

Google Cloud

Next '22

Democratizing Data with BigQuery and Data Mesh

Oct/
11–13





Vikas Kumar

AVP
CNA Insurance



Stuart Moncada

Group Product Manager,
Cloud Data Analytics
Google Cloud

What will you learn from this session?

01 Transformation Strategy

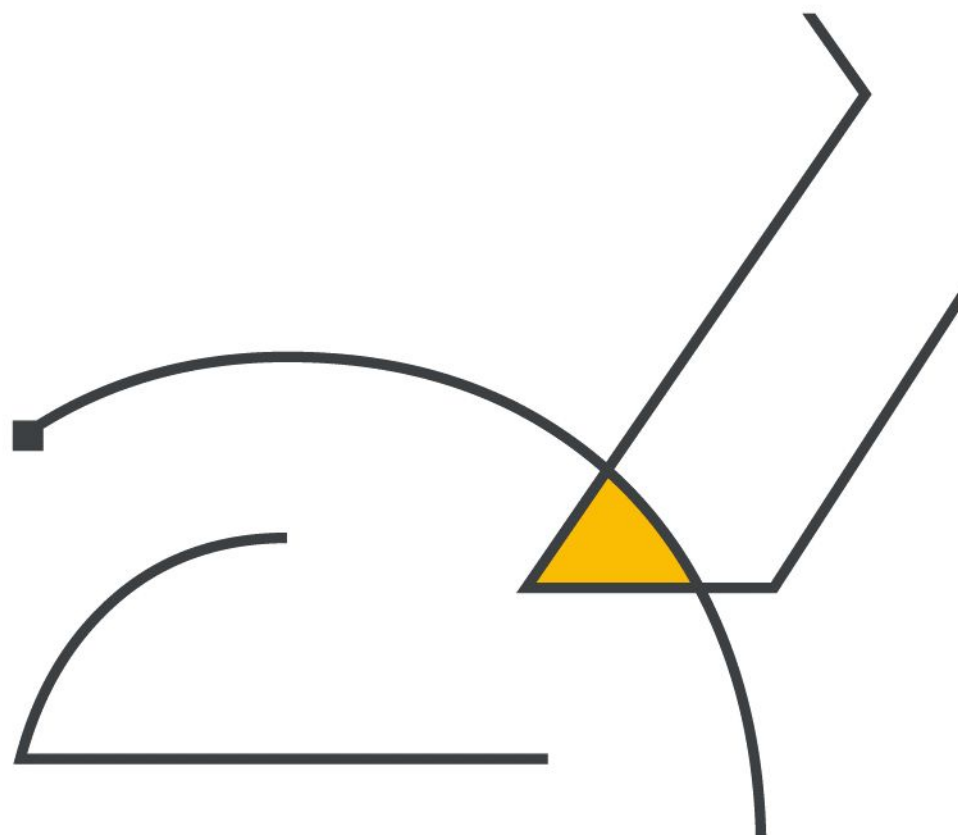
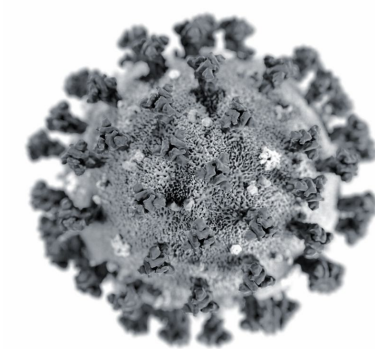
02 Cloud Journey

03 Organization Challenges

04 Business Value

05 Democratize Data and Platform

Data and Innovation more critical than ever




 Fortune

Tech industry's massive 2022 layoffs are sign of a 'boomerang' effect, LinkedIn economist says

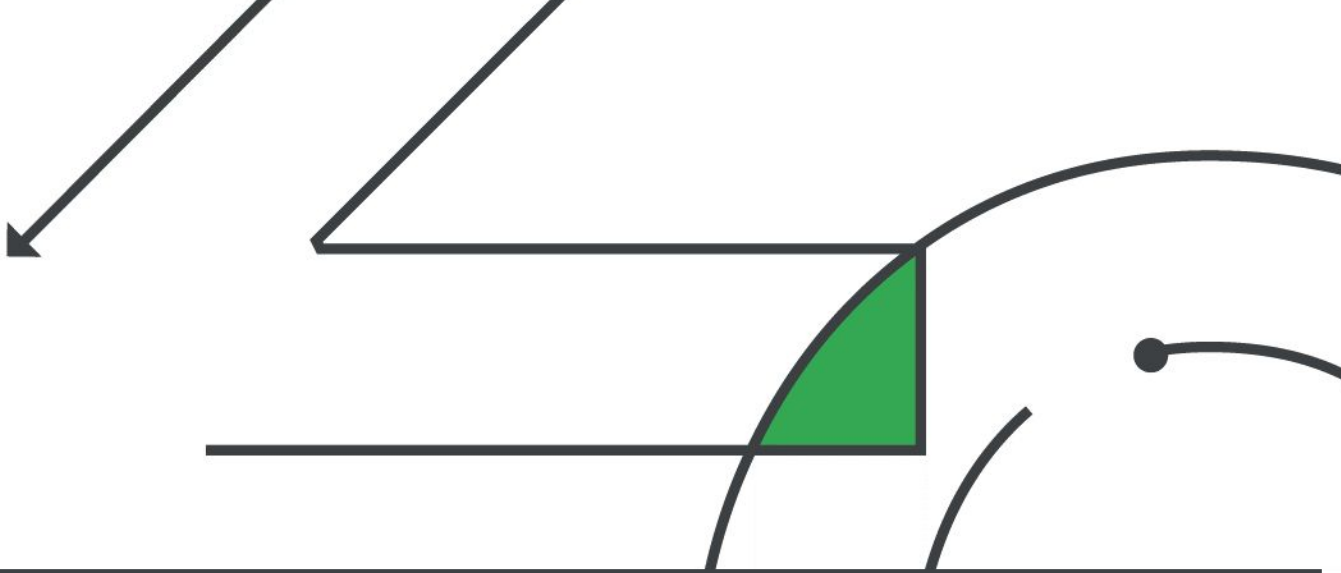
 CNBC

Inflation rose 8.6% in May, highest since 1981

 Wall Street Journal

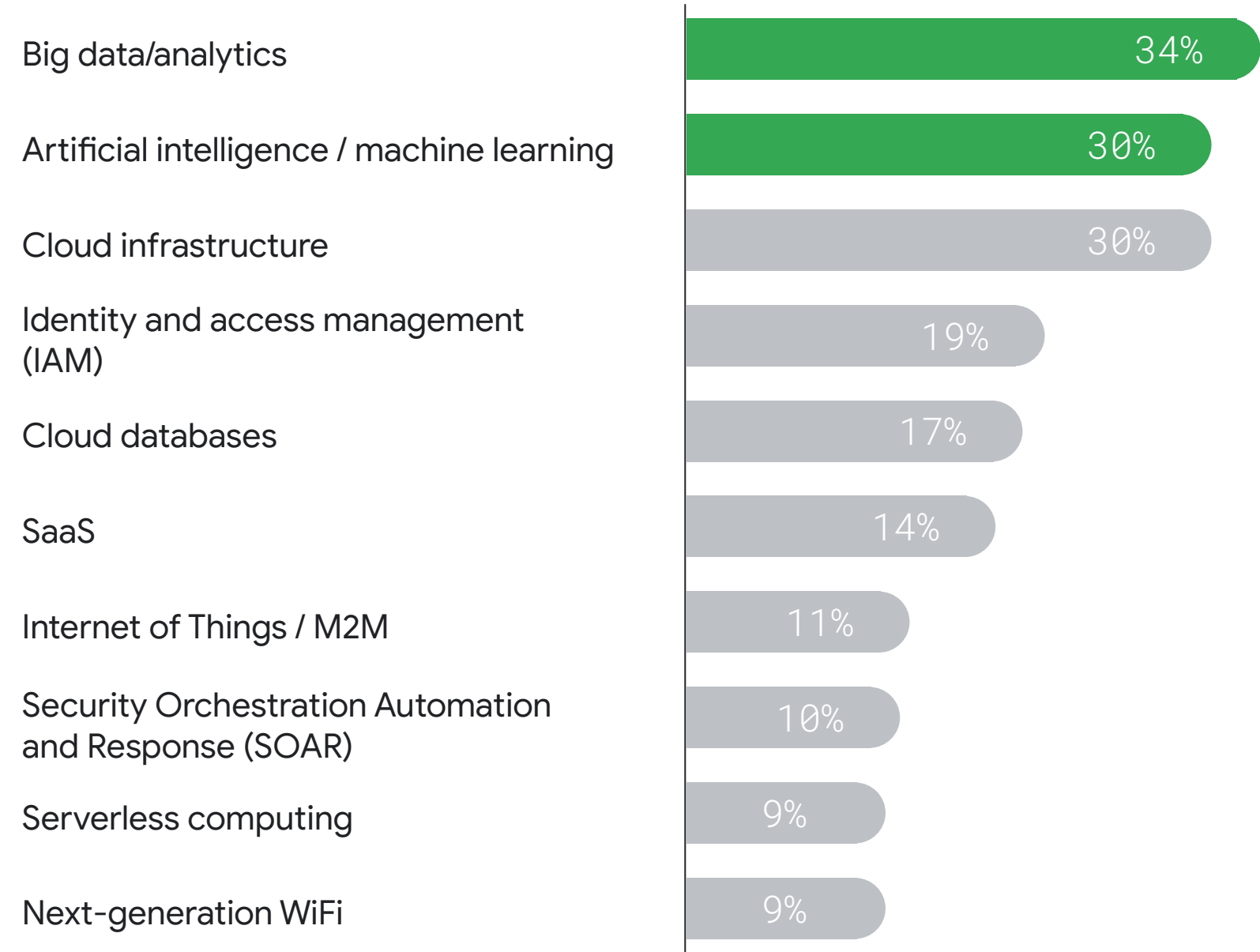
More Supply-Chain Disruptions Are Coming

