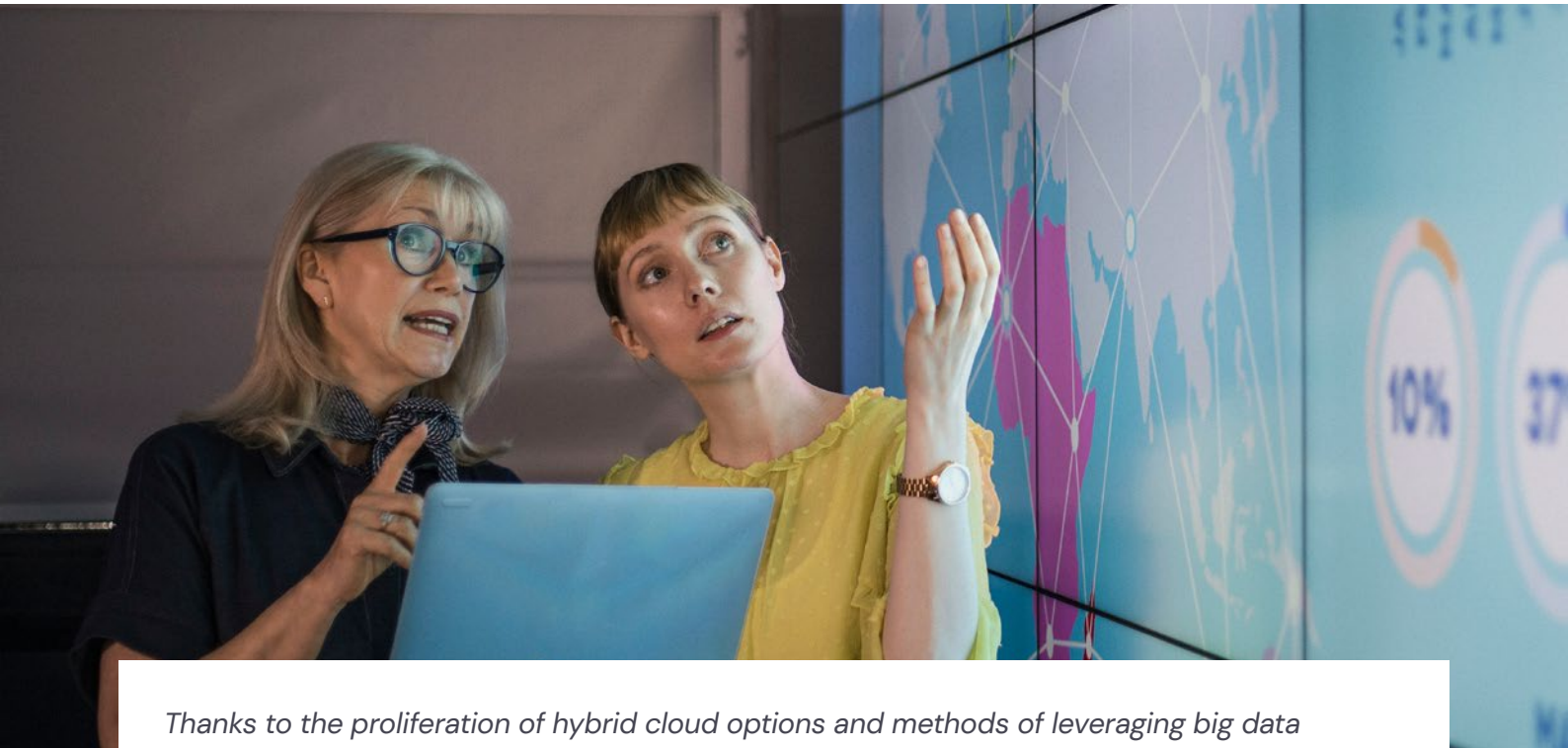# Transforming Data into Insights on Google Cloud

Thanks to the proliferation of hybrid cloud options and methods of leveraging big data adoption, the potential to unlock data value is expanding and becoming more promising. From storage to story — realizing the data tension between the data lake and harnessing its potential toward consumption-ready data products — ensuring trust, ownership, and traceability is imperative to leveraging the best of data lakes to enable just-in-time, empowered consumption/utilization across the data value chain and lifecycle.

At Kin + Carta, we deliver solutions that enable the creation of data products and intelligent experiences. As a certified Google Cloud Partner we utilize Google's suite of **Data Analytics** and **AI/ML** solutions to build systems that realize the highest potential of data lifecycle value across the greatest range of stakeholders.

We recently held a broad industry roundtable with IT executives examining the status of data-driven transformation. Four key themes included the need to:

– **Reduce Barriers for Adoption**
– **Enable Data-as-a-Product**
– **Build with Governance**
– **Sustain Momentum**

Following these key themes are core opportunities to leverage Google Cloud to accelerate data transformation into insights.

Reducing barriers for adoption

From engineering to data science, bringing familiar toolchains and practices closer to the center of data gravity for users may yield greater willingness to experiment, unlock, and realize value from data in the cloud.

