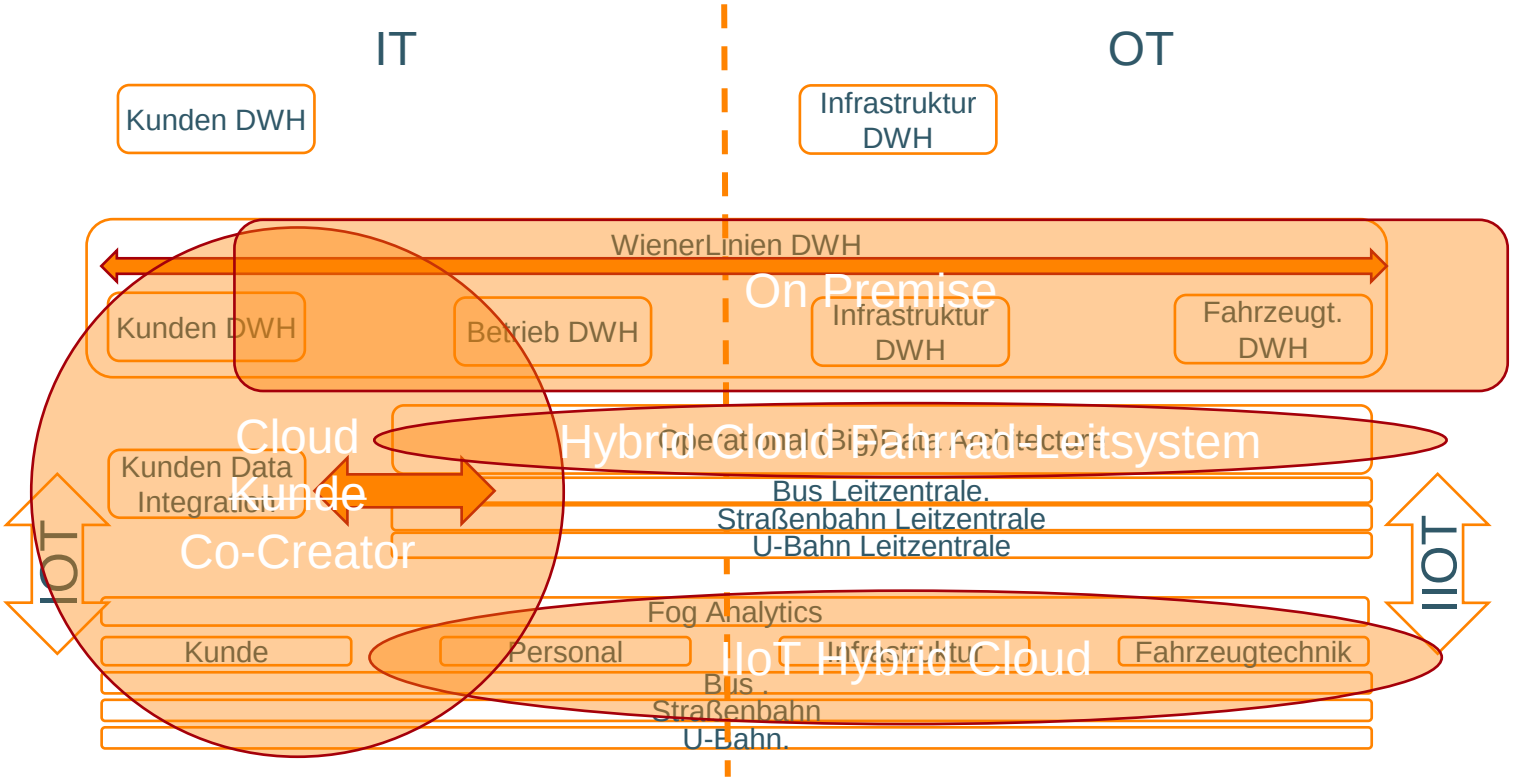
Future City E-Bike

Neue Business Modelle bis Differenzierung



- „Continuous Improvement Process“ → Think big, Start small
- Neue Business Modell-Entwicklung mit Business Model Canvas
- In der Zukunft → Echtzeit- Fahrzeug-Steuerung durch IoT und Dataplattformen
- Über 10 Disziplinen übergreifend über die Wiener Linien
- Daten & Information von allen WL Hauptabteilungen in der Zukunft notwendig