Technology is altering the way organizations operate



Organizations see data and AI/ML potential

but are struggling to make it a reality



10%

of organizations achieve significant financial benefits from AI

Q: Which of these technologies has the most potential to significantly alter the way your business operates over the next 3 to 5 years?
CIO Magazine, Oct 2021

Boston Consulting Group
Are You Making the Most of Your Relationship with AI?, Oct 2020

Companies are unable to realize measurable value from data



Data is big and multi-format



Data requires more than SQL



Data reaches everyone

More data copies

More data silos

More capacity

More tech islands

More security risk

More integrations

High costs

Constant Capacity Planning

Low productivity

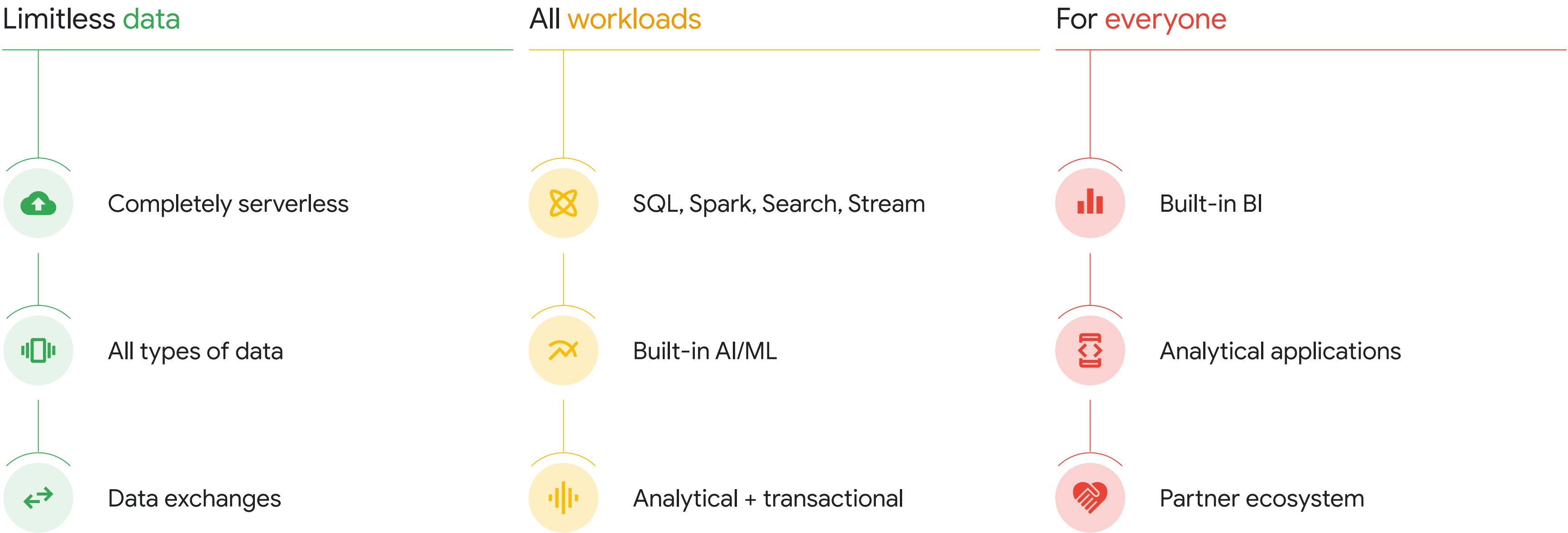
Limited access

Data unavailable

Poor SLAs

Unclear compliance

Characteristics of a modern data cloud



Cost effective | Highly productive | Always on | Easy to secure | Clear compliance | Open extensions

Limitless data

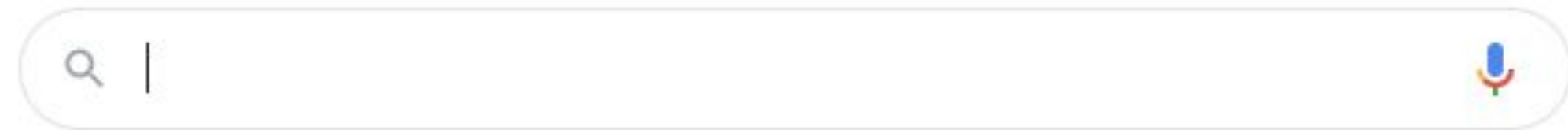
Completely Serverless

Serverless

- Simplifies capacity management
- Dynamically adjusts demand

Completely Elastic

- Auto-start and auto-pause
- Dynamically adjusts demand
- No performance cliff due to local capacity
- Immune to large-scale hardware failures

The Google logo is displayed in its standard multi-colored font.A standard Google search bar with a magnifying glass icon on the left and a microphone icon on the right.

