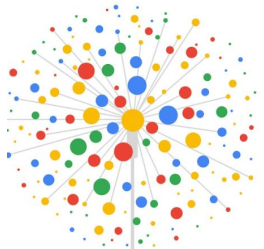
Thyssenkrupp Steel
Data Analytics Workshop

WELCOME!

Google Cloud

Agenda 1.7.

- 13:00 - 14:00 Kennenlernen & Data Platform Überblick
- 14:00 - 14:30 Set up, Projekt erstellen, Daten hochladen
- 14:30 - 15:15 ETL Prozess (Dataprep, Dataflow, Datafusion)
- 15:15 - 16:00 Big Query (Data Import, Federated and Saved Queries)
- 16:00 - 17:00 AI/ML (Big Query ML, Sheet Visualization, Vertex AI Notebooks)



Google Cloud

The *Art* of Building a Cool Data Platform on Google Cloud

Stiv Sterjo, Smart Analytics & AI Specialist

Google Cloud

Organizations are hindered by traditional data warehouses limitations

1

Systems can't **scale** fast enough to keep up with **data growth**

telegraphmediagroup

2

Data is not **fresh** or current enough

Blue Apron

3

Access restrictions on data prevent teams from working effectively

4

Systems don't support **machine learning** and **AI initiatives**

5

Legacy systems bust customer **budgets** and renewing **licenses** is a hassle

HSBC

Google Cloud

DATA
CONSUMERS

Google Cloud

DATA
SHOP

- Instant Archive Retrieval
- Lifecycle Rules
- Signed-URLs
- Geo-Redundancy
- Private Google Network

Cloud Storage

Custom data experiences with Looker

Modern BI & Analytics	Integrated Insights	Data-driven Workflows	Custom Applications
Interactive dashboards	Reports in your inbox or Slack	Feed ML/AI workflows	Advanced developer tools
Self-serve data exploration	Data in the tools people already work in	Dynamic real-time alerting	Productize analytics

Looker

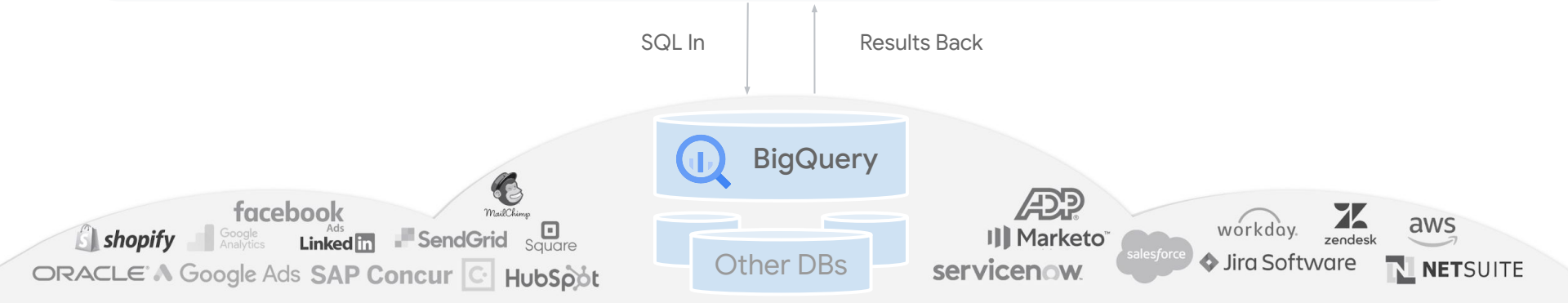
API-first & cloud native for integrating into existing workflows

Semantic modeling layer for enterprise-wide governance

In-database architecture for access to real-time data

No performance bottlenecks by pushing the processing to the database layer

perfect when combined with BigQuery



DATA
CONSUMERS

Google Cloud

DATA
SHOP

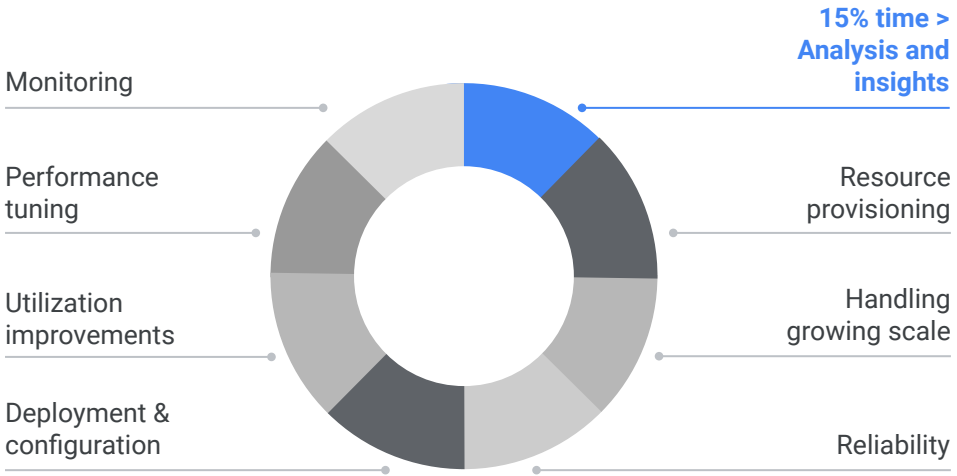
- Instant Archive Retrieval
- Lifecycle Rules
- Signed-URLs
- Geo-Redundancy
- Private Google Network

Cloud Storage

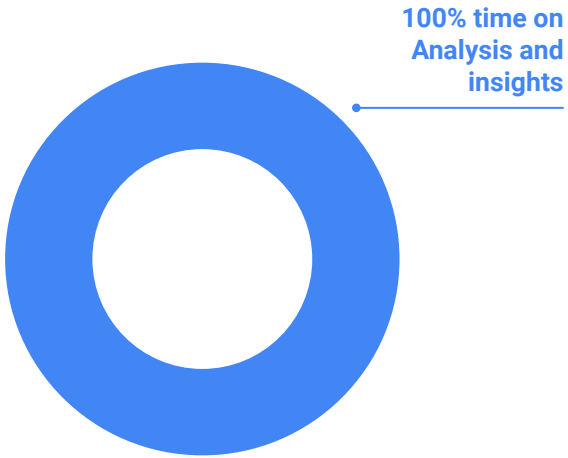
BigQuery

BigQuery | Serverless data warehouse

Traditional data warehouses



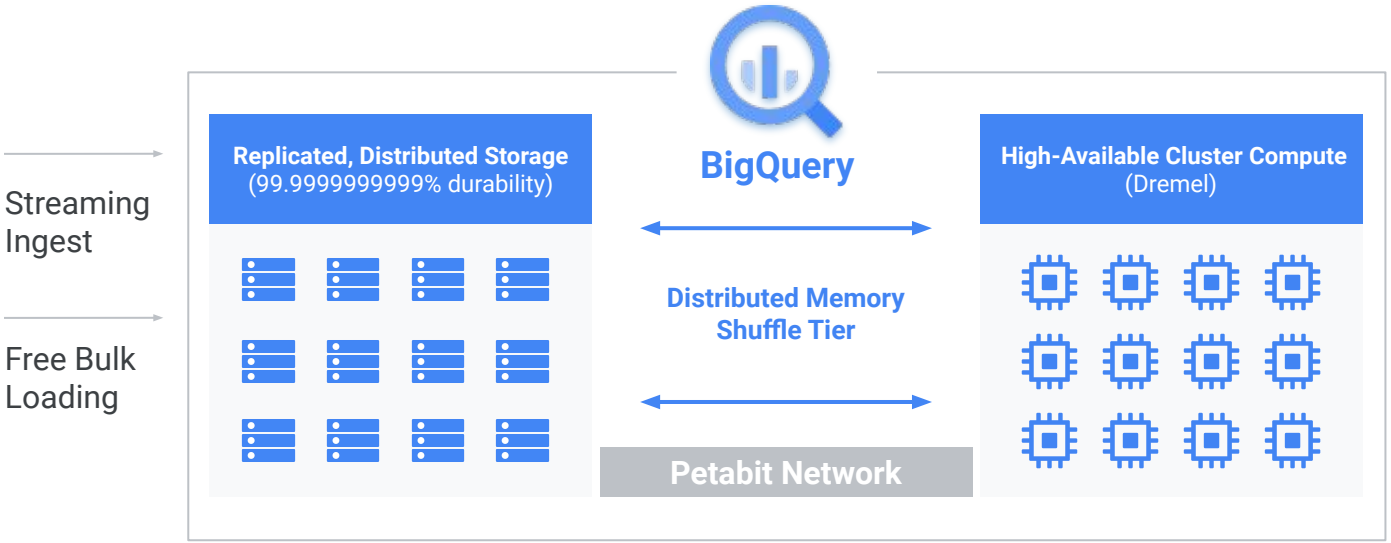
BigQuery's serverless analytics



Google Cloud

BigQuery | Architecture

Decoupled storage and compute for maximum flexibility



Google Cloud

BigQuery

Data warehouse with customers ranging from TB to 100+ PB

- Cloud-scale enterprise data warehouse
- Standard SQL(ANSI 2011) with DML Support
- Encrypted, durable, highly available

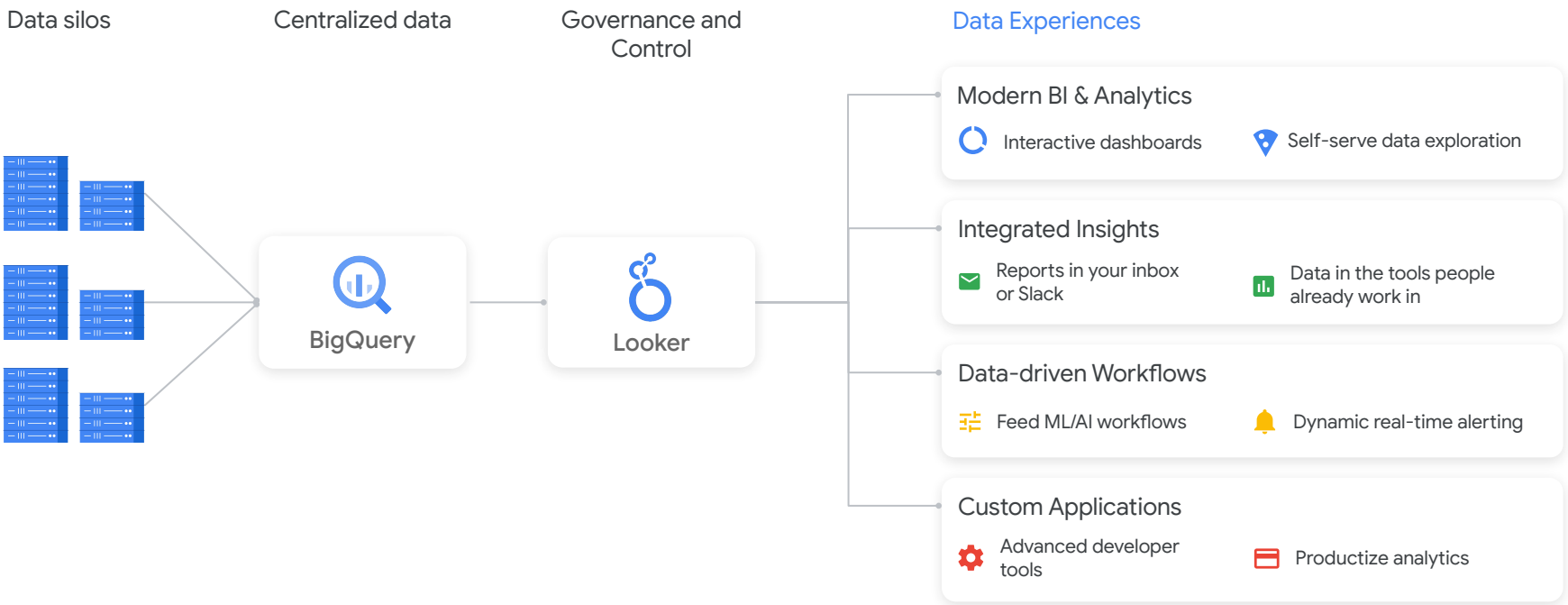
Spotify	**Serverless** platform	Unique
King	**Real-time** insights	Unique
20th Century Fox	Built-in **ML**	Unique
AirAsia	**Insights** for everyone	Unique