Foundational to Google Cloud's Data Cloud is [BigQuery](#), the serverless, fully-managed data warehouse with low-latency petabyte-scale querying power for data analysis. With its architecture decoupling storage and compute, BigQuery is designed to serve as a landing zone to quickly stage data through integration tools like [Data Fusion](#), a low/no-code ETL solution with a vast range of transformation plug-ins to simplify building data pipelines. The BigQuery [Data Transfer Service](#) brings SaaS and enterprise data stores into BigQuery quickly and reliably without writing a single line of code. Continuous, near real-time replication of CDC-based workloads is managed with [Datastream](#), to quickly land data into BigQuery for downstream analysis and reporting.

For baseline model development, [BigQuery ML](#) allows data analysts to leverage native SQL skills to generate ML and evaluate models directly from BigQuery datasets while reducing data movement and security issues by interacting in a single interface.

Accelerating data science time-to-value is supported by Google Cloud [AutoML](#) (integrated with the Vertex AI platform), as well as through Kin + Carta's [Octain](#) solution. Both let you rapidly build models with relatively small data sets to prove use-case impact and begin further iteration against larger datasets to support deployment in production. Google Cloud's integrated AI/ML products are built atop Google's [world-class AI](#) capabilities to help you build custom models within minutes. They offer [pretrained models](#) to solve common-use cases while letting you develop custom models as needed.

For data activation and reporting, [Looker](#)'s self-service BI platform serves the spectrum of data analysts, data scientists, and business users. Its powerful modeling language removes data movement concerns to ensure quality controls and simplify adoption. Along with a [marketplace of prebuilt data connectors](#) and plugins, Looker is designed to readily spin up visualizations and dashboards for common use-cases and is accessible from any environment.
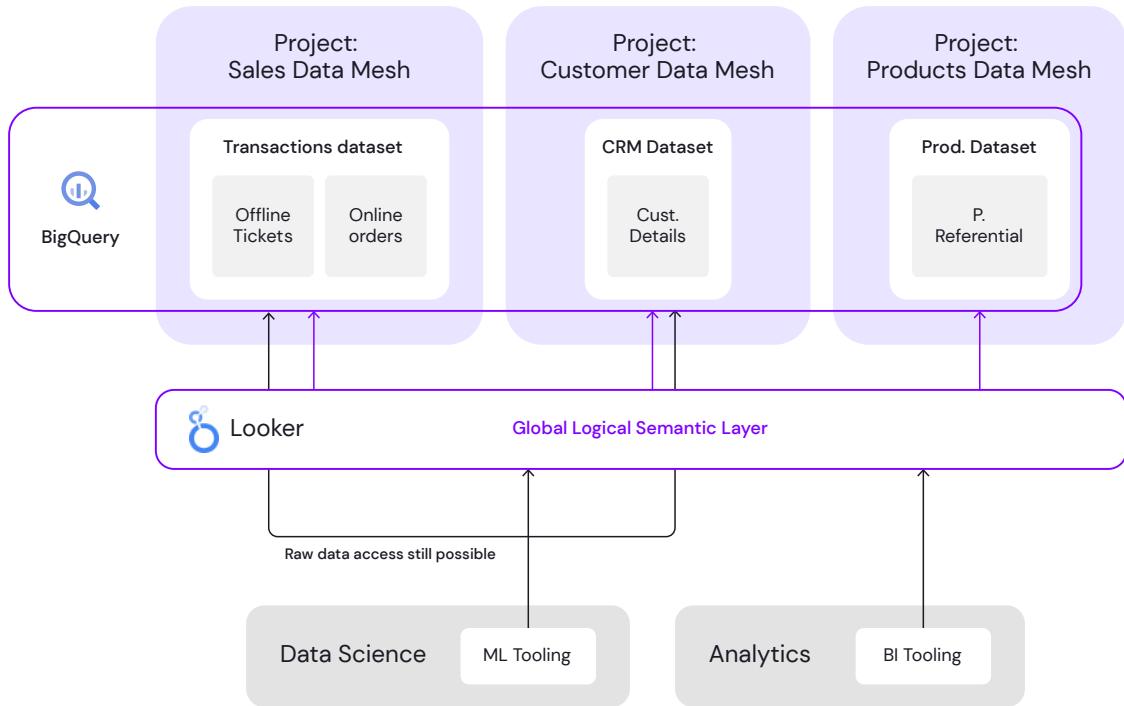
Enable Data-as-a-Product

Data-led transformation means rethinking the stakeholder relationship to the data lifecycle. The [Data Mesh architecture pattern](#) is a response to new challenges of scale, quality, access, and activation with four core principles: domain-driven ownership of data, data-as-a-product, self-service data platform, and federated computational governance.

When thinking about data-as-a-product, it's important to enable quality controls to the source of data-users to engender responsible use of data platforms and increase the flexibility and velocity to generate new use-cases and value at scale.