Google Search

I'm Feeling Lucky

Limitless data

BigLake

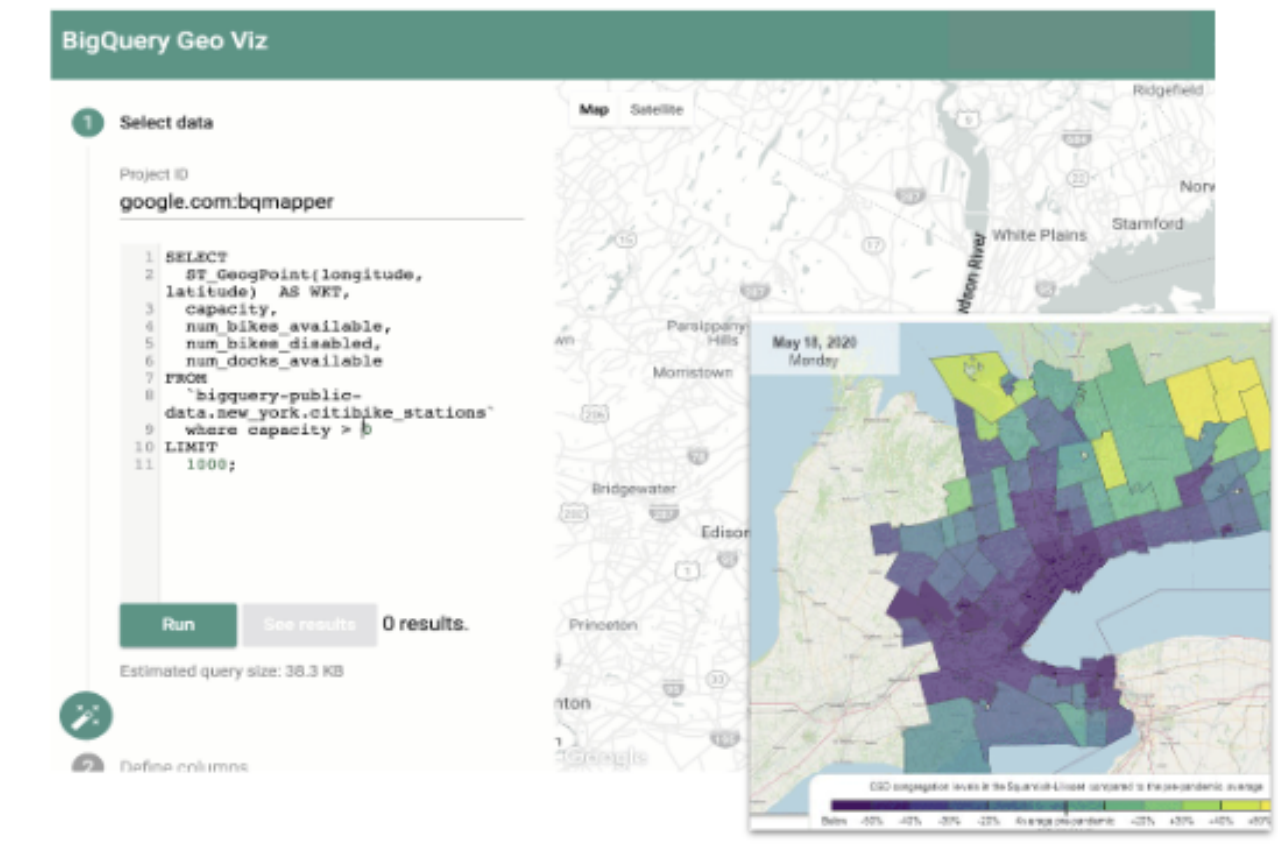
All your data types in one platform

- Structured
- Semi-structured
- Unstructured (text, images)
- Parquet
- JSON
- Nested Tables
- Geospatial

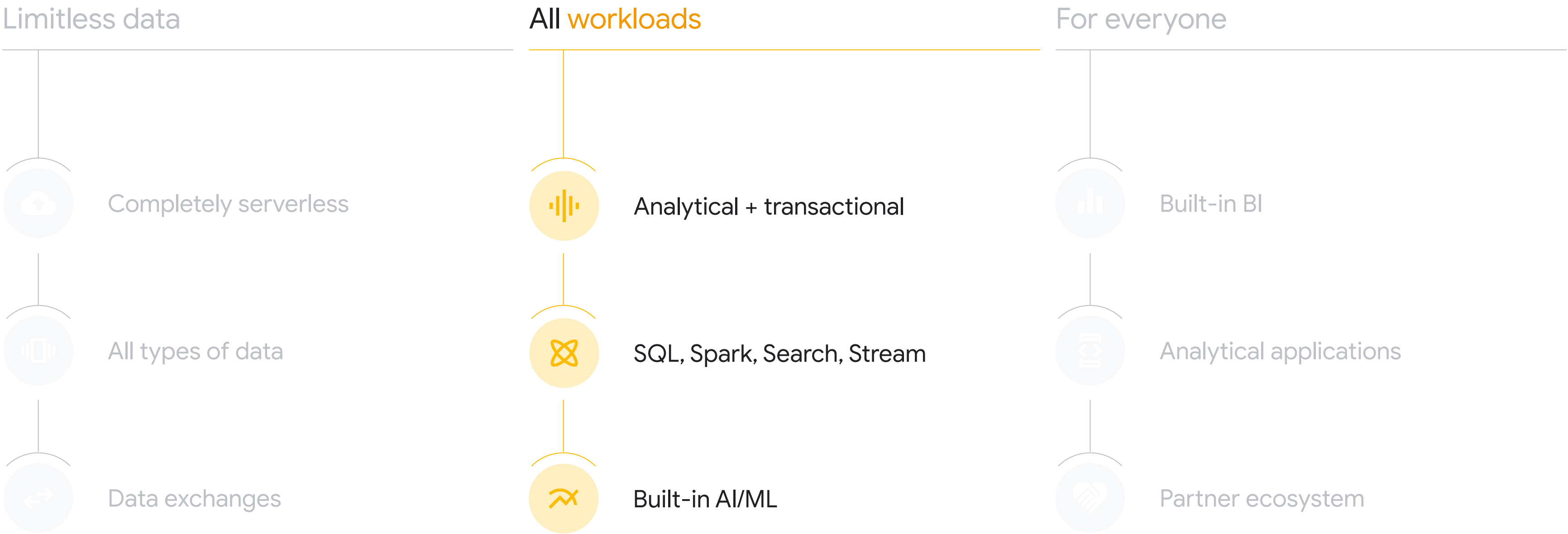
BigLake (Lakehouse)

Best components of a data lake and data warehouse

A modern and open lake house architecture reduces cost, and scales across more workloads efficiently.



Characteristics of a modern data cloud



Stream Analytics

Why Stream?

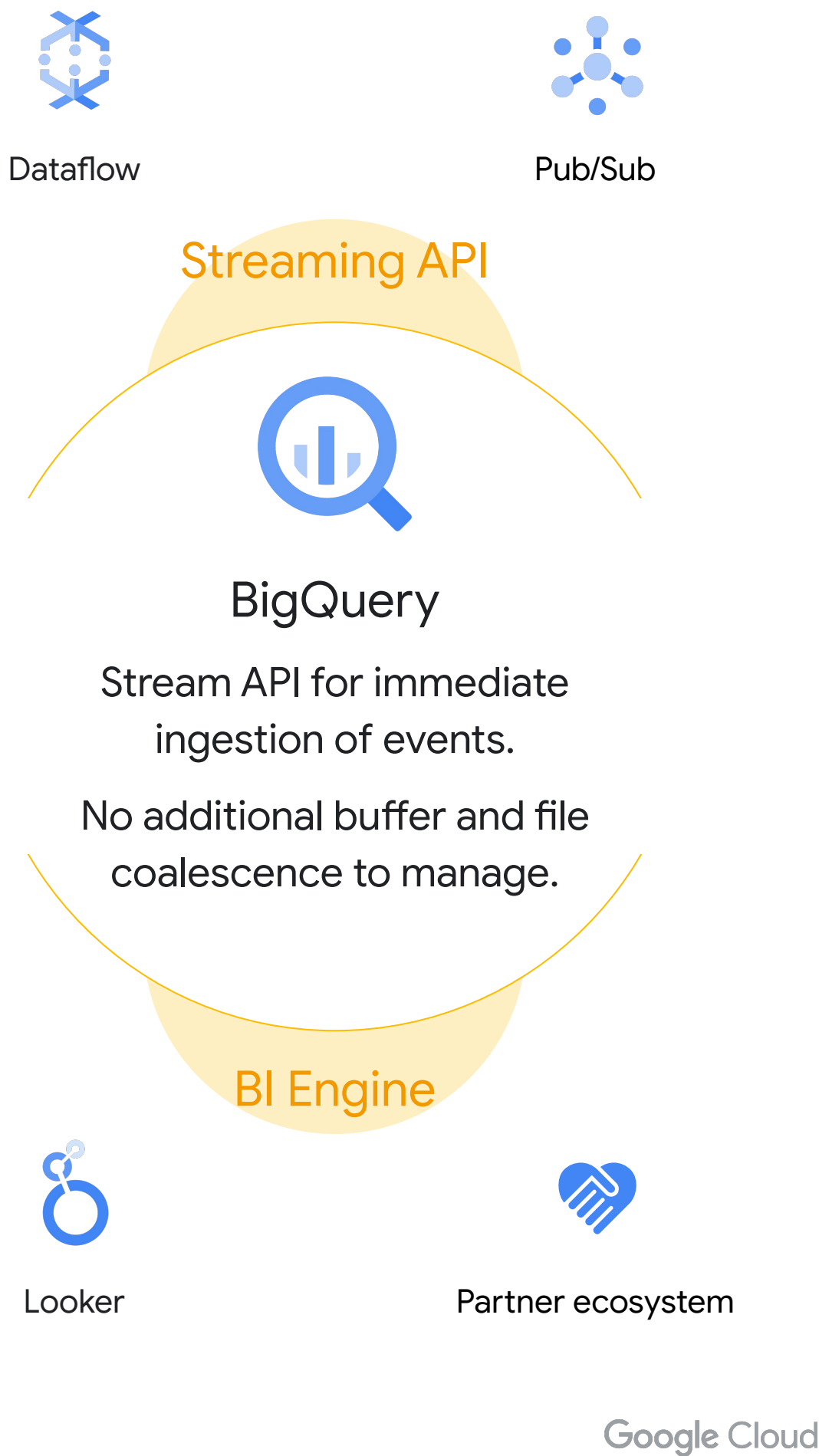
Leverage event-driven analysis with built-in streaming capabilities

vs

Leverage historical analysis and batch processing

Analyze business events in real-time

Move your business to event-driven action, logstream, clickstream, and sensor data to enable use cases like anomaly detection and continuous intelligence.