Google Cloud

BigQuery is named a Leader in The Forrester Wave™: **Cloud Data Warehouse, Q1 2021**

Google Cloud

Better together: BigQuery + Looker powering data experiences



Google Cloud

DATA
CONSUMERS

Google Cloud

DATA
SHOP

Cloud Storage

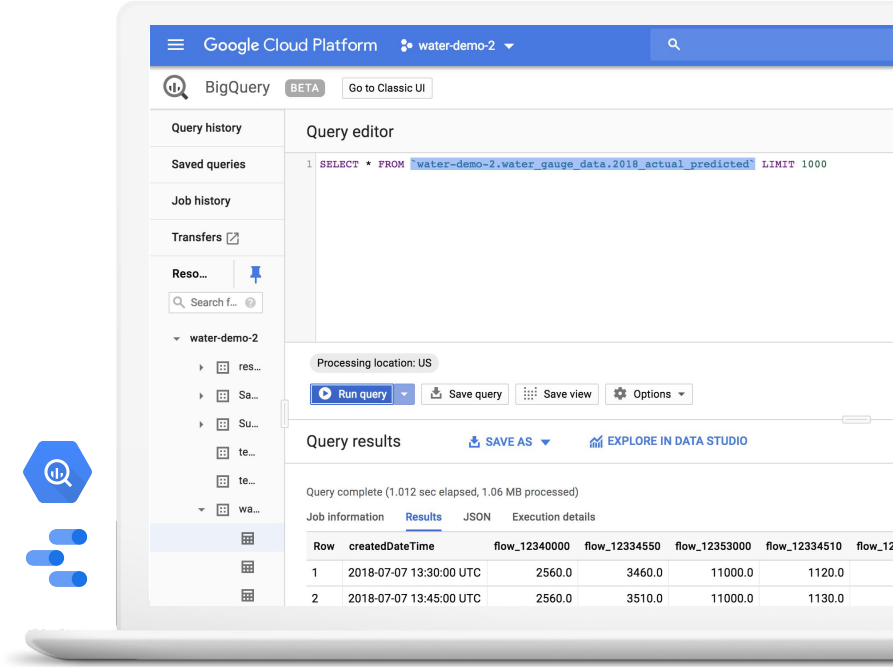
- Instant Archive Retrieval
- Lifecycle Rules
- Signed-URLs
- Geo-Redundancy
- Private Google Network

BigQuery

- Expiration Rules
- Partitioning/Clustering
- Serverless/Autoscaling
- No Tuning/No Ops
- Ad-Hoc Materialization
- Compressed Storage
- Cost Savings

See your BigQuery data in one click with Data Studio Explorer

Tight integration in the BigQuery UI brings visual exploration of your query results in one simple click.



Google Cloud

Accelerate data exploration and unlock insights
BigQuery BI Engine

Super fast reporting

Enable real-time dashboarding over data inside BigQuery with sub-second query response and with high concurrency.

Simplified architecture

Simplify BI architecture and accelerate time to insight by reducing additional ETL, extracts or BI server management.

Smart tuning

Optimize visual analytics over big data sets with automatic performance tuning and capacity flexibility.

Google Cloud

AppSheet enables anyone to create apps fast with no-code

1. Connect your data

2. Customize your app

3. Deploy your app

Google Cloud

DATA
CONSUMERS

Google Data Studio

AppSheet

Looker

Google Cloud

DATA
SHOP

Cloud Storage

- Instant Archive Retrieval
- Lifecycle Rules
- Signed-URLs
- Geo-Redundancy
- Private Google Network

BigQuery

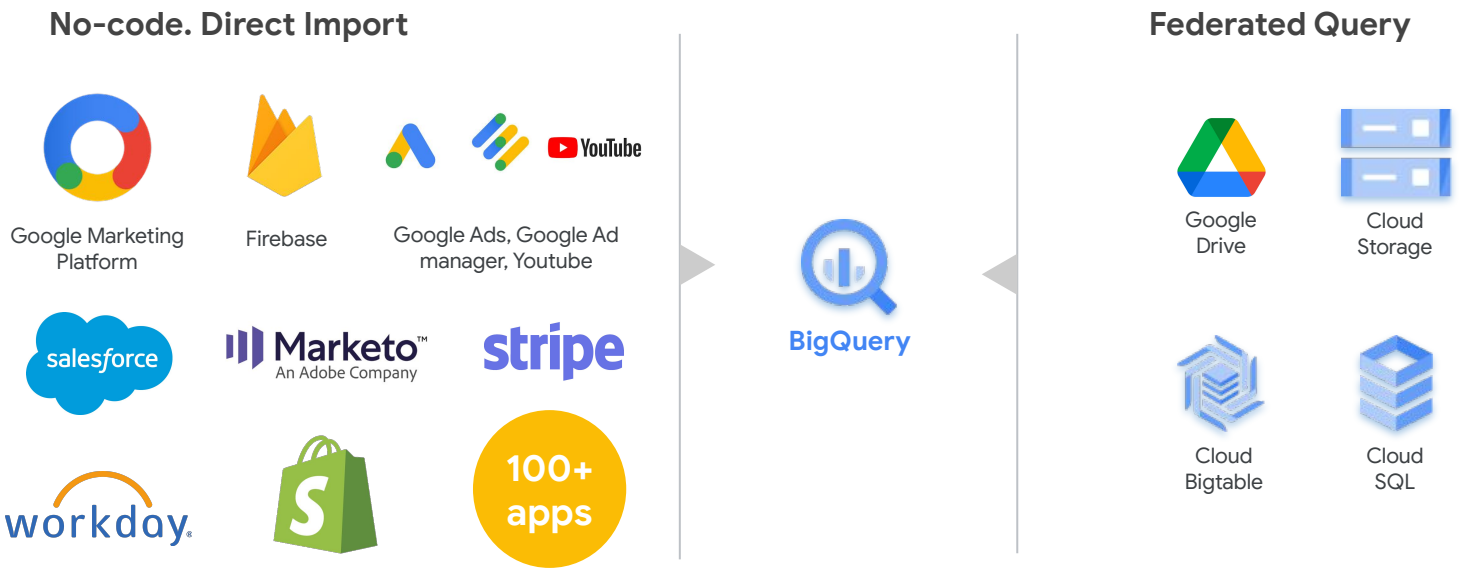
- Expiration Rules
- Partitioning/Clustering
- Serverless/Autoscaling
- No Tuning/No Ops
- Ad-Hoc Materialization
- Compressed Storage
- Cost Savings

INGEST &
PROCESS

DATA
SOURCES

Automated data transfer and federation

Use BigQuery to analyze all data important to an organization. Users should be free to make optimal decisions for storage and compute.



Google Cloud

Cloud IoT Core

Fully managed service to easily and securely connect, manage, and ingest data from globally dispersed devices

Protocol Bridge

- **Bi-directional messaging**
- **MQTT and HTTP protocol** endpoints
- Automatic load balancing
- **Global data access** with Pub/Sub

Cloud IoT Core

Device Manager

- **Device identity management**
- Configure, update and **control** individual devices
- Role level access control
- Console and APIs for device deployment and monitoring