High Level Data Architecture für Wien Mobil Bike



CLOUDERA

THE ENTERPRISE DATA CLOUD COMPANY

We believe that data can make what is impossible today, possible tomorrow

We empower people to transform complex data into clear and actionable insights

We deliver cloud-native services to manage and secure the data lifecycle in any cloud or data center

TODAY'S ENTERPRISE NEEDS A MODERN DATA ARCHITECTURE

Cloud Experience

- Easy to use with self-serve capabilities
- Elasticity and agility to meet changing demands of workloads and company
- Simple to manage and maintain environments and applications

Hybrid, Multi-Cloud

- Move data and applications without rewriting and retraining
- Separate data management strategy from infrastructure strategy
- Manage all environments from a single pane of glass

EDGE 2AI

Multi-Function & Open

- Deploy one platform to address current and future workload needs
- Connect disparate workload types to develop applications on one platform
- Open source, open APIs

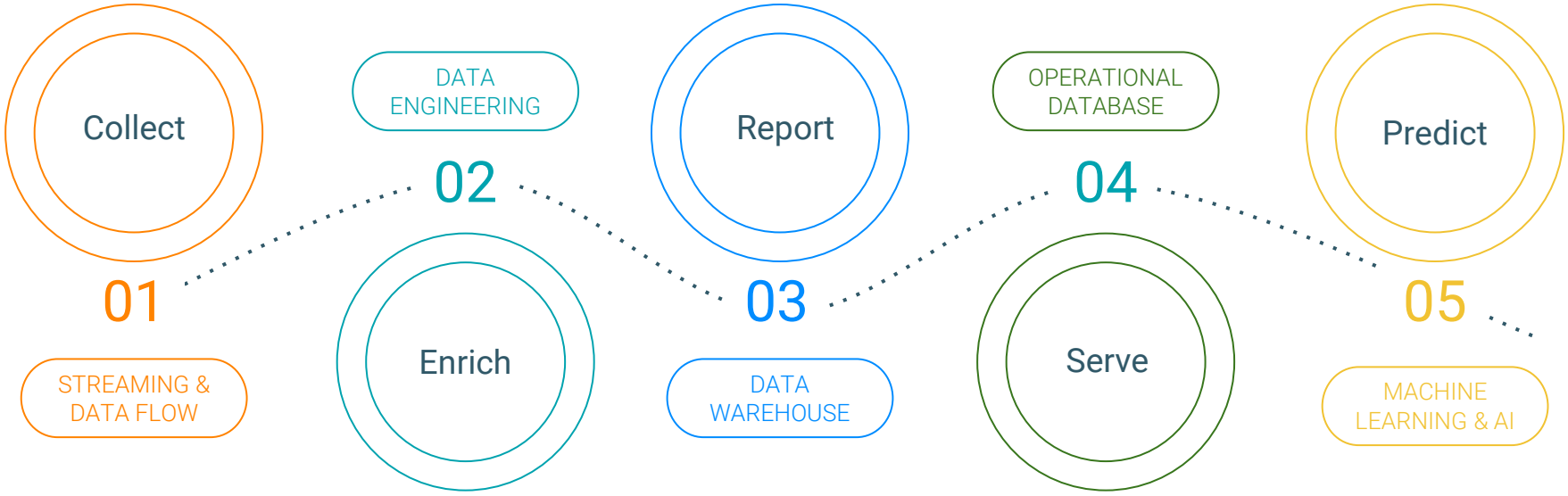
CLOUDERA SDX

Secure & Governed

- Manage data security and governance centrally
- Automate application security at all layers
- Reduce time to value with enterprise-grade productivity tools