Built-in AI/ML

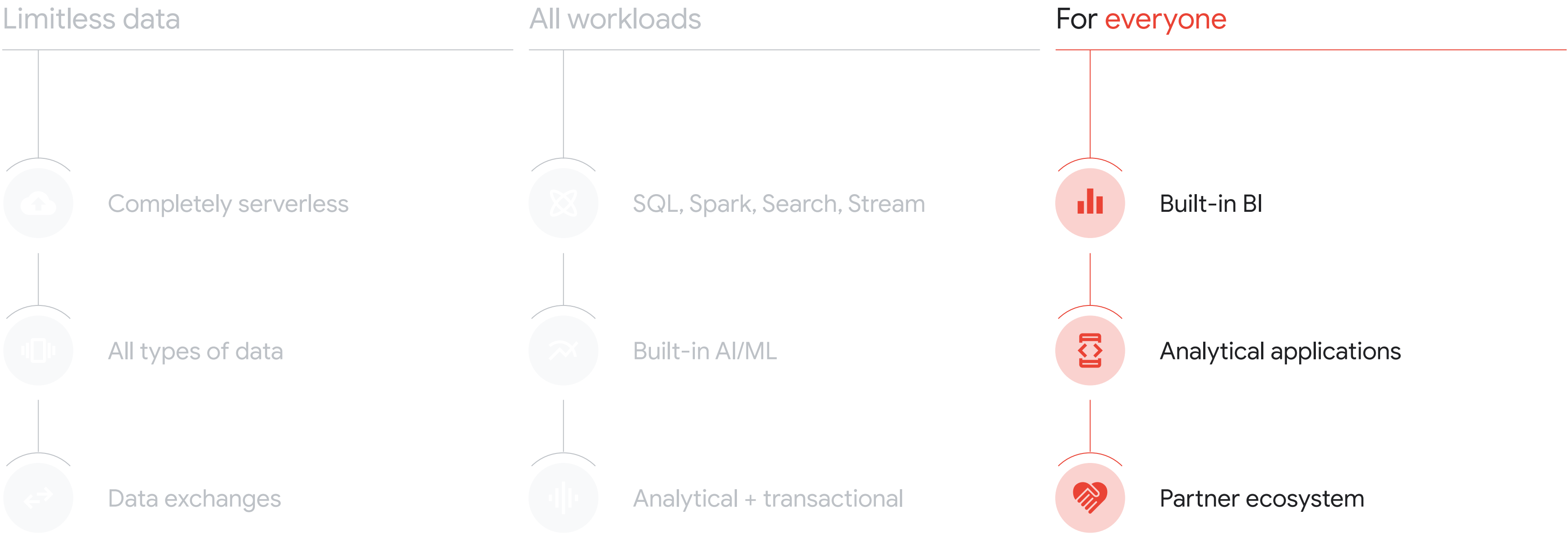
Machine Learning for all, Built-in ML with SQL

- Execute and automate ML initiatives within BigQuery using predefined models
- Leverage external models developed in Tensorflow directly from SQL
- Export developed models for use in Vertex AI

Direct integration with Vertex AI Workbench



Characteristics of a modern data cloud



Building the future requires a comprehensive transformation



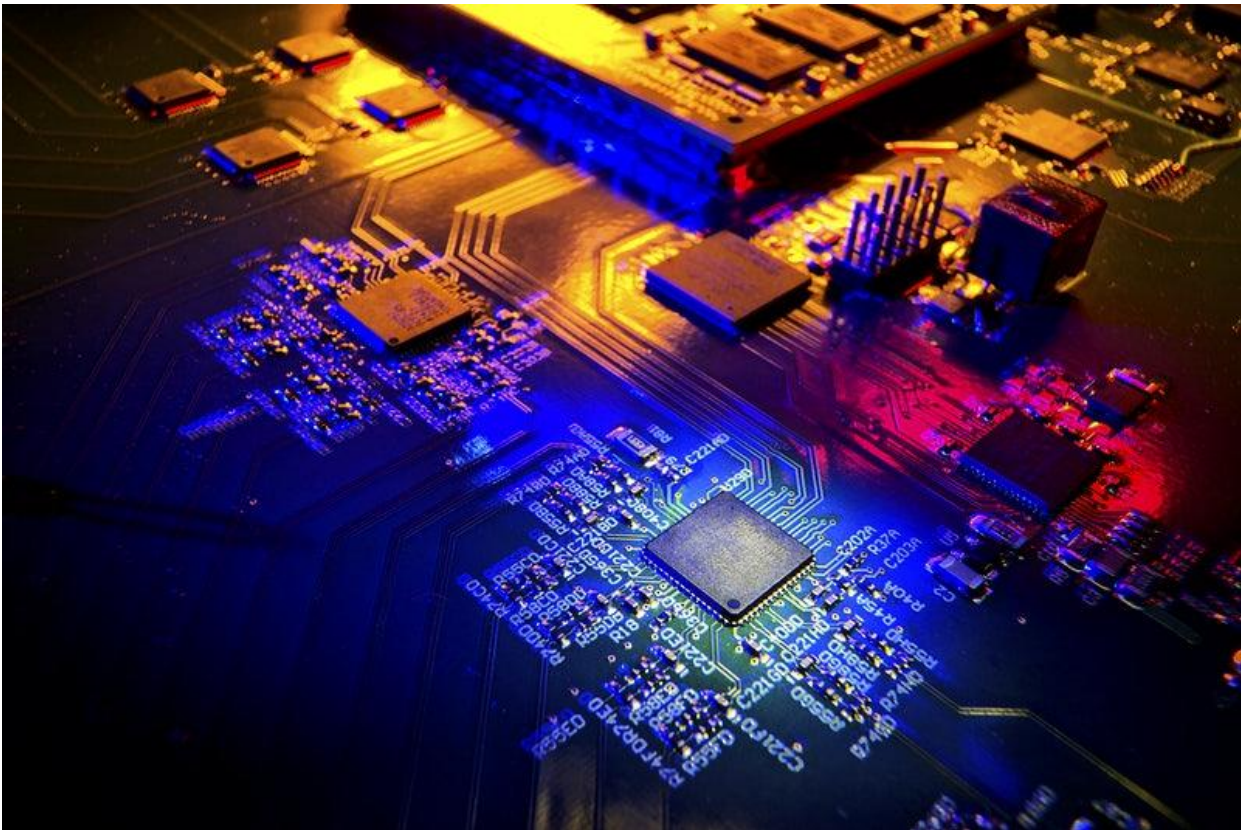
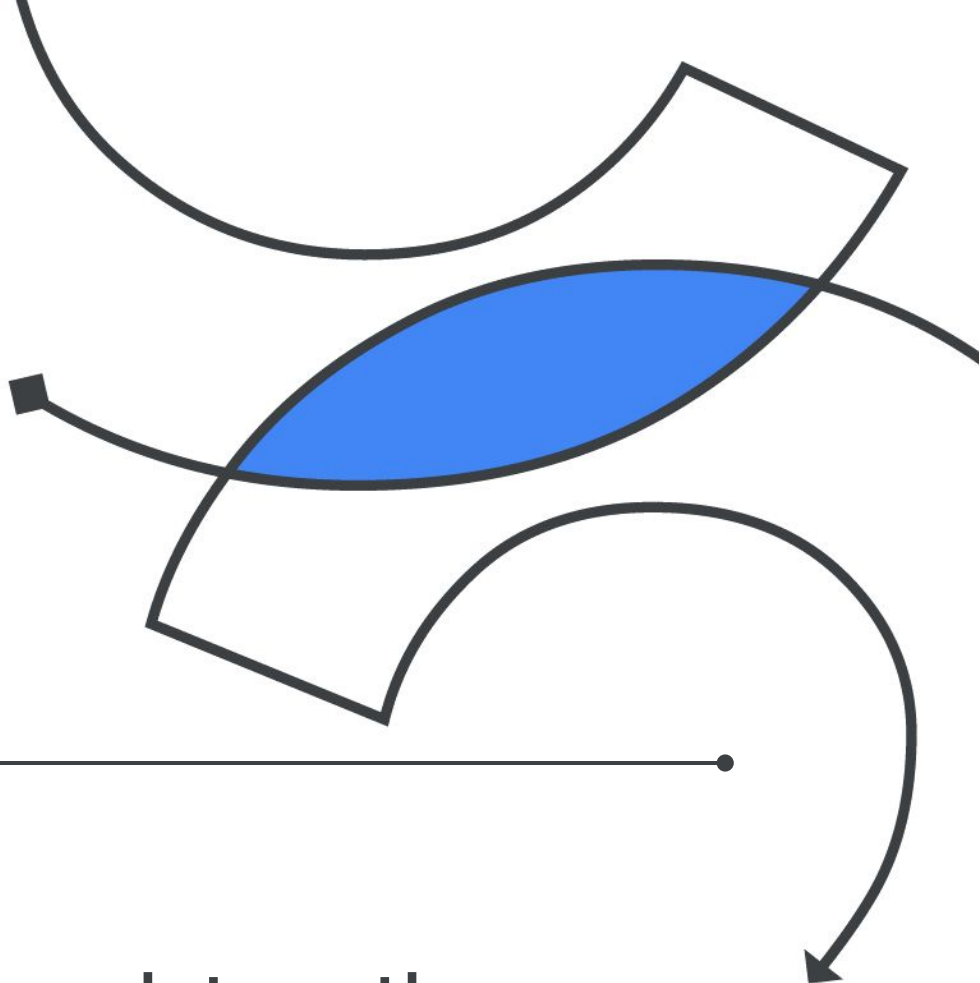
Talent