Google Cloud

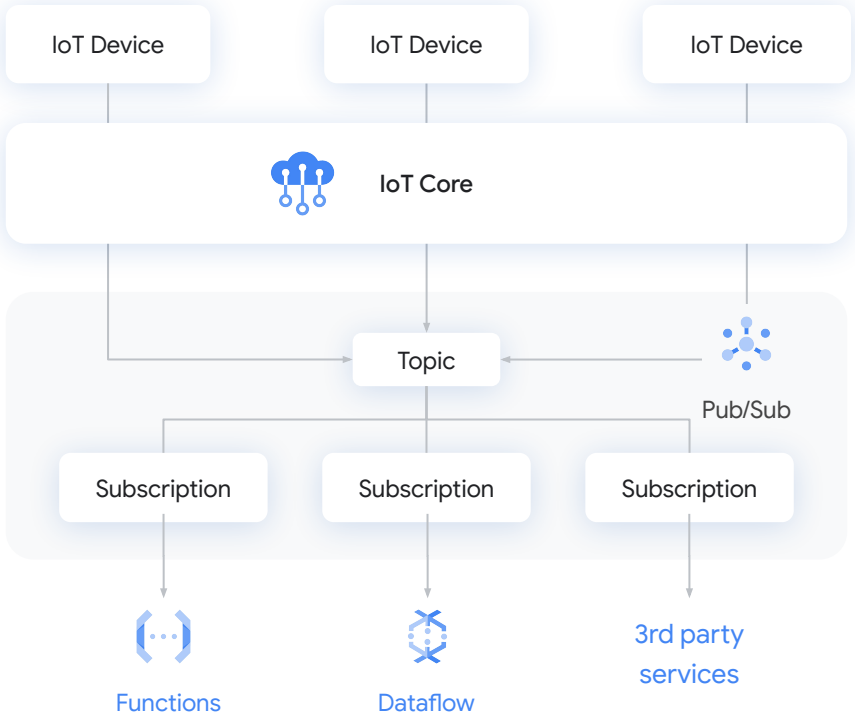
Ingestion at scale with Cloud Pub/Sub

Durable message persistence

Global service

Powerful real-time features with Dataflow & Cloud Functions

Separation of application and devices



Google Cloud

BigQuery for real-time processing

Cloud Pub/Sub

Serverless, reliable, scalable event delivery system

Zero management: just create a topic and set ACL

Cloud Dataflow

Serverless data processing pipelines based on Apache Beam

Unified batch+stream

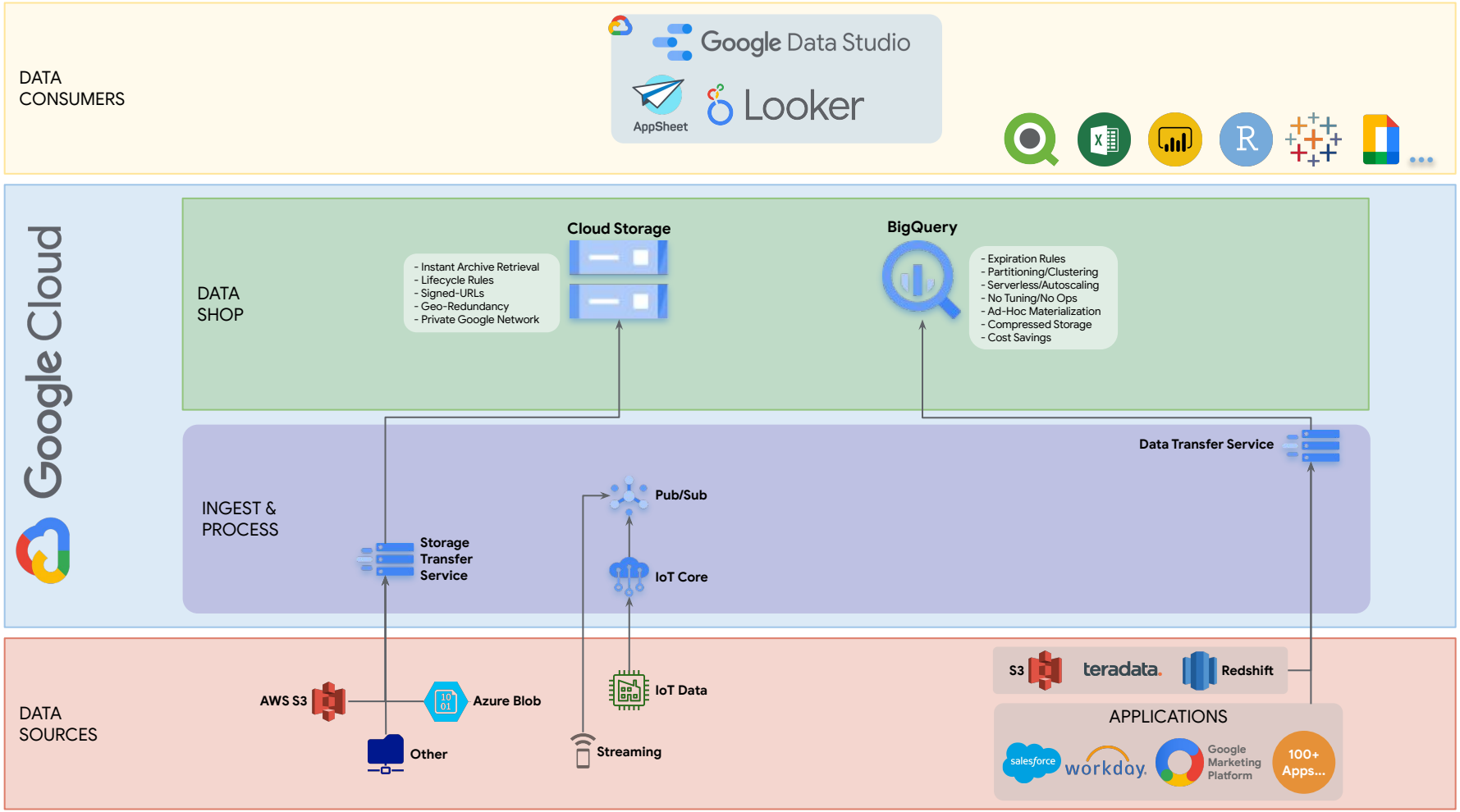
Java + Python

BigQuery

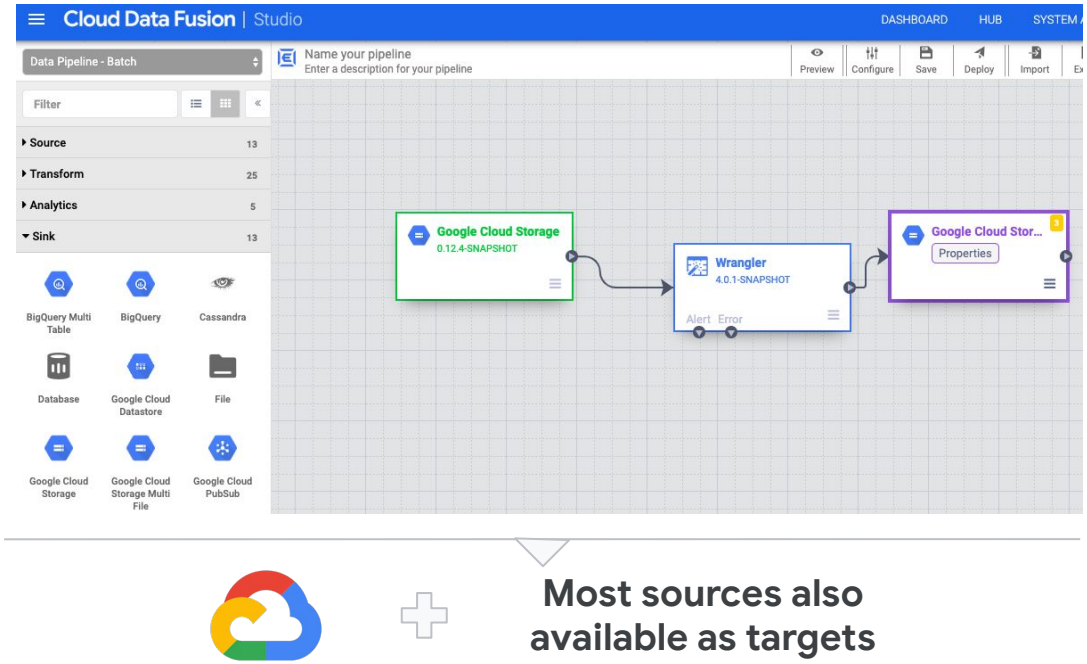
Stream API for immediate ingestion of events.

No additional buffer and file coalescence to manage.

Google Cloud


DATA CONSUMERS
Google Data Studio
AppSheet
Looker
Google Cloud
DATA SHOP
Cloud Storage
- Instant Archive Retrieval
- Lifecycle Rules
- Signed-URLs
- Geo-Redundancy
- Private Google Network
BigQuery
- Expiration Rules
- Partitioning/Clustering
- Serverless/Autoscaling
- No Tuning/No Ops
- Ad-Hoc Materialization
- Compressed Storage
- Cost Savings
INGEST & PROCESS
Data Transfer Service
Pub/Sub
Storage Transfer Service
IoT Core
DATA SOURCES
AWS S3
Azure Blob
IoT Data
Other
Streaming
S3
teradata.
Redshift
APPLICATIONS
salesforce
workday.
Google Marketing Platform
100+ Apps...

Blend data from various sources and formats with Data Fusion



Most sources also available as targets

Google Cloud

Data Fusion

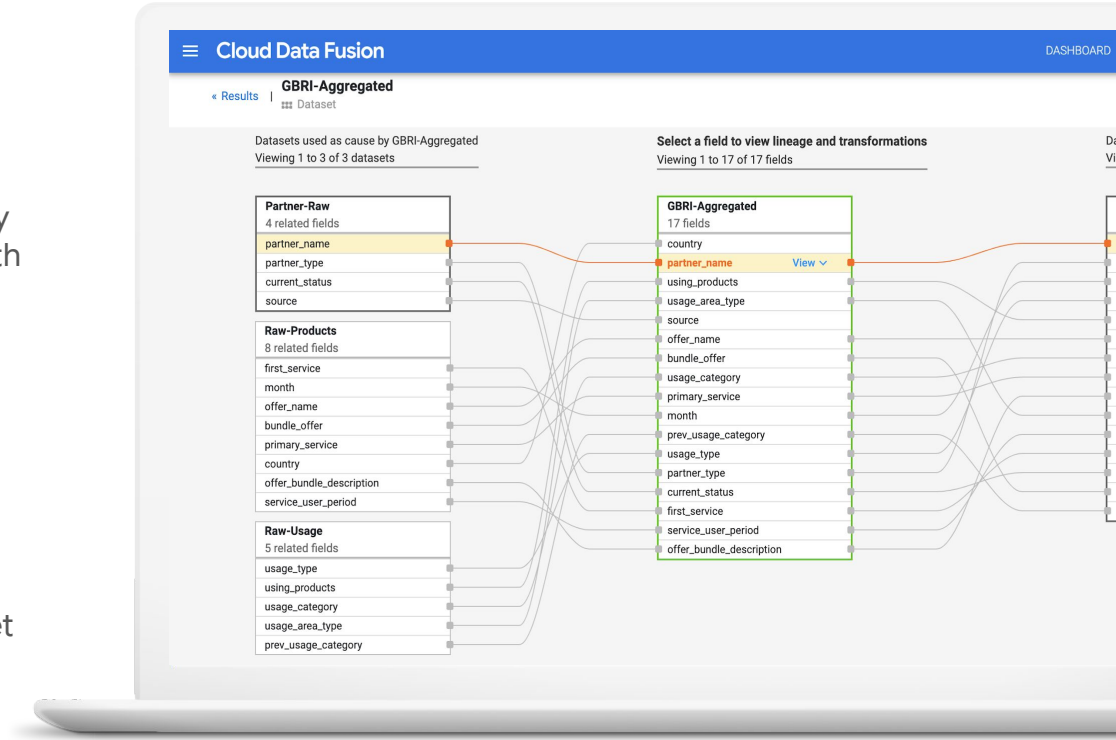
Data Lineage

Understand the flow of how your data is being prepared for business.

Granular visibility into flow of data:
Understand the impact of changes or identify root cause of discrepancy of data visually with field level lineage

Operational View:
Understand the processes and runs that transform the data

Openess or share lineage:
Use REST API to retrieve and integrate dataset level or field level lineage with other tools.



Google Cloud

Data Fusion

Open and community based

Cloud **Data Fusion** is Google's cloud native manage service for enterprise data integration and ingestion powered by CDAP.

CDAP

cdap.io

CDAP is an 100% open-source framework to build data analytics applications for on-premise and cloud.

API

Runtime

Pre-built Accelerators

Library

Apache 2.0 License

In production at scale for +5 years

Execution Options:

Cloud Dataproc

AWS EMR

Azure HDInsight

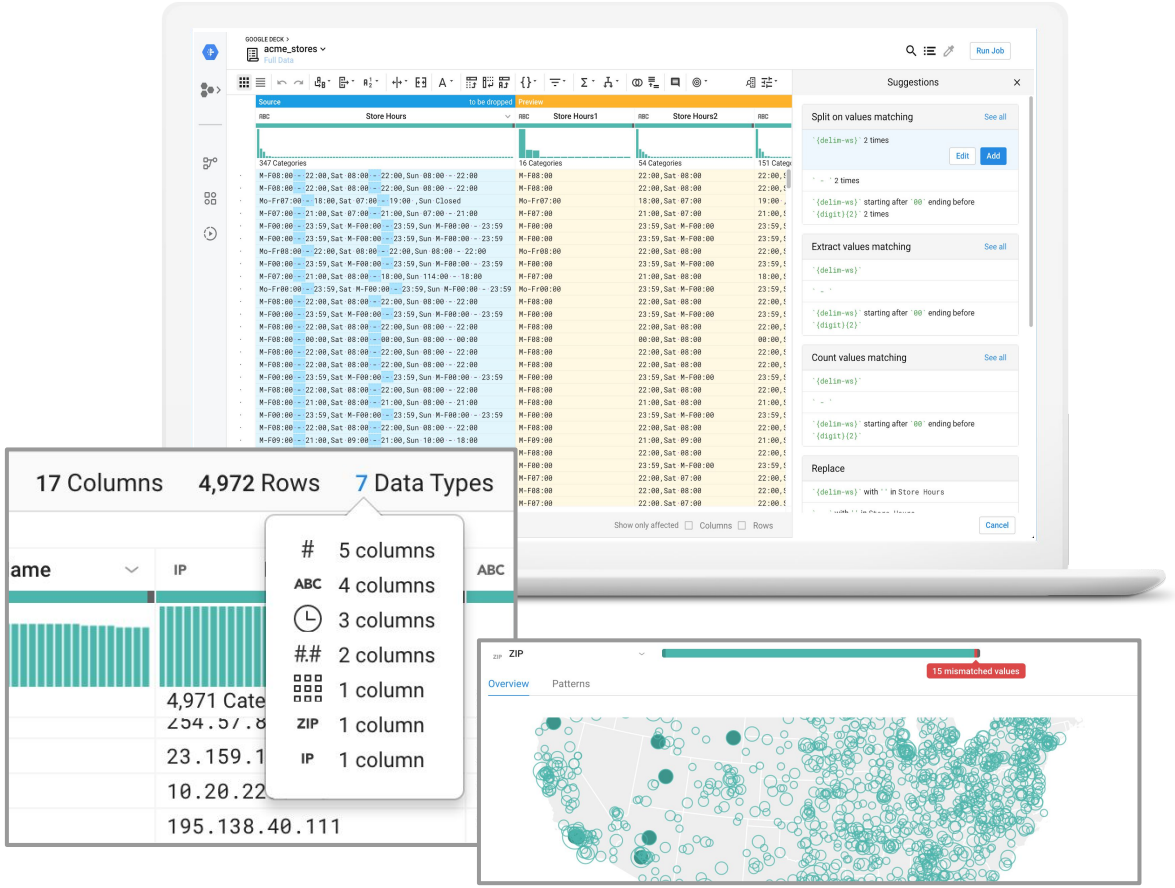
Cloudera Data Platform / HDP

Google Cloud

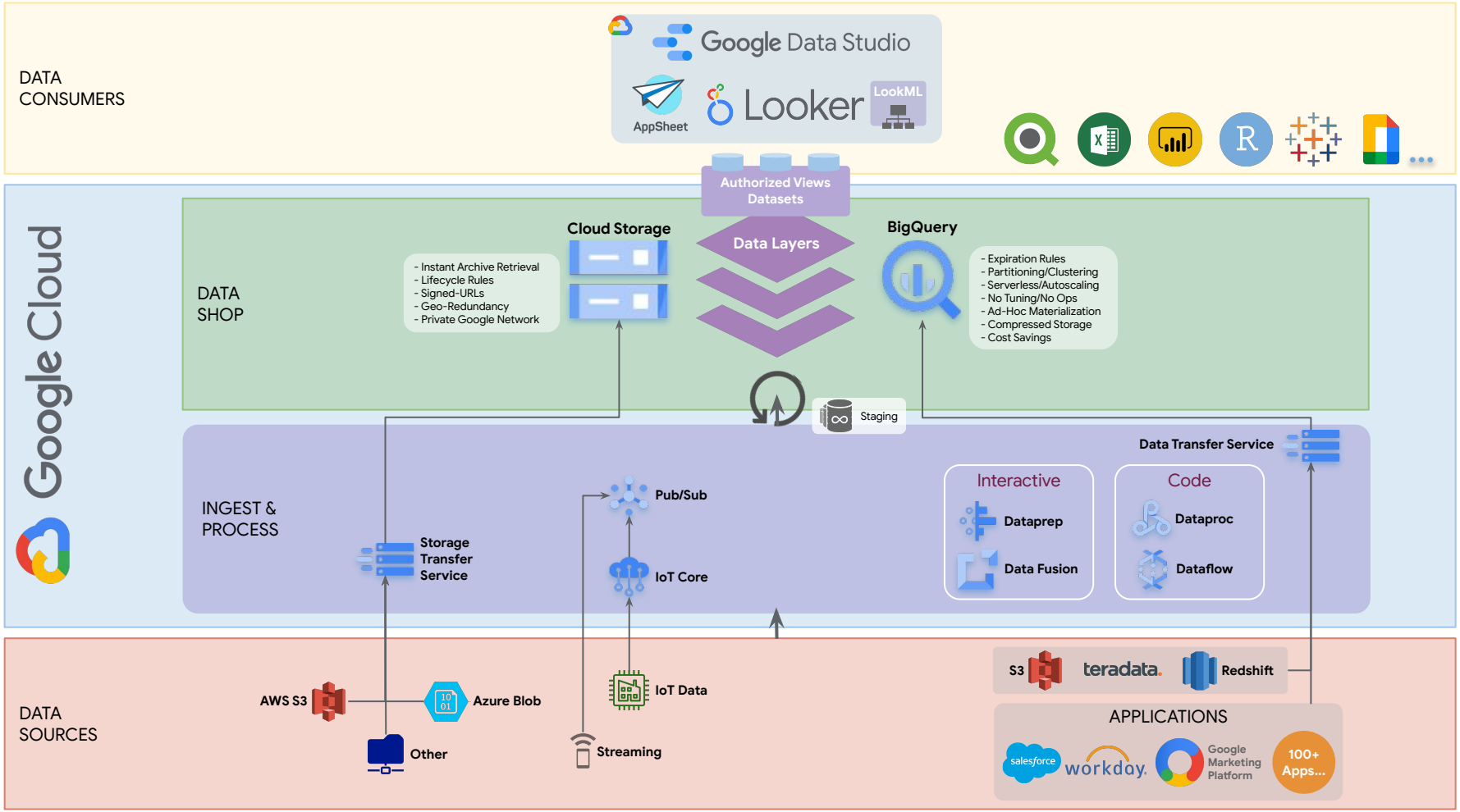
Dataprep

Built-in intelligence

- Get ML-guided transformation suggestions automatically
- Standardize, structure and join datasets easily with a guided approach
- One-click feature engineering and calculation
- Instantly understand data distribution and patterns

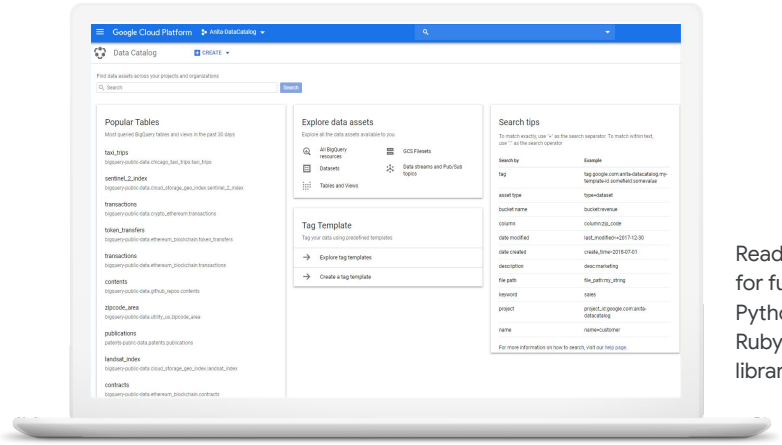


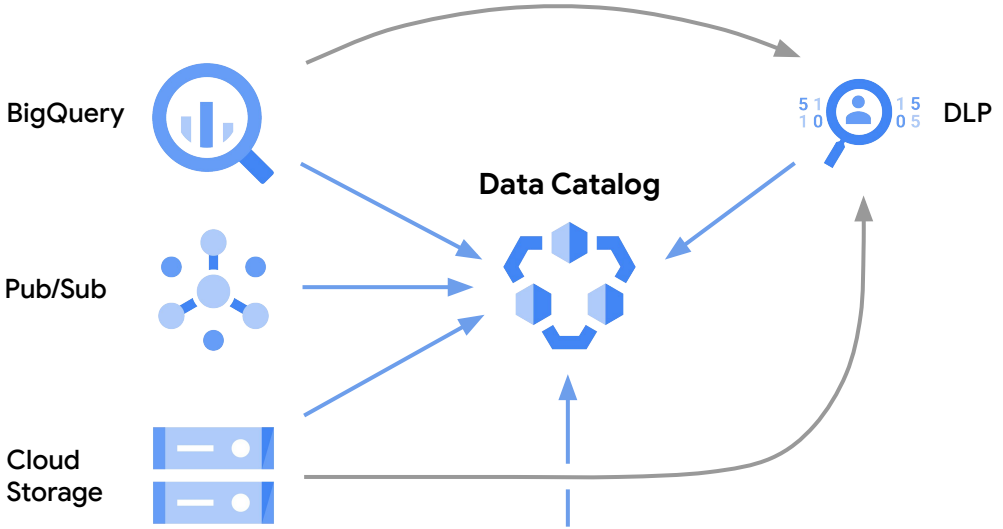
Google Cloud


DATA CONSUMERS
Google Data Studio
AppSheet
Looker
LookML
Google Cloud
DATA SHOP
Authorized Views Datasets
Data Layers
Cloud Storage
- Instant Archive Retrieval
- Lifecycle Rules
- Signed-URLs
- Geo-Redundancy
- Private Google Network
BigQuery
- Expiration Rules
- Partitioning/Clustering
- Serverless/Autoscaling
- No Tuning/No Ops
- Ad-Hoc Materialization
- Compressed Storage
- Cost Savings
Staging
INGEST & PROCESS
Storage Transfer Service
Pub/Sub
IoT Core
Data Transfer Service
Interactive
Dataprep
Data Fusion
Code
Dataproc
Dataflow
DATA SOURCES
AWS S3
Azure Blob
Other
IoT Data
Streaming
S3
teradata.
Redshift
APPLICATIONS
salesforce
workday.
Google Marketing Platform
100+ Apps...

Data Catalog

1. Auto-syncs ‘technical metadata’ from GCP data assets in near real-time
2. Auto-tags PII data through DLP integration
3. Supports non-GCP data assets through open-source connectors





Read, Write, and Search **API** for full metadata access: Python, Java, and Node.js, Go, Ruby, C#, and PHP language libraries

Hive, Oracle, SQL Server, Teradata, Redshift, MySQL, PostgreSQL, Looker, Tableau

Google Cloud

Templates and Tags

Define *template* for creating tags

Attributes

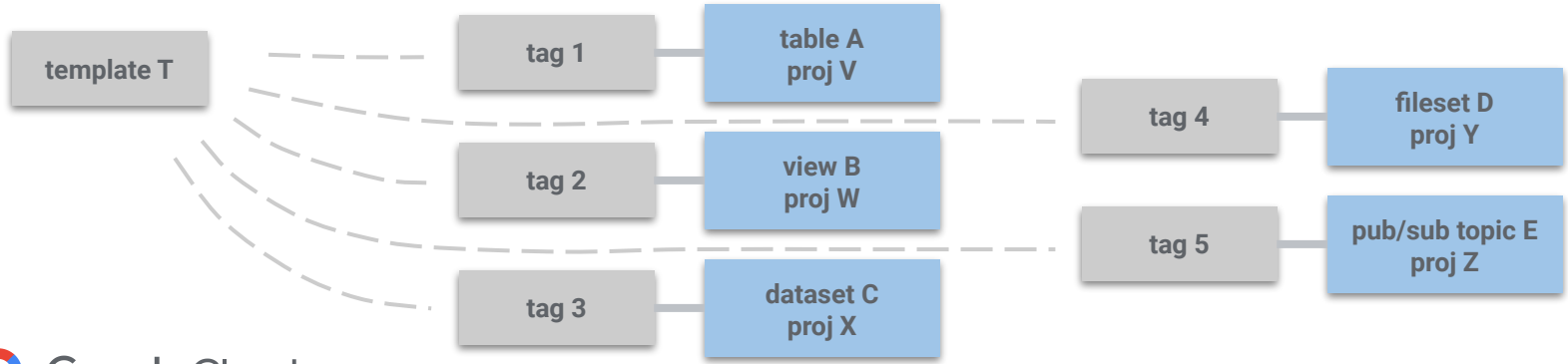
ID	Display Name	Type	
approved_by_data_governance	All approved data assets need to marked with this seal of approval. Data assets without this seal of approval will be deleted after 30 days. This is needed to enforce compliance	BOOL	
data_classification	Data Classification	ENUM	Values
data_lifecycle	Data Lifecycle	ENUM	
delete_by_date	This data asset needs to be deleted by the specified date	STRING	
has_pii	Has PII	BOOL	
retain_till_date	This data asset needs to be retained till the specified date	STRING	

Public
Sensitive
Confidential
Regulatory

Attach *tags* on individual data assets using a template

Data Governance Template (hr-240999108536-2.data_governance_template)

Attribute	Display name	Value
approved_by_data_governance	All approved data assets need to marked with this seal of approval - Data assets without this seal of approval will be deleted after 30 days - This is needed to enforce compliance	true
data_classification	Data Classification	Sensitive
data_lifecycle	Data Lifecycle	In_prod
delete_by_date	This data asset needs to be deleted by the specified date	10/21/2020
has_pii	Has PII	false
retain_till_date	This data asset needs to be retained till the specified date	11/21/2019