CLOUDERA - THE ENTERPRISE DATA CLOUD COMPANY

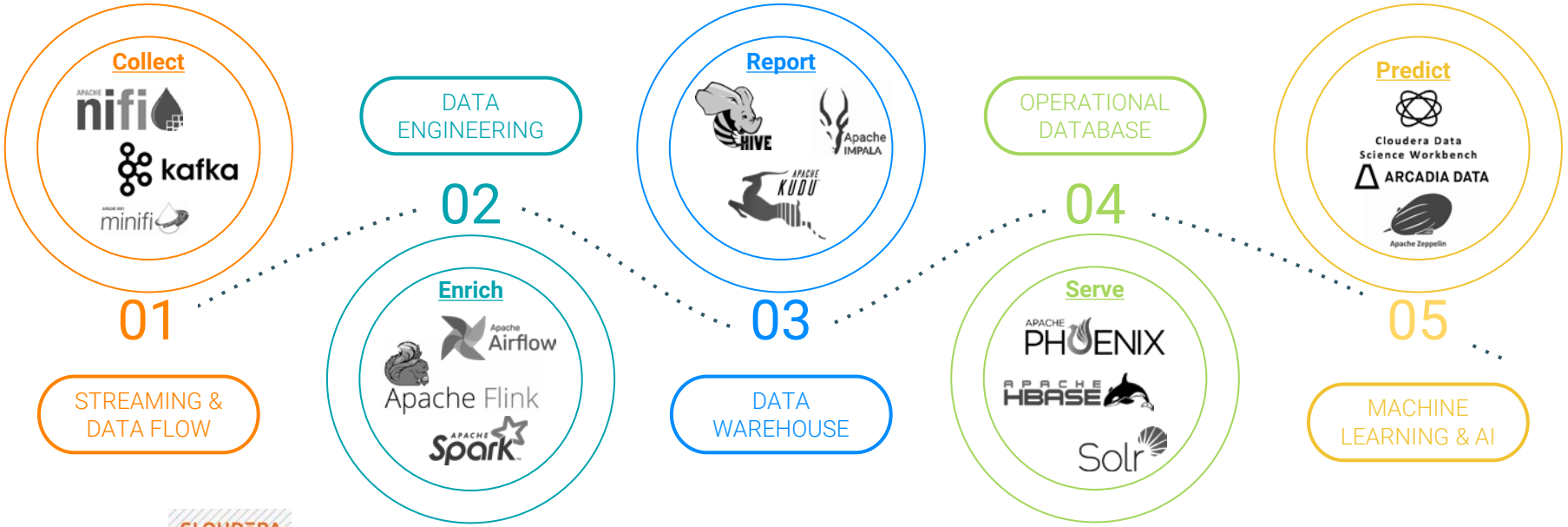
Manage and secure the data lifecycle in any cloud or datacenter



CLOUDERA SDX SECURITY | GOVERNANCE | LINEAGE | MANAGEMENT | AUTOMATION

DELIVERING THE BEST IN OPEN SOURCE FOR DATA AND ANALYTICS

Best-fit technology integrated into one platform for the entire business



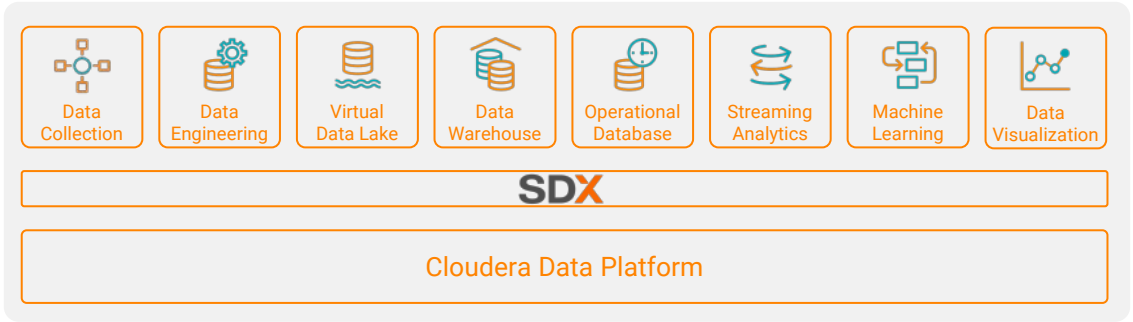
CLOUDERA SDX SECURITY | GOVERNANCE | LINEAGE | MANAGEMENT | AUTOMATION

Apache Ranger Apache Atlas

A HYBRID / MULTI-CLOUD DATA PLATFORM *AND* AN INTEGRATED SUITE OF SECURE ANALYTIC APPS

Real-time
Batch
Structured
Unstructured

Data Sources



Analysts
Engineers
Scientists
Developers

Data Users

Data Lifecycle
integration for better user productivity and faster time to value

Hybrid & Multi-Cloud
to leverage existing investments and reduce risk

Secure & Governed
to simplify data protection, sharing and compliance

Open & Extensible
to support more use cases faster and at lower cost

TH  N  Y  U