Technology Infrastructure



How we work together



Augment transformation with modernized work

Product based

- Agile and product
- Cloud Native
- Engineering Practices
- Modern architecture standards
- CICD and DevOps
- Advance Analytics
- Build with AI / ML



Framework for transformation approach

Inception of data transformation journey:

01

Formulate
transformation
strategy & kick off

02

Specify what success
looks like (prove value),
Socialize plan
& alignment

03

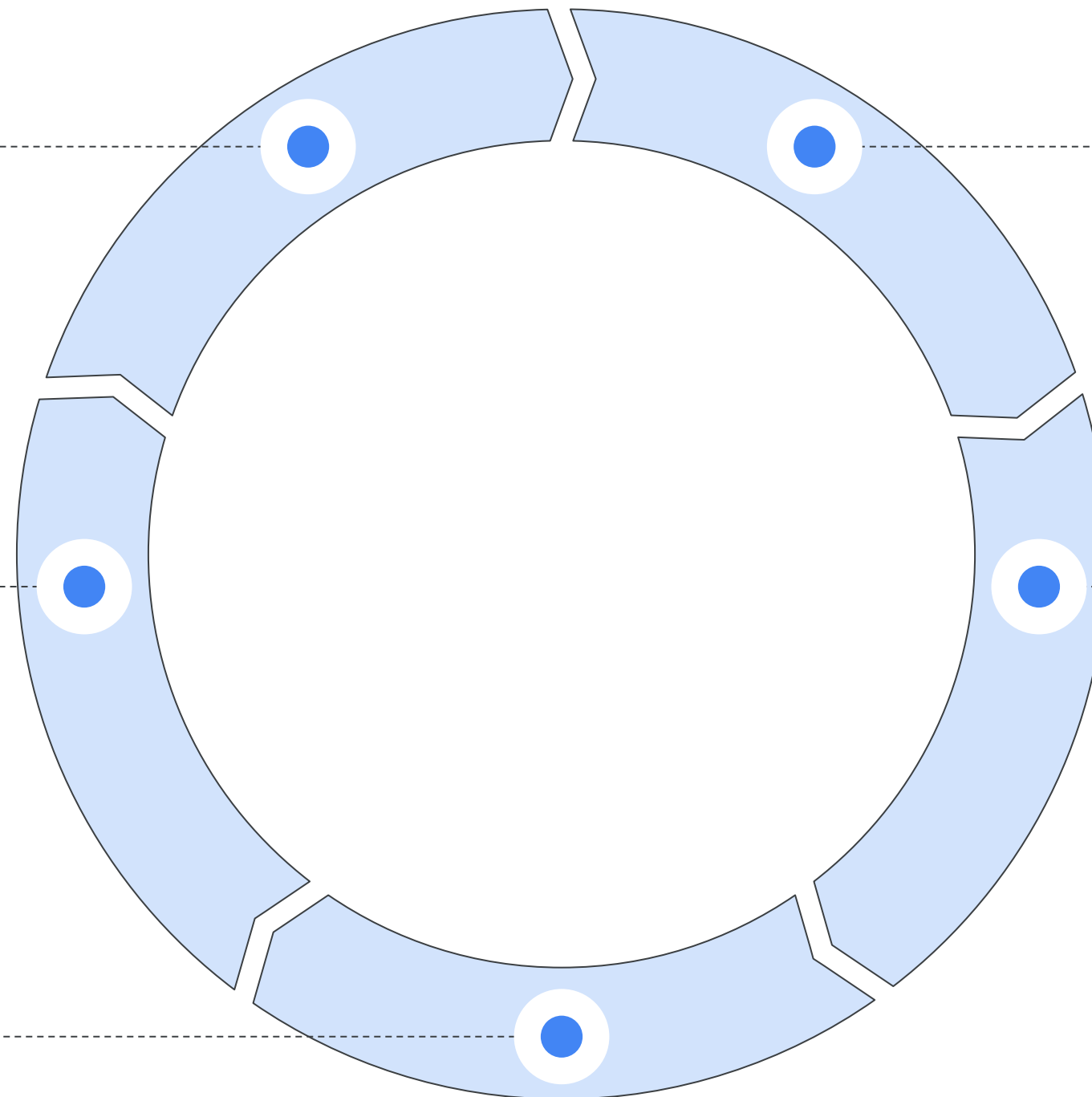
Identify system(s) and
deployment strategy, share
roll out plans with business
leaders