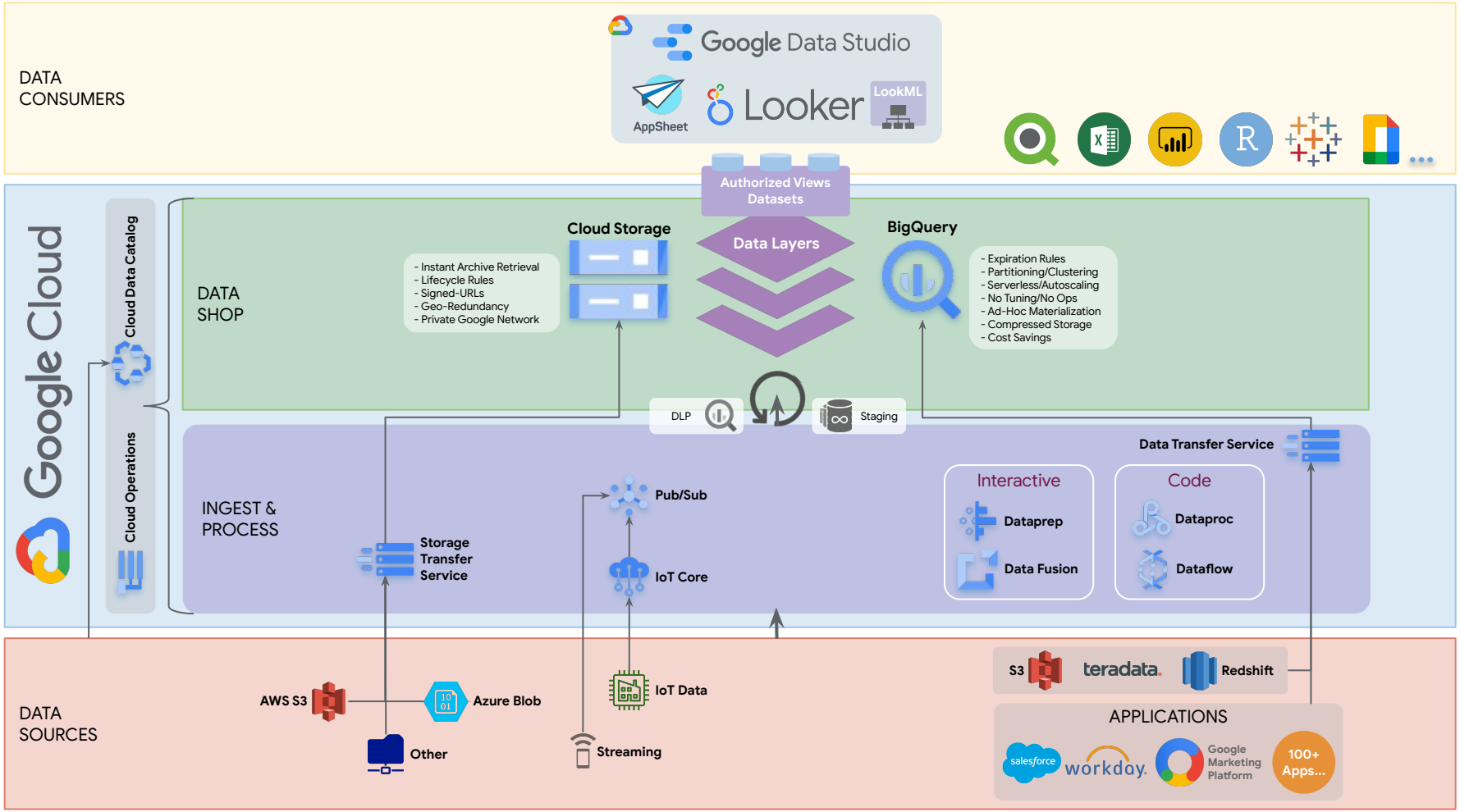


Google Cloud

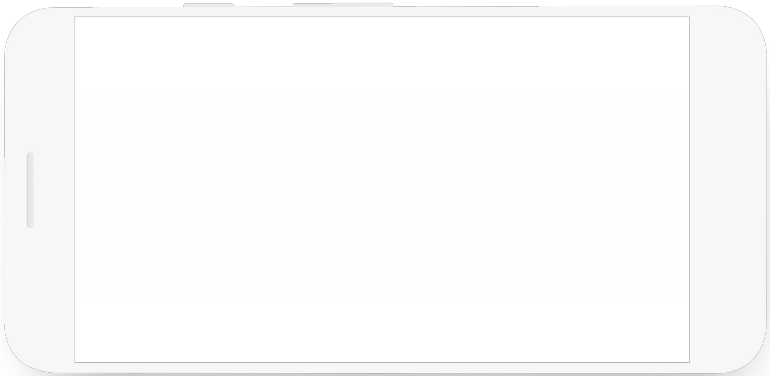
Anonymization made simple with DLP

- **Classification** of sensitive data like Personally Identifiable Information (PII)
- Data masking, format-preserving encryption, **transformations**
- Re-identification **risk analysis** (k-anonymity)

ID	Job Title	Phone	Comments
359740	Senior Engineer	307-964-0673	Please email them at jane@imadethisup.com
981587	VP, Engineer	713-910-6787	none
394091	Lawyer	692-398-4146	Updated phone to: 692-398-4146
986941	Senior Ops Manager	294-967-5508	none
490456	Junior Ops Manager	791-954-3281	Tried to verify account with their SSN 222-44-5555


DATA CONSUMERS
Google Data Studio
AppSheet
Looker
LookML
Google Cloud
Cloud Data Catalog
Cloud Operations
DATA SHOP
Authorized Views Datasets
Data Layers
Cloud Storage
- Instant Archive Retrieval
- Lifecycle Rules
- Signed-URLs
- Geo-Redundancy
- Private Google Network
BigQuery
- Expiration Rules
- Partitioning/Clustering
- Serverless/Autoscaling
- No Tuning/No Ops
- Ad-Hoc Materialization
- Compressed Storage
- Cost Savings
DLP
Staging
INGEST & PROCESS
Storage Transfer Service
Pub/Sub
IoT Core
Data Transfer Service
Interactive
Dataprep
Data Fusion
Code
Dataproc
Dataflow
DATA SOURCES
AWS S3
Azure Blob
Other
IoT Data
Streaming
S3
teradata.
Redshift
APPLICATIONS
salesforce
workday.
Google Marketing Platform
100+ Apps...

BigQuery ML for predictive analytics



1. **Execute** ML initiatives without moving data from BigQuery
2. **Iterate** on models in SQL in BigQuery to increase development speed
3. **Automate** common ML tasks, and hyperparameter tuning

Google Cloud

Built-in models

OR

Import TensorFlow model

Classification

Logistic regression

DNN classifier (Beta)

XGBoost classifier (Beta)

Other models

k-means clustering

Recommendation: Matrix factorization (Beta)

Regression

Linear regression

DNN regressor (Beta)

XGBoost classifier (Beta)

Model import

Import TensorFlow and XGBoost models for prediction (Beta)

```
CREATE MODEL project_id.mydataset.mymodel
 OPTIONS(
      MODEL_TYPE='TENSORFLOW',
      MODEL_PATH="gs://bucket/path/to/saved_model/*")
```

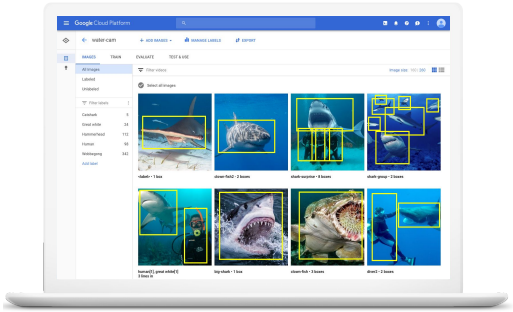
AI Platform for every level of expertise

our data + our models

Pre-trained APIs

No training data needed,
get started right away

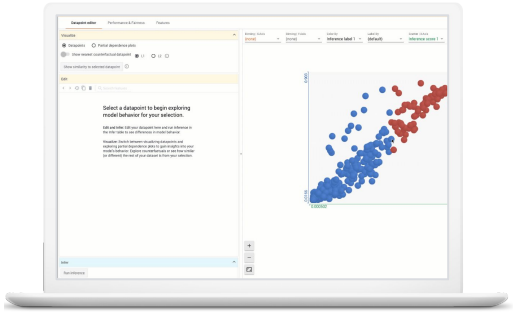
customer data + our models



Custom AI with AutoML

Easily create custom models
(A no-code approach)

customer data + customer model

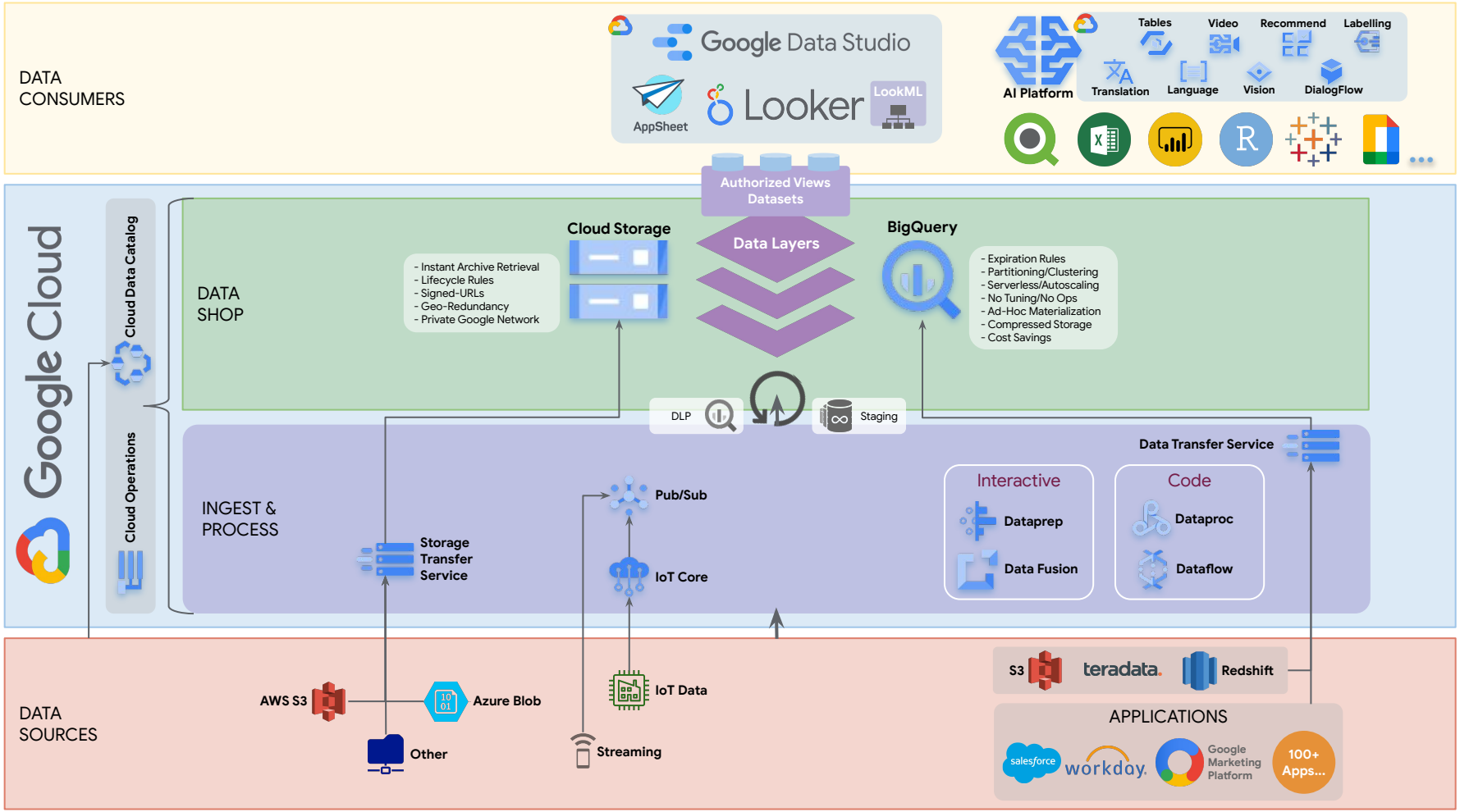


End-to-end AI with core tools

Help data scientists and ML
engineers build and deploy AI

BigQuery ML

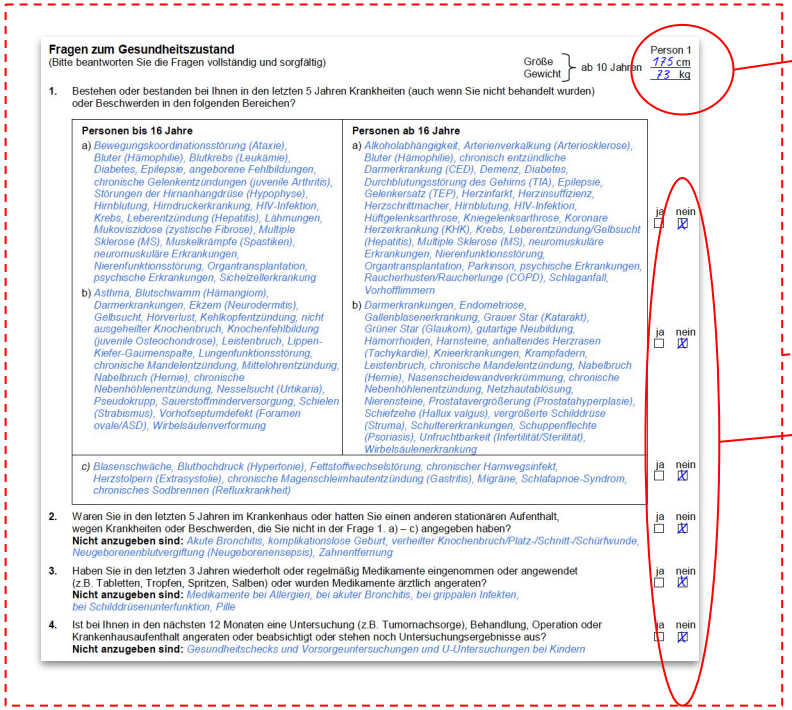
Google Cloud


DATA CONSUMERS
Google Data Studio
AppSheet
Looker
LookML
AI Platform
Tables
Video
Recommend
Labelling
Translation
Language
Vision
DialogFlow
Google Cloud
Cloud Data Catalog
Cloud Operations
DATA SHOP
Authorized Views Datasets
Data Layers
Cloud Storage
- Instant Archive Retrieval
- Lifecycle Rules
- Signed-URLs
- Geo-Redundancy
- Private Google Network
BigQuery
- Expiration Rules
- Partitioning/Clustering
- Serverless/Autoscaling
- No Tuning/No Ops
- Ad-Hoc Materialization
- Compressed Storage
- Cost Savings
DLP
Staging
INGEST & PROCESS
Data Transfer Service
Interactive
Dataprep
Data Fusion
Code
Dataproc
Dataflow
Storage Transfer Service
Pub/Sub
IoT Core
DATA SOURCES
AWS S3
Azure Blob
Other
IoT Data
Streaming
S3
teradata.
Redshift
APPLICATIONS
salesforce
workday.
Google Marketing Platform
100+ Apps...

Printed/Handwritten Documents
Document AI
Vision AI
Text-to-Speech
Intranet/FAQ
DialogFlow
BigQuery
DialogFlow
DialogFlow
I need to know the length for screw SCH5678
Sure, screw SCH5678 is 10mm long. Do you need anything else?
Yes, I need to talk to an expert please!
Of course. Please wait...
Record Message
PDF
Documents
DialogFlow
I need to know the type of this plug
Sure, please take a picture of the plug
This is a DUB1234 plug
Take Picture
Natural Language
Speech-to-Text
Translation
Expert

Document AI - Challenges in Invoices or other documents

Where traditional document analysis can be boosted with Machine Learning



Fragen zum Gesundheitszustand
(Bitte beantworten Sie die Fragen vollständig und sorgfältig)

Größe / Gewicht } ab 10 Jahren — Person 1: 175 cm / 73 kg

1. Bestehen oder bestanden bei Ihnen in den letzten 5 Jahren Krankheiten (auch wenn Sie nicht behandelt wurden) oder Beschwerden in den folgenden Bereichen?

Personen bis 16 Jahre	Personen ab 16 Jahre
a) *Bewegungskoordinationsstörung (Ataxie), Bluter (Hämophilie), Blutkrebs (Leukämie), Diabetes, Epilepsie, angeborene Fehlbildungen, chronische Gelenkentzündungen (juvenile Arthritis), Störungen der Hirnanhangdrüse (Hypophyse), Hirnblutung, Hirndruckerkrankung, HIV-Infektion, Krebs, Leberentzündung (Hepatitis), Lähmungen, Mukoviszidose (zystische Fibrose), Multiple Sklerose (MS), Muskelkrämpfe (Spastiken), neuromuskuläre Erkrankungen, Nierenfunktionsstörung, Organtransplantation, psychische Erkrankungen, Sichelzellerkrankung*	a) *Alkoholabhängigkeit, Arterienverkalkung (Arteriosklerose), Bluter (Hämophilie), chronisch entzündliche Darmerkrankung (CED), Demenz, Diabetes, Durchblutungsstörung des Gehirns (TIA), Epilepsie, Gelenkersatz (TEP), Herzinfarkt, Herzinsuffizienz, Herzschrittmacher, Hirnblutung, HIV-Infektion, Hüftgelenksarthrose, Kniegelenksarthrose, Koronare Herzerkrankung (KHK), Krebs, Leberentzündung/Gelbsucht (Hepatitis), Multiple Sklerose (MS), neuromuskuläre Erkrankungen, Nierenfunktionsstörung, Organtransplantation, Parkinson, psychische Erkrankungen, Raucherhusten/Raucherlunge (COPD), Schlaganfall, Vorhofflimmern*
b) *Asthma, Blutschwamm (Hämangiom), Darmerkrankungen, Ekzem (Neurodermitis), Gelbsucht, Hörverlust, Kehlkopfentzündung, nicht ausgeheilter Knochenbruch, Knochenfehlbildung (juvenile Osteochondrose), Leistenbruch, Lippen-Kiefer-Gaumenspalte, Lungenfunktionsstörung, chronische Mandelentzündung, Mittelohrentzündung, Nabelbruch (Hernie), chronische Nebenhöhlenentzündung, Nesselsucht (Urtikaria), Pseudokrupp, Sauerstoffminderversorgung, Schielen (Strabismus), Vorhofseptumdefekt (Foramen ovale/ASD), Wirbelsäulenverformung*	b) *Darmerkrankungen, Endometriose, Gallenblasenerkrankung, Grauer Star (Katarakt), Grüner Star (Glaukom), gutartige Neubildung, Hämorrhoiden, Harnsteine, anhaltendes Herzrasen (Tachykardie), Knieerkrankungen, Krampfadern, Leistenbruch, chronische Mandelentzündung, Nabelbruch (Hernie), Nasenscheidewandverkrümmung, chronische Nebenhöhlenentzündung, Netzhautablösung, Nierensteine, Prostatavergrößerung (Prostatahyperplasie), Schiefzehe (Hallux valgus), vergrößerte Schilddrüse (Struma), Schultererkrankungen, Schuppenflechte (Psoriasis), Unfruchtbarkeit (Infertilität/Sterilität), Wirbelsäulenerkrankung*

ja ☐ nein ☑ | ja ☐ nein ☑

c) *Blasenschwäche, Bluthochdruck (Hypertonie), Fettstoffwechselstörung, chronischer Harnwegsinfekt, Herzstolpern (Extrasystolie), chronische Magenschleimhautentzündung (Gastritis), Migräne, Schlafapnoe-Syndrom, chronisches Sodbrennen (Refluxkrankheit)*

ja ☐ nein ☑

2. Waren Sie in den letzten 5 Jahren im Krankenhaus oder hatten Sie einen anderen stationären Aufenthalt, wegen Krankheiten oder Beschwerden, die Sie nicht in der Frage 1. a) – c) angegeben haben?
Nicht anzugeben sind: *Akute Bronchitis, komplikationslose Geburt, verheilter Knochenbruch/Platz-/Schnitt-/Schürfwunde, Neugeborenenblutvergiftung (Neugeborenensepsis), Zahnentfernung*

ja ☐ nein ☑

3. Haben Sie in den letzten 3 Jahren wiederholt oder regelmäßig Medikamente eingenommen oder angewendet (z.B. Tabletten, Tropfen, Spritzen, Salben) oder wurden Medikamente ärztlich angeraten?
Nicht anzugeben sind: *Medikamente bei Allergien, bei akuter Bronchitis, bei grippalen Infekten, bei Schilddrüsenunterfunktion, Pille*

ja ☐ nein ☑

4. Ist bei Ihnen in den nächsten 12 Monaten eine Untersuchung (z.B. Tumornachsorge), Behandlung, Operation oder Krankenhausaufenthalt angeraten oder beabsichtigt oder stehen noch Untersuchungsergebnisse aus?
Nicht anzugeben sind: *Gesundheitschecks und Vorsorgeuntersuchungen und U-Untersuchungen bei Kindern*

ja ☐ nein ☑

Recognize handwritten records

Automatically recognize complete document type

Read form fields & check boxes

Google Cloud

The AI-Powered Enterprise Data Warehouse

DocAI + EKG + CMS = Unstructured Data ETL

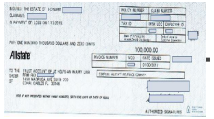
Capture
Unstructured Content

Process
Best-in-class AI

Store
Unified Data Lake

Analyze
Data Warehousing

Use
Advanced analytics



```
{ Type: Check
  Amount: $100
  To:Allstate
…}
```

```
{ Type: Check
  Amount: $100
  To: Allstate
    Insurance,Inc
…}
```

"10 checks indicate $100 payments to Allstate Insurance, Inc"

"We're seeing new payment patterns to All-state Insurance, Inc correlating with Jumbo Loan volumes in the North East"

Unstructured Documents

Document AI

Get structured data from unstructured content

Enterprise Knowledge Graph (EKG)

Normalize, validate & link entities across your data

Content Warehouse

Integrated unstructured + structured storage

Human in the Loop (HITL)

Comprehensive tooling for human review of AI model creation & outputs

BigQuery
analysis engine

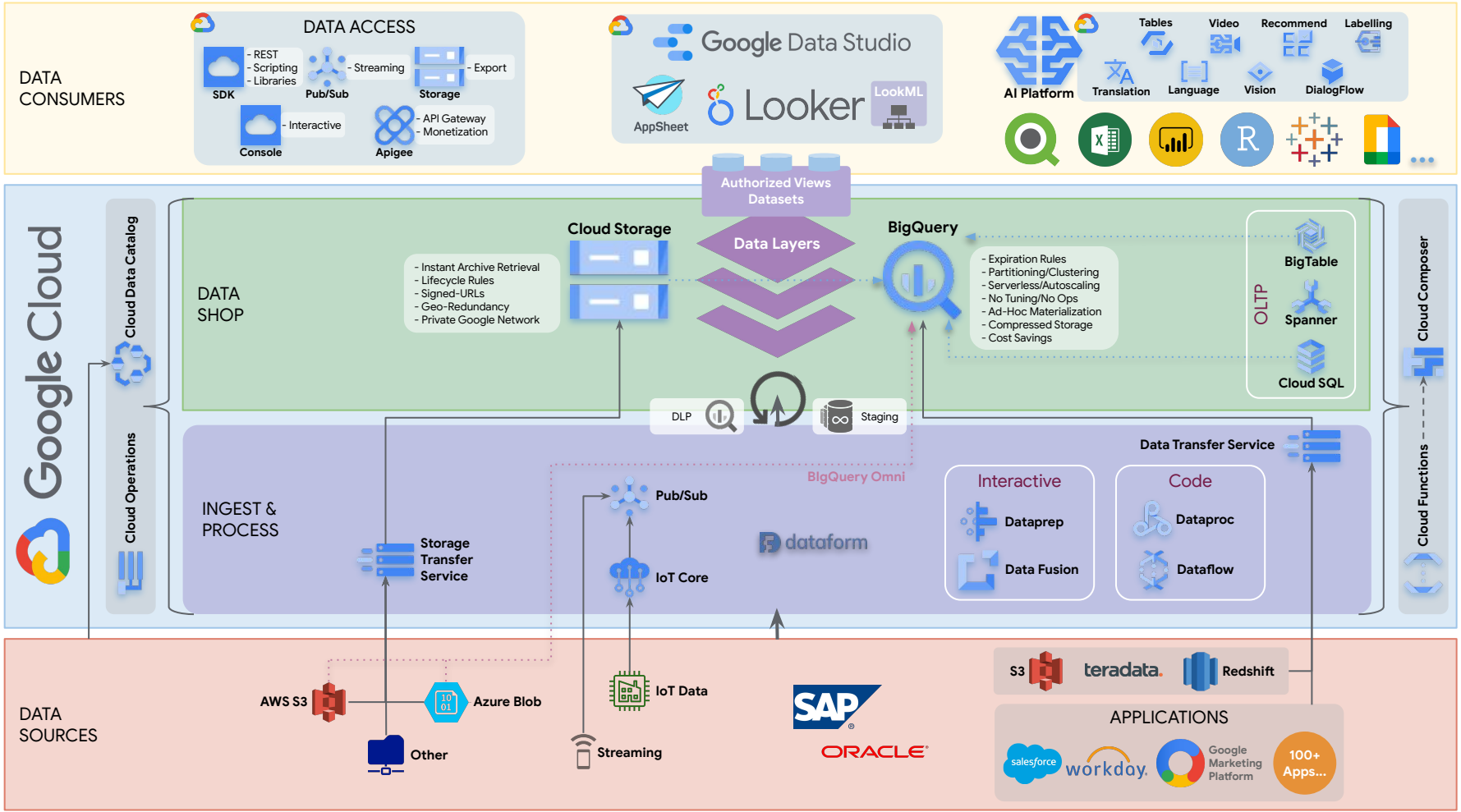
looker

Unified Analytics

Easily join structured & unstructured data into analysis, models, and processes

Google Cloud

Transformative Data Platform



Swiss steel upgrades its supply chain systems to improvement procurement and optimize efficiency