05

Feedback from business
and iterate

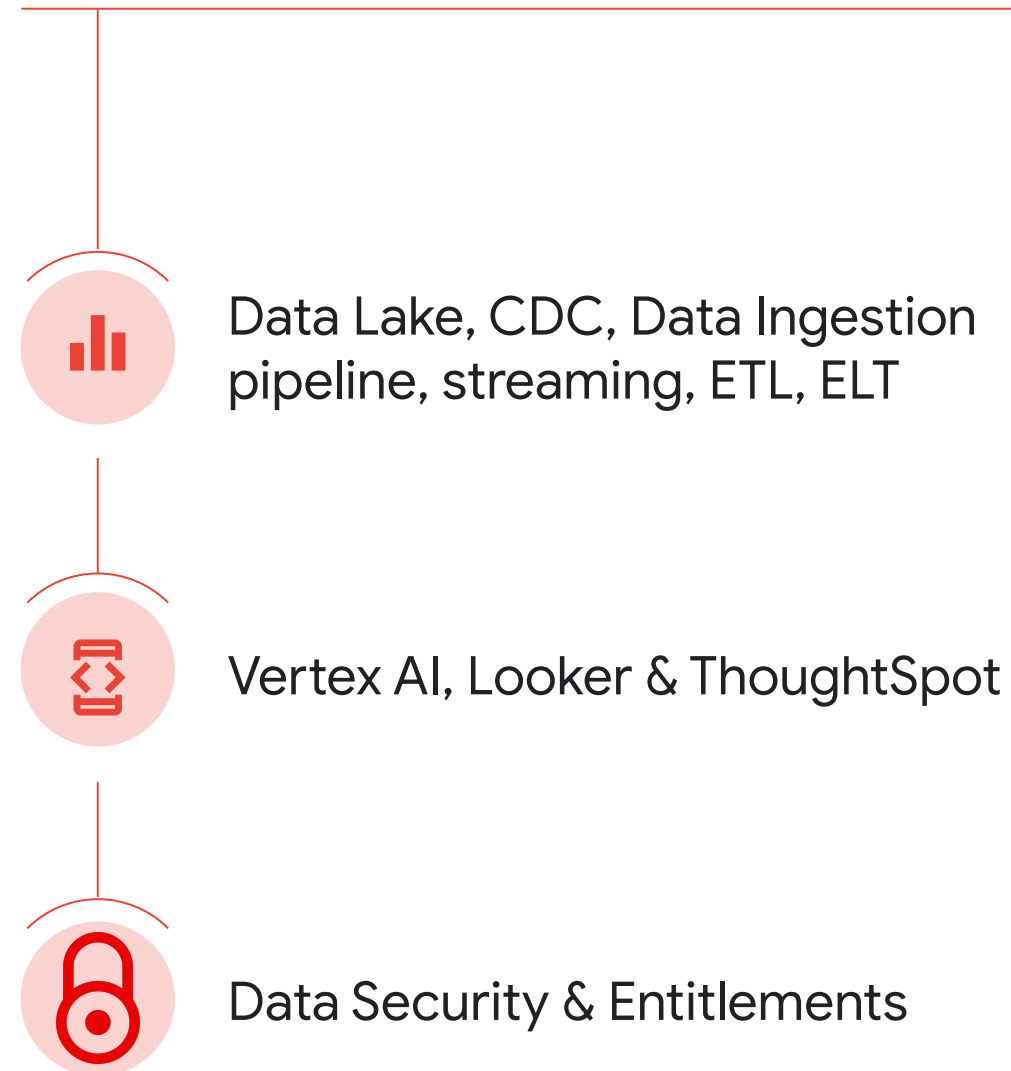
04

Build Platform and
execute

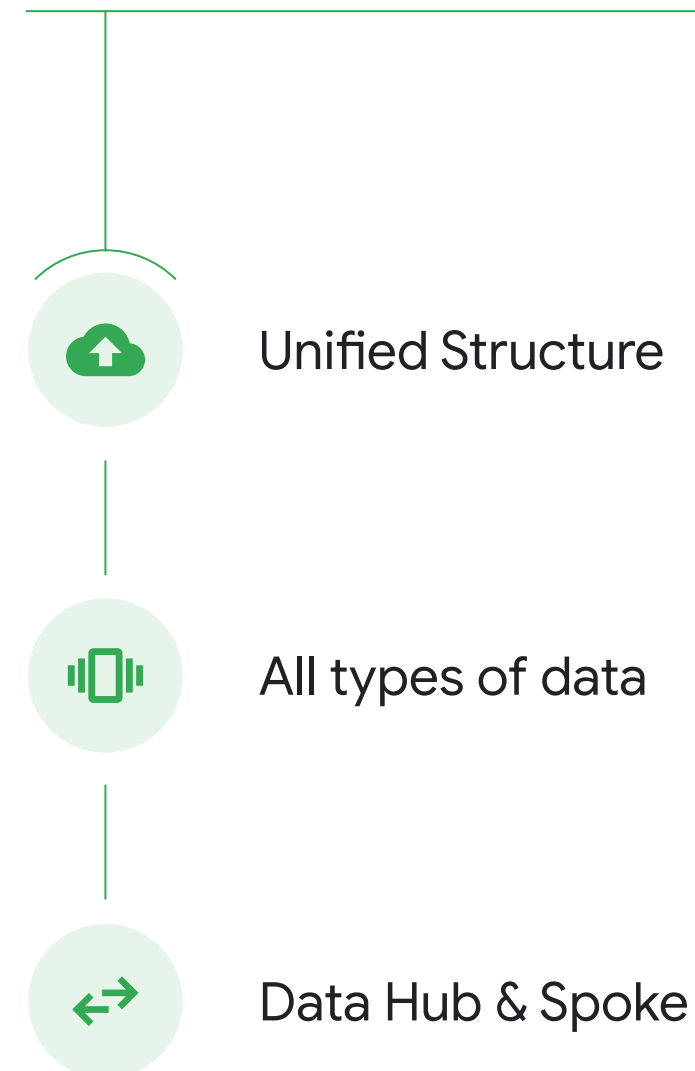


Democratize platform, data & AI/ML on cloud

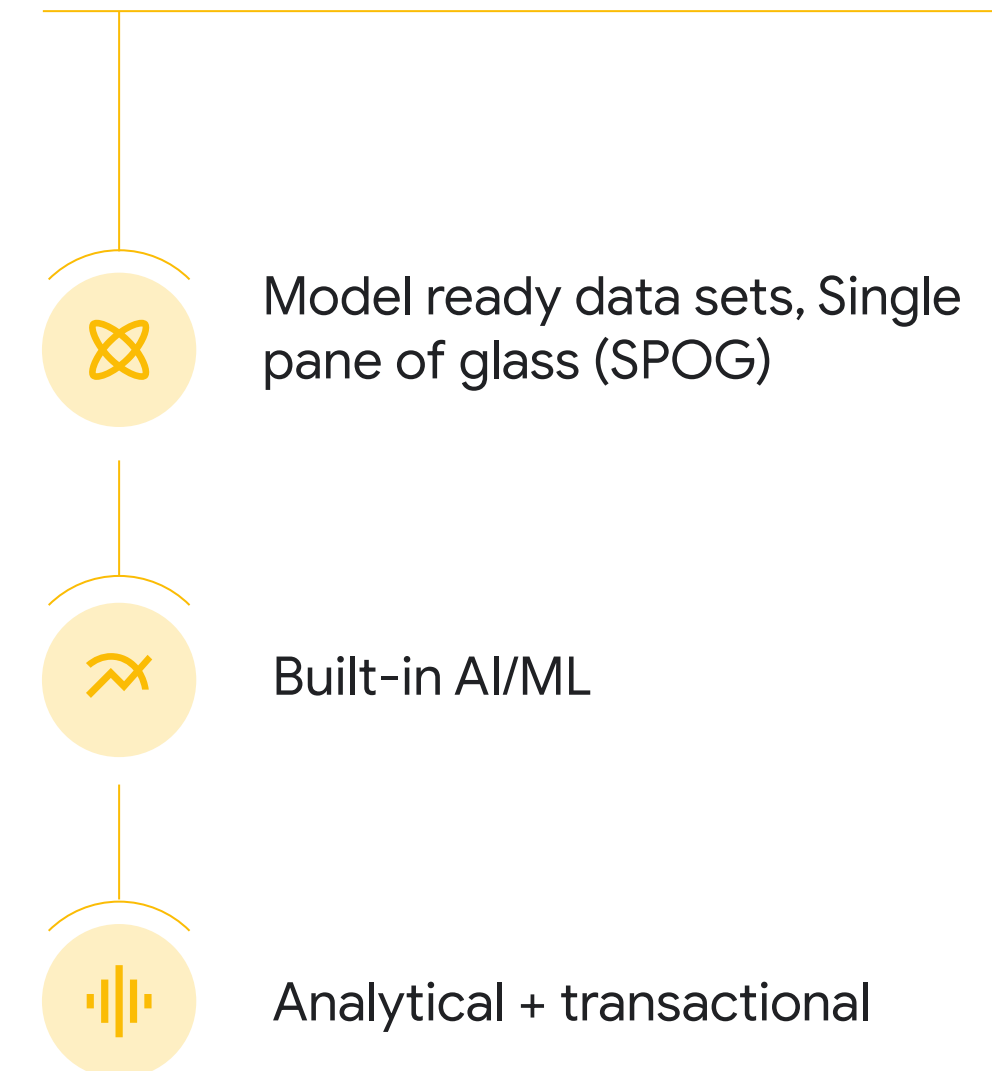
Platform for **everyone**



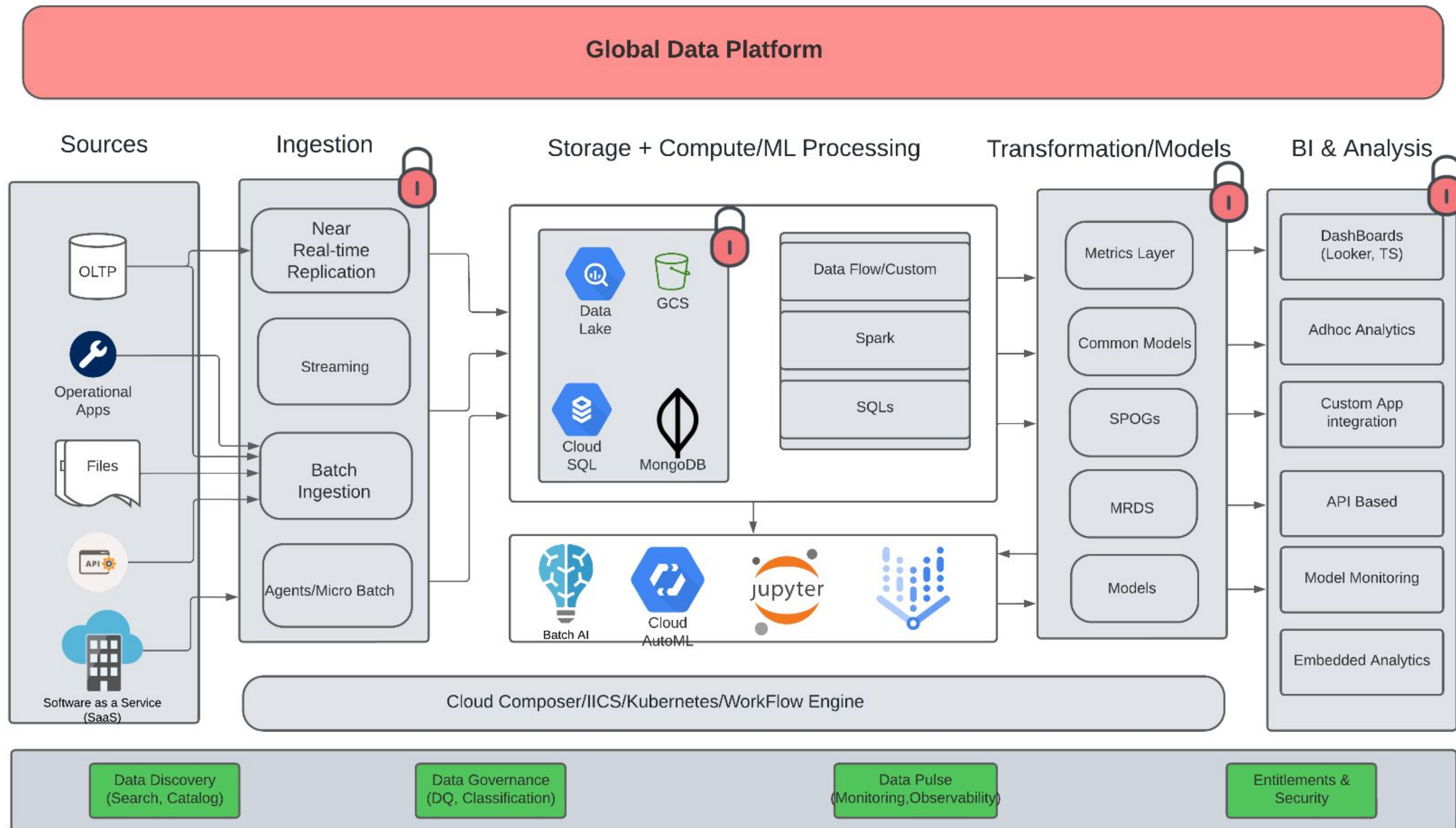
Limitless **data**