- Resulted in expected total **cost savings of nearly €500,000** for the first year
- Enabled **safer work environment** with smart devices for employee reporting and data collection
- Increased ability to find bottlenecks through networked supply chain, creating **cost saving opportunities**

> Google Cloud Platform improved our supply chain workflow and helped us avoid numerous indirect costs.”

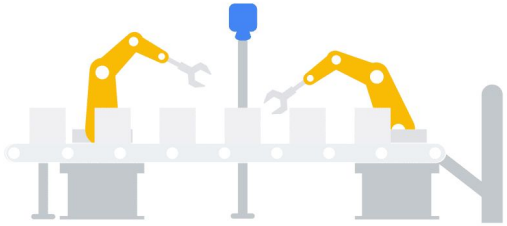
Daniel Jung, Head of Strategic Procurements
Swiss Steel

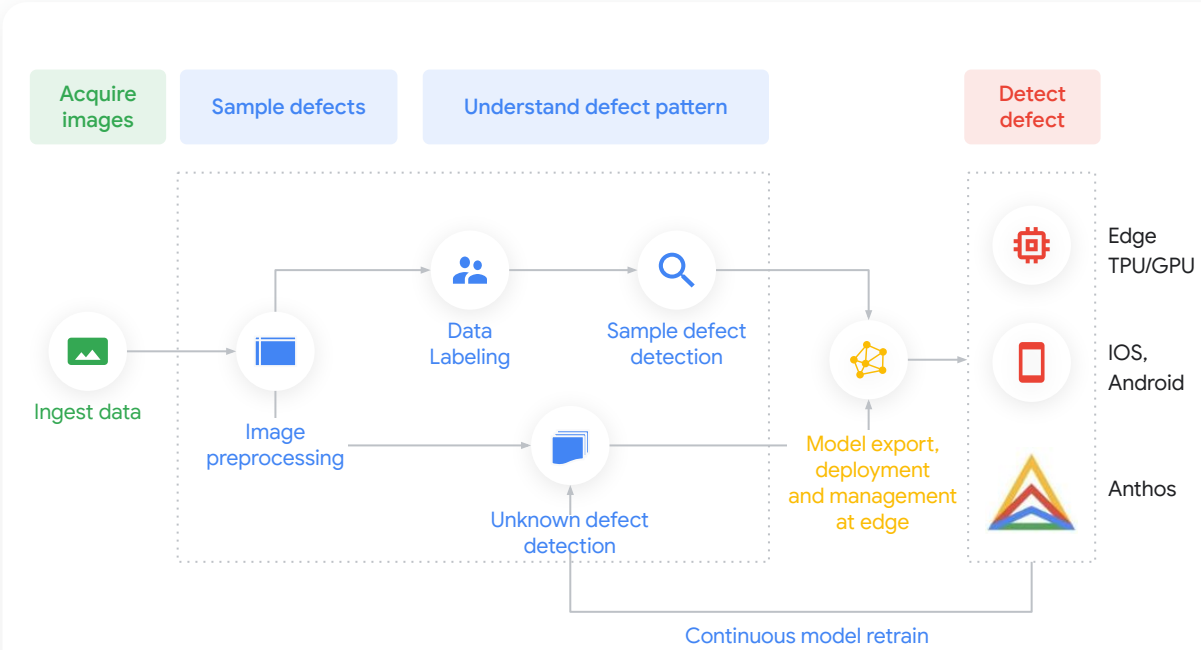
SWISS STEEL

Google Cloud

Use our best-in-class AI/ML platform to increase automation

Example: Google Cloud visual inspection for production processes





Google Cloud

Google Cloud Visual Inspection makes it easy for manufacturers to extract value from AI

High Accuracy

Train high quality defect detection model optimized for inspection use cases

Less Training Data Required

Minimize the labeled data required for model training, powered by unsupervised learning and transfer learning

On-Premises Support

Run inference on prem to provide low latency

Quick Setup

Set up and enable production workload within several weeks

High Scalability

Scale up or down easily with Google Cloud's flexible infrastructure to optimize cost

Integration with Google Cloud Products

Leverage solutions like Cloud AI Data Labeling, Anthos, and ML Edge for factory automation

Google Cloud

Google Cloud is ranked as a Leader for Vision AI products

Google

Wave position
LEADER

- Data
- Capabilities
- Pretrained models
- Development
- Deployment
- Solutions
- Ease of use
- Vision
- Roadmap
- Market approach

Differentiated | On par | Needs improvement

Products evaluated
Vision and Video Intelligence (both API and AutoML)

REFERENCE QUOTES

"We have built a very robust data pipeline and modeling capabilities with a small team and extremely cost effectively."

"Google has been a very strong partner, and their PaaS services are more stable and complete on release than competitors'."

FORRESTER®

"[Google Cloud] Is the platform to pick, that can do the most for the most. From business users, to developers, and data scientists, Google has powerful CV tools that all of them can use."

Google Cloud

Source: Forrester, The Forrester New Wave™: Computer Vision (CV) Platforms Q4, 2019

LG CNS

LG CNS reduces manual inspection time with Google Cloud

Saved **$1M per product line** per year by using Google Edge TPUs and Cloud AI

Automated the defect detection process, with **99.9%** accuracy, a **60%** increase from a previous solution

Increased throughput by accelerating the visual inspection process

Google Cloud

Source: LG CNS

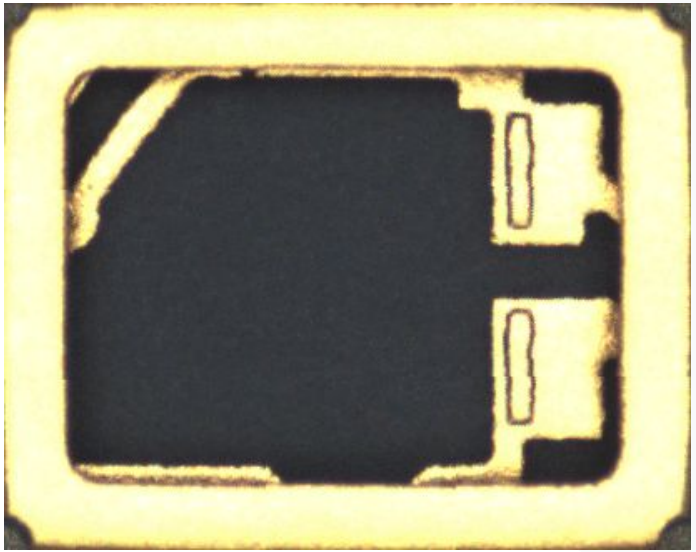
KYOCERA

Kyocera automates complex defect detection on electronic equipment

Achieved **98.5%** defect detection accuracy at an average prediction latency of **8.98** ms

Accelerated and simplified the manual product inspection process

Overcame detection challenges including multiple partially visible objects, constantly moving targets, and miniscule difficult-to-detect blemishes



Google Cloud

&

thyssenkrupp

That was so cool!

Let's rock together...

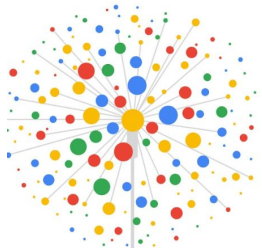
Google Cloud

Thyssenkrupp Steel Innovation Workshop Day 2 (7.7.)

Google Cloud

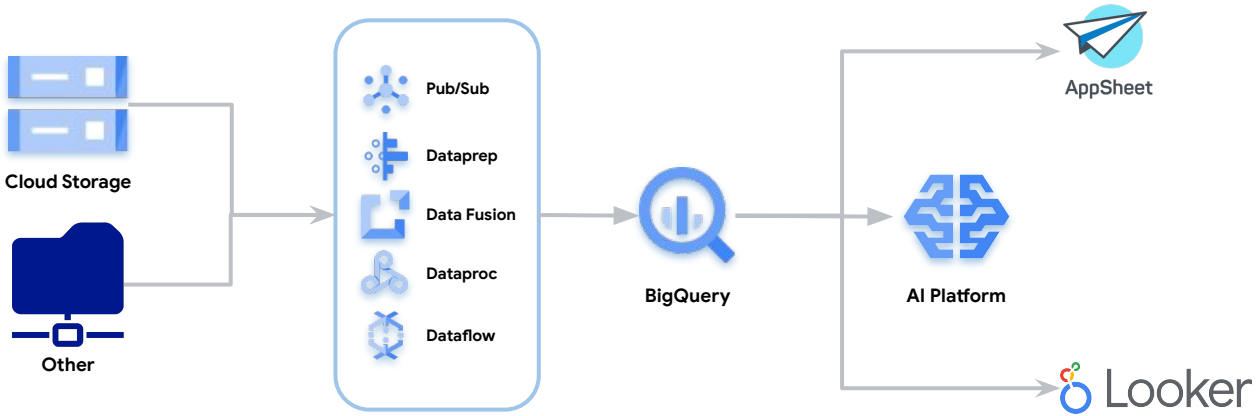
Agenda 7.7.

- 09:00 - 10:00 Recap 1.7. & offene Themen: Security, SAP
- 10:00 - 11:00 Google's BI Tool: Looker
- 11:00 - 12:00 API Management & No-Code Platform: Apigee & Appsheet
- 12:00 - 13:00 Q&A, Feedback, und nächste Schritte



Google Cloud

Solutions Recap



Google Cloud

BigQuery is named a Leader in The Forrester Wave™: Cloud Data Warehouse, Q1 2021

Google: a Leader in IaaS Platform Native Security

Of the 7 vendors evaluated, Google Cloud scored **highest overall in the current offering category**, and received the highest scores possible in the following current offering criteria: global data centers, guest OS & container protection, storage & data security, solution delivery, and navigation & integrated environment.

Google Cloud also received the **highest scores possible** in the strategy category criteria of: execution roadmap, market approach: developers and sales, partner ecosystem, commercial model, and plans in security posture management, hypervisor security, guest OS & container, and network security.

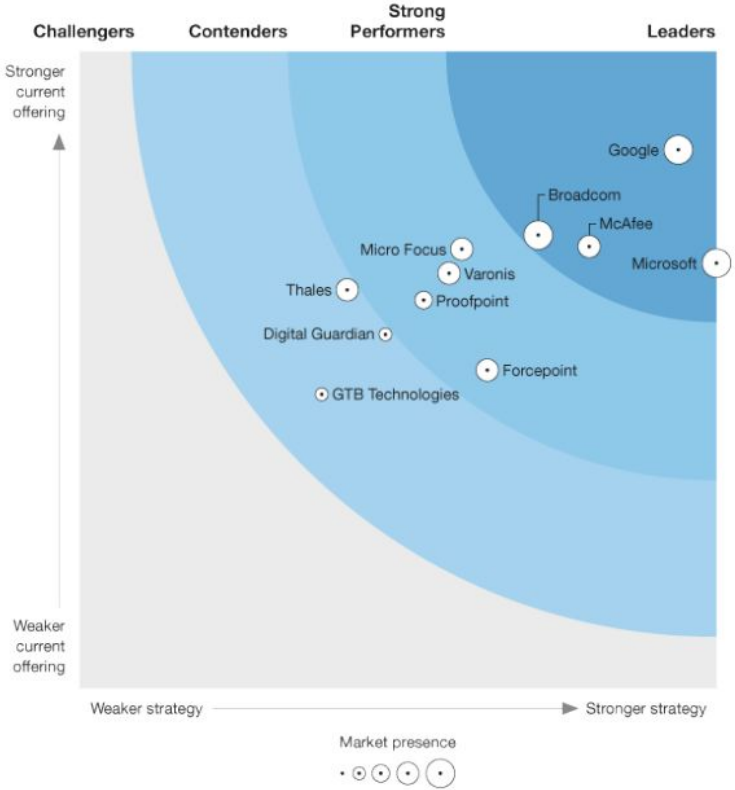
The Forrester Wave™: IaaS Platform Native Security Q4 2020. The Forrester Wave™ is copyrighted by Forrester Research, Inc. Forrester and Forrester Wave™ are trademarks of Forrester Research, Inc. The Forrester Wave™ is a graphical representation of Forrester's call on a market and is plotted using a detailed spreadsheet with exposed scores, weightings, and comments. Forrester does not endorse any vendor, product, or service depicted in the Forrester Wave. Information is based on best available resources. Opinions reflect judgment at the time and are subject to change.

Google Cloud

Google: a Leader in securing unstructured data

Of the 11 vendors evaluated, Google Cloud scored **highest in the current offering category**, and received the highest scores possible in the following current offering criteria: data intelligence, access control, deletion, obfuscation, security and risk, and APIs and integrations.

Google Cloud also received the **highest scores possible** in the strategy category for: data security platform vision, data security execution roadmap, performance, planned enhancements, zero trust enabling partner ecosystem, and diversity, equity & inclusion.



The Forrester Wave™: Unstructured Data Security Platforms Q2 2021. The Forrester Wave™ is copyrighted by Forrester Research, Inc. Forrester and Forrester Wave™ are trademarks of Forrester Research, Inc. The Forrester Wave™ is a graphical representation of Forrester's call on a market and is plotted using a detailed spreadsheet with exposed scores, weightings, and comments. Forrester does not endorse any vendor, product, or service depicted in the Forrester Wave. Information is based on best available resources. Opinions reflect judgment at the time and are subject to change.

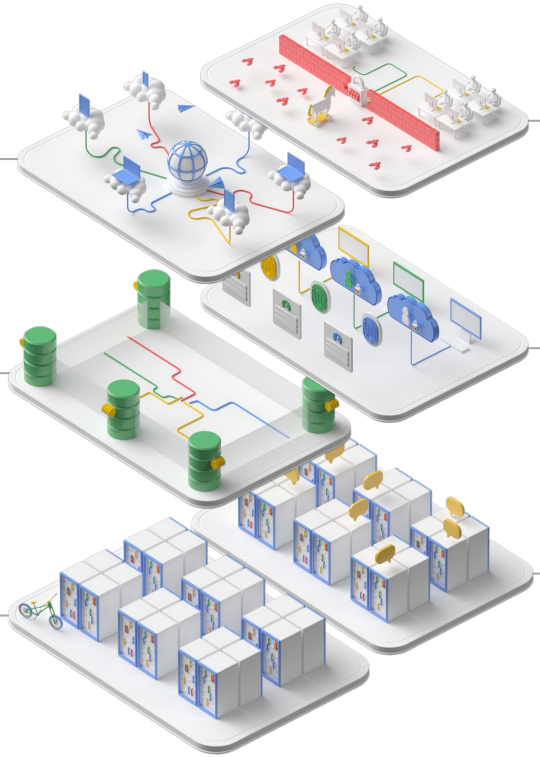
Google Cloud

Defense in depth at scale

Internet communication

Storage services

Hardware infrastructure



Operational & device security

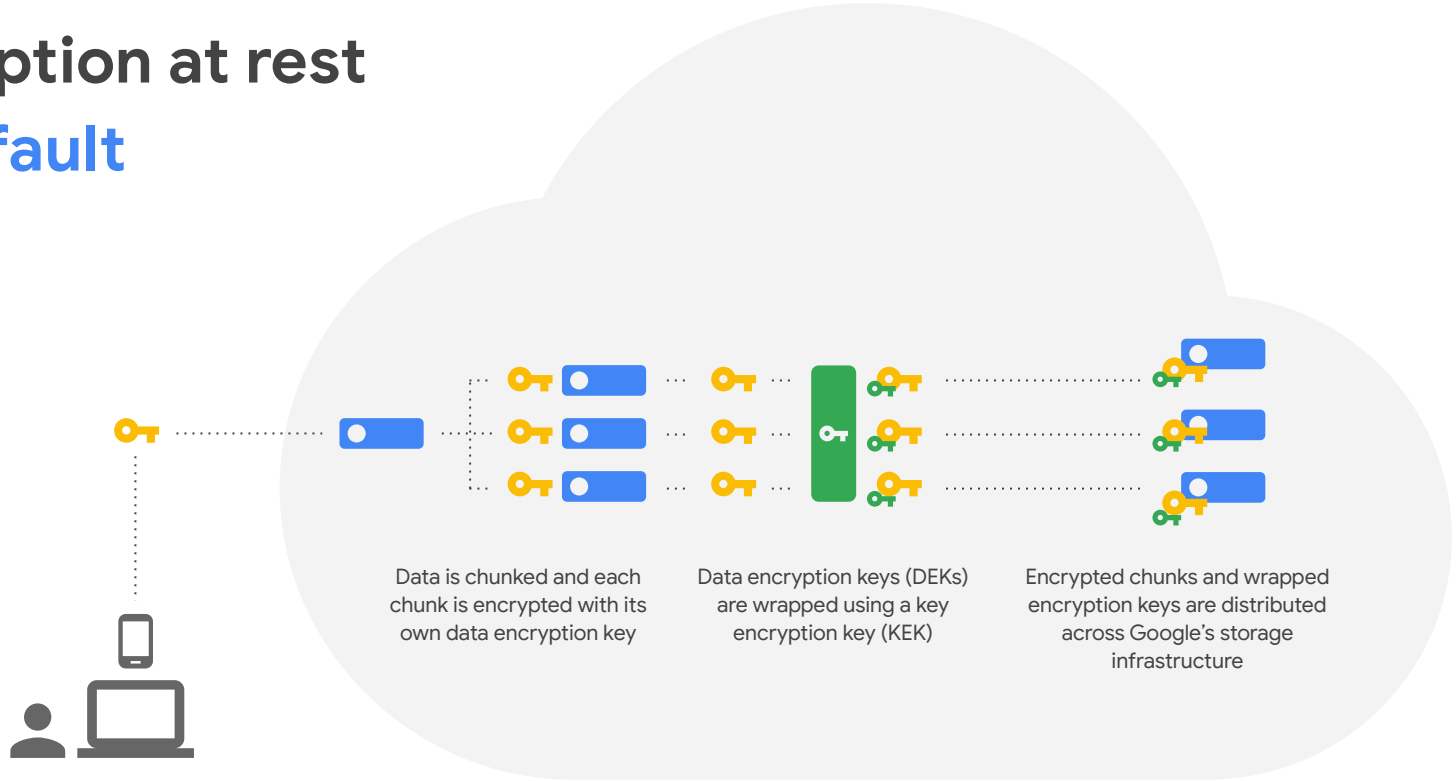
Identity

Service deployment

Google Cloud

Encryption at rest
by default

Connections require TLS



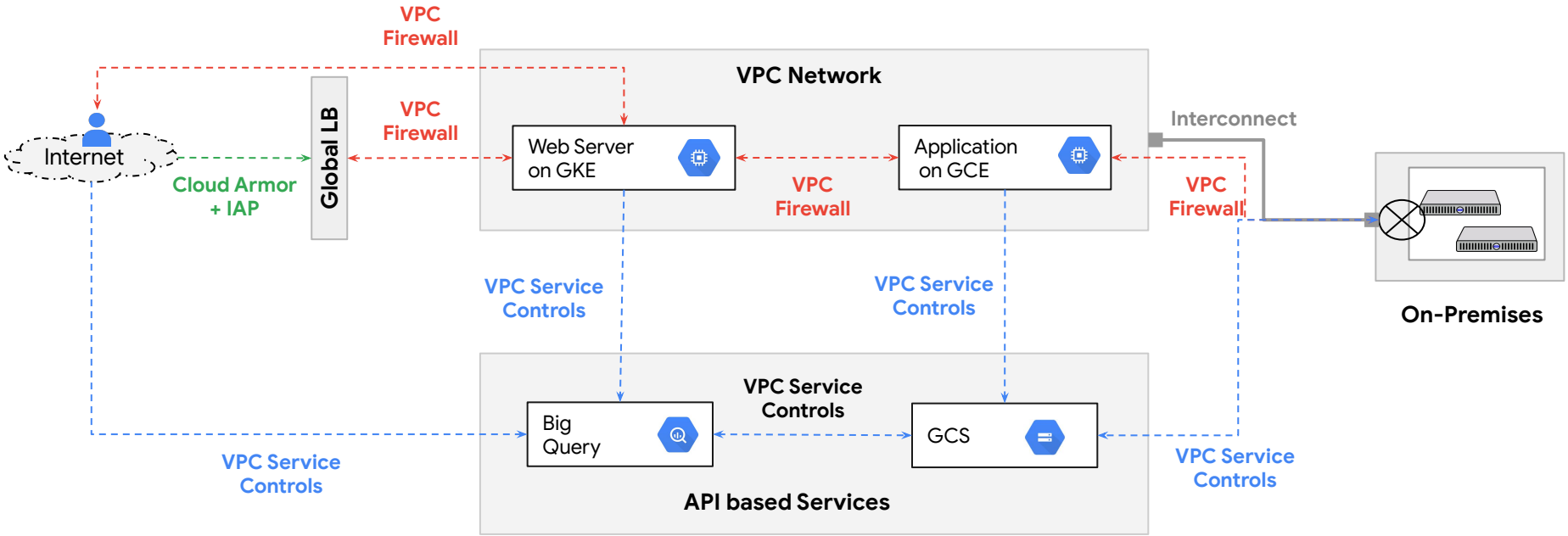
Google Cloud

Define a security perimeter with VPC Service Controls (GA)

Define security perimeters around sensitive data in Google Cloud Platform services

- Mitigate data exfiltration risks
- Privately access GCP services from on-premises
- Enforce context-aware access from the internet
- Centrally manage security policies

VPC Service Controls - Example Architecture



Google Cloud

SAP & GCP

SAP

SAP Tooling

Batch
- Data Services
- Data Hub / Data Intelligence
- BW OpenHub Service

Real Time
- SLT + Data Services
- SLT + Data Hub / Data Intelligence

Federation
- Smart Data Access

Visualization
- Analytics Cloud
- Business Objects
- Data Warehouse Cloud

Partner Tooling

On GCP Marketplace

- Informatica — IICS - Data Integration
- Qlik — Qlik Replicate
- Datavard — Datavard Glue
- software AG — webMethods

EMERGING

- HVR
- Dell Boomi
- Aecorsoft

Google Tooling

Batch
- Cloud Data Fusion
  - SAP Table Plugin (Via JCO) *Now GA*

Real Time
- BigQuery
  - Real Time Direct Connector for SAP (Via SLT) *Target GA H1 '22*

BigQuery

Google Cloud