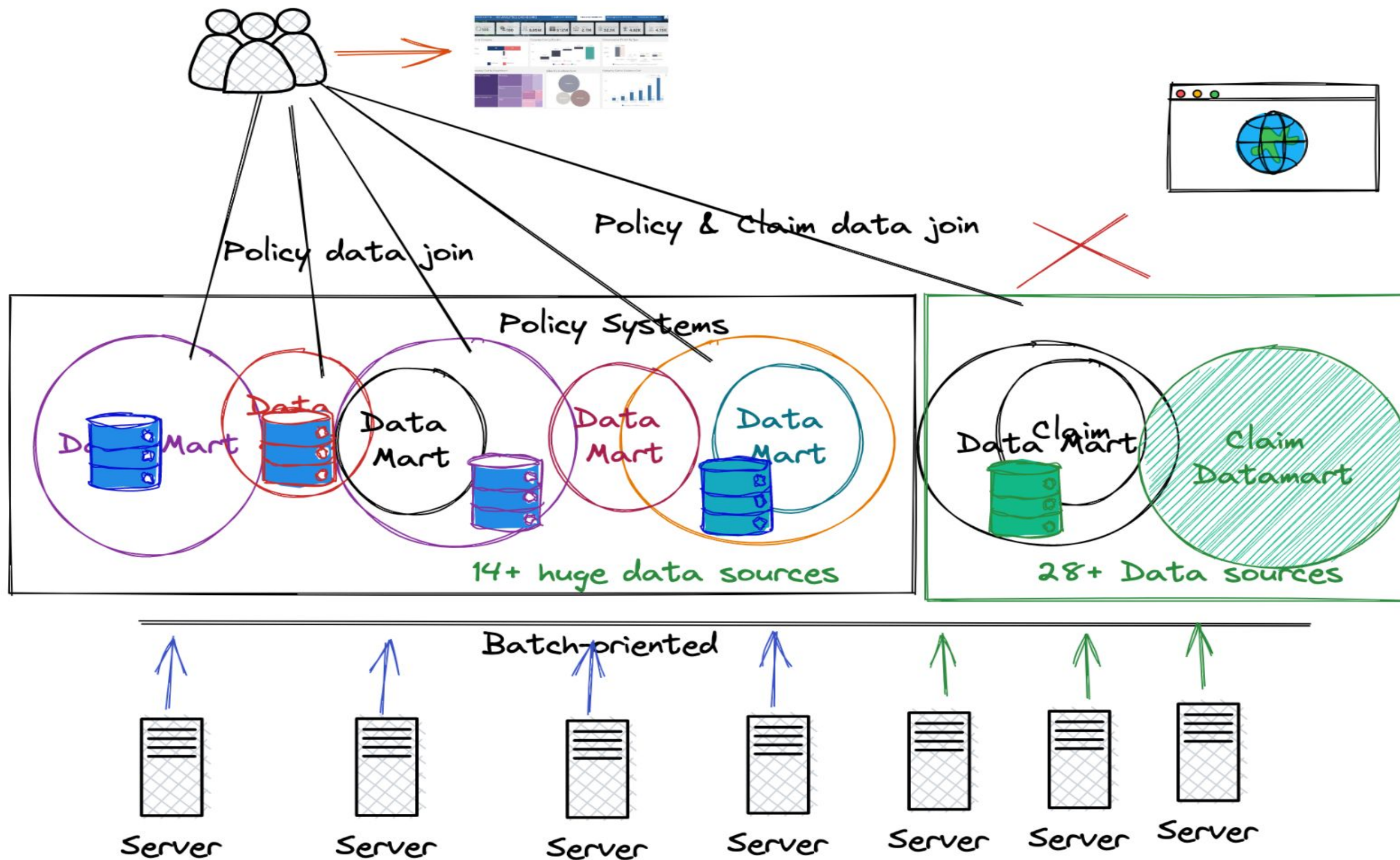
All **analytics & AI / ML**



Next-Gen platform for everyone



Legacy Architecture

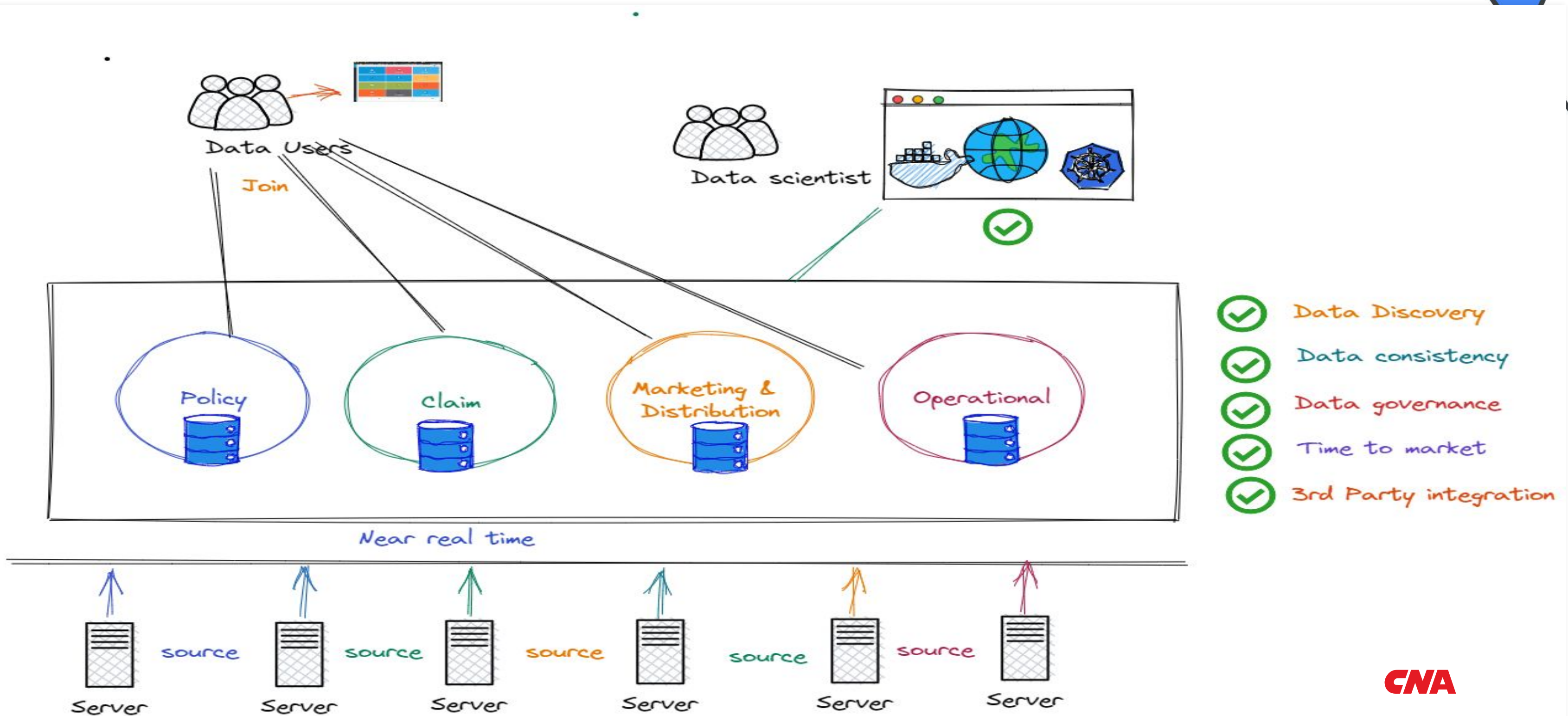


3rd party / SaaS
on demand integrations

Major Challenges

Data Discovery?
Data Governance?
data duplication
/inconsistencies
time to market?
AI/ML?

Next-Gen data mesh architecture



Data ecosystem to help solve cultural boundaries

All Types Of Data

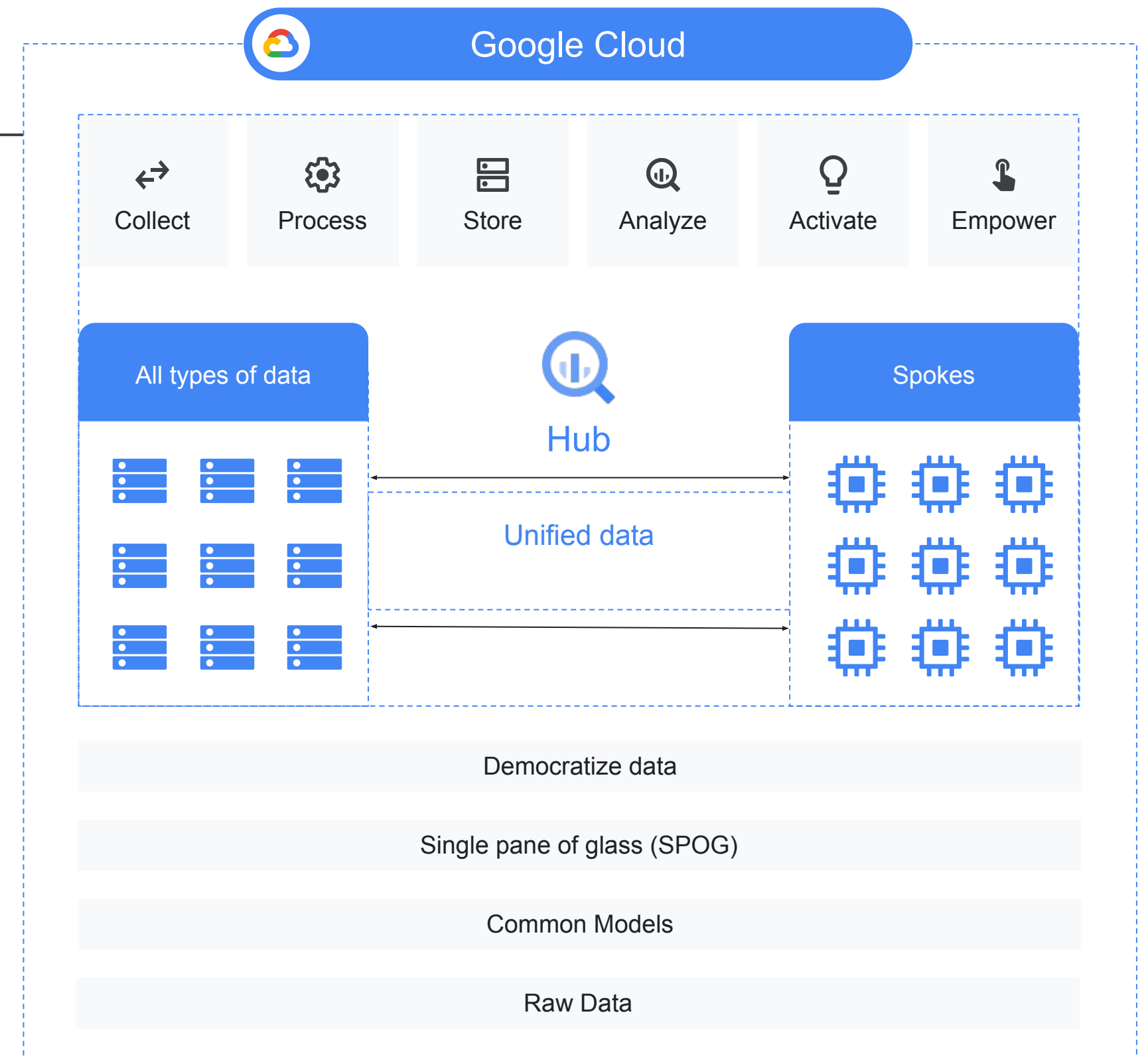
- Bigquery as Data Lake for structure data sets
- Data ingestion happen with various stream patterns
- Model ready data sets to serve AI/ML use cases and Vertex AI Platform.

Unified structure

- Simplify data structure / content with context
- Ability to keep hierarchy with Bigquery arrays

Hub & spoke

- Empower business domains to consume cloud data from hub
- Spokes made of business domains that own their own data
- Provision for spoke data sets to use by other spoke
- Spokes own workloads to manage and scale.



All analytical & AI/ML for data insights

BigQuery, Bigtable

- Raw data and transformed data
- Common data assets
- Build model ready data sets
- Offline store & online store for model integration

Feature Store

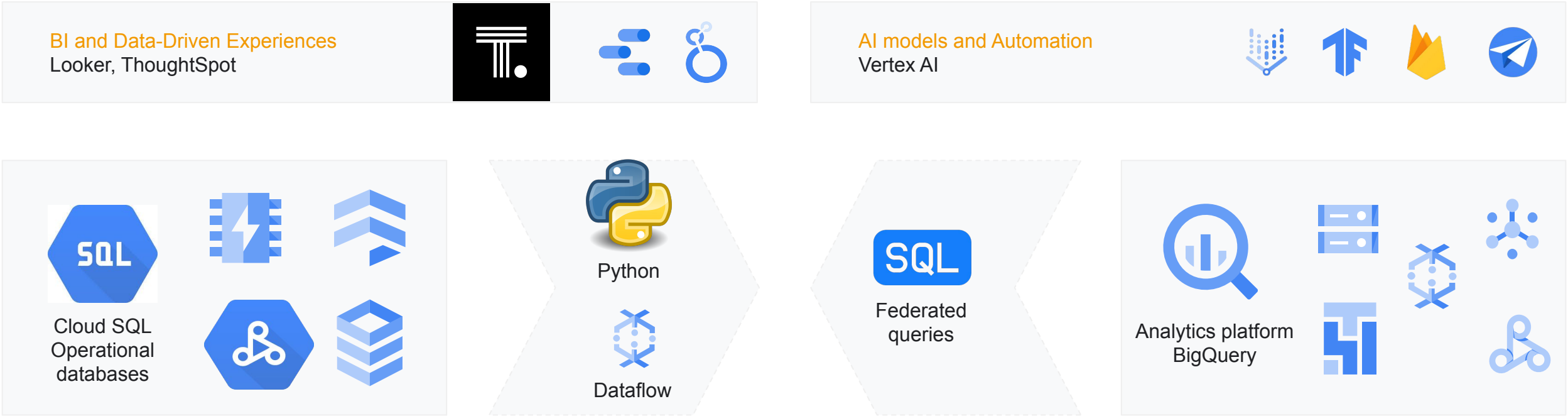
- Common data assets and model ready data set have direct integration into Vertex AI platform.
- Data scientist build models, use Vertex AI to deploy batch and real-time model.

Direct integration with Vertex AI Workbench

Data scientists use Dataflow, Dataproc and Tensorflow with data from BigQuery with a simple, seamless integration

Complete MLOps environment

Vertex AI provides an environment for continuous training and deployment workflows with kube flow ML Pipelines



Business value

- Meaningful data context, easy to discover
- Utilized data with unified data products
- Near real time data processing to enable data insights, near real time prediction
- Single source of truth for data consumers
- Cloud implementation offers seamless integration with 3rd party data providers



What we've achieved so far

Platform was built within 9 months

Delivered business value within year.

We have moved majority of our raw data into BQ.

Enabled hub & spoke model, Onboarded a major business domain within 9 months from dayone

Building AI/ML models using Google Vertex AI



Next steps

AI/ML Predictions

Self-serve one click platform provisioning

Expand hub & spoke model

Continuous cloud enablement by integrating with Marketplace cloud providers

Thank you

Google Cloud

Next '22