Looker's modern analytics capabilities support data mesh principles via domain-centric, data activation and federated access patterns. It uses a modeling layer that abstracts underlying data connections from querying and business logic, along with robust governance controls. It's designed to empower self-service analysis and experimentation across various data sources with guardrails to ensure a single version of the truth across data platforms. Looker further supports data products through its API and extension framework, embedding visualizations directly into

applications, triggering business workflows, and exposing insights where consumer-endpoints can be best targeted.

To further enable data-as-a-product, you can tap into Analytics Hub to securely exchange data assets across your organization and meet quality, security and cost controls. This service lets you curate a library of internal and external assets to deliver data at scale to exactly the right audiences. Because it is built atop the power of BigQuery with storage decoupled from compute, publishers can share datasets with multiple subscribers without duplication, effectively federating access for domain owners to augment analytics and ML initiatives with trusted datasets.
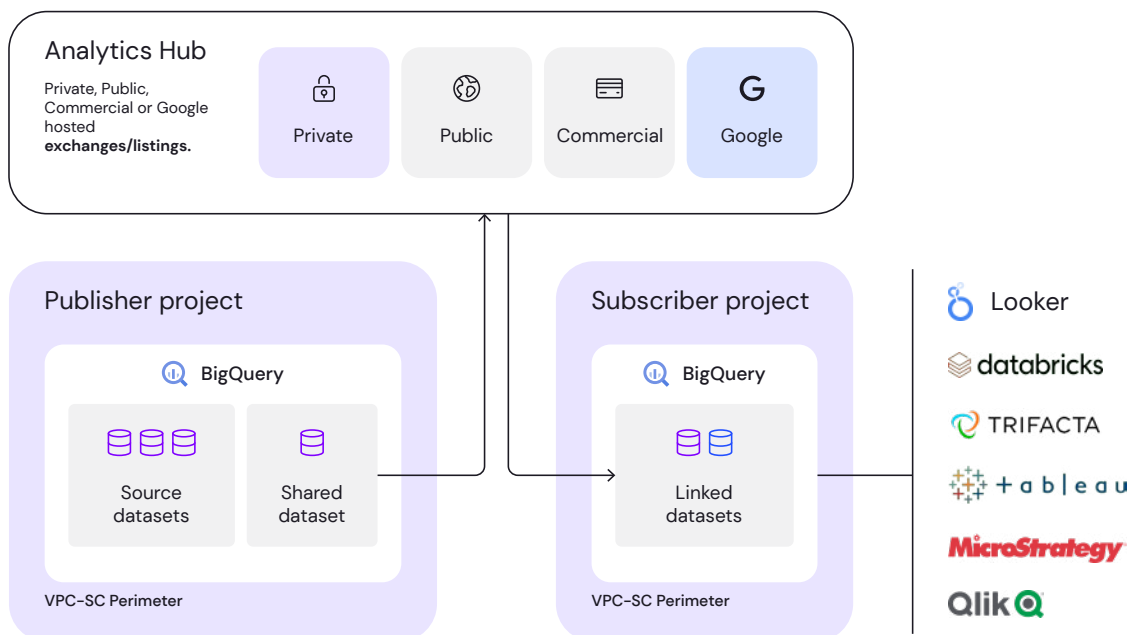
Build with Governance

Google Cloud services such as Data Fusion, BigQuery, Analytics Hub and Dataplex all readily support best-practices for designing a data mesh architecture. Foundational concerns including governance, IAM, security and auditing, are supported at every level of implementation with all data handling in Google Cloud encrypted by default, both in [transit](#) and at [rest](#).

[Dataplex](#) suite helps organizations discover, manage, monitor, and govern data across disparate data stores with consistent controls, providing access to trusted data and powering analytics at scale. Dataplex lets you automatically capture data lineage to understand, trace dependencies, and effectively troubleshoot data issues.

Dataplex is also integrated with [Data Catalog](#) for technical and business metadata management and leverages [Data Loss Prevention](#) to automatically detect, obfuscate and de-identify sensitive information in your data, minimizing the risk of compliance violations.

Control over data access and analysis in [BigQuery](#) is possible with features such as [authorized views](#) and [column-level](#) and [row-level](#) access policies, allowing datasets to be managed securely with clear data provenance and versioning.

Robust auditing and metadata are standard in [Data Fusion](#)'s integrations to support governance and compliance alongside the partner ecosystem. This extends data lineage controls across on-prem and hybrid cloud environments, lowering TCO, promoting self-service and standardization, and reducing repetitive work.

Sustain momentum

Data-led transformation requires the tools, people and processes to manage change at the rate it can be metabolized. This requires strong cloud foundations to support workload migrations, enable automation and increase the scale and velocity of change.

Kin + Carta's [Data Launchpad](#) leverages the Google Cloud [Cloud Adoption Framework](#) to guide enterprise cloud readiness and evaluate opportunities for database, analytics and data application workload migration. We leverage the [Cloud Foundation Toolkit](#) to provision infrastructure as code (IaC) for repeatable pipeline deployments, managing velocity of data-driven applications at scale. This approach ensures the right workloads and migration patterns to lower risk of configuration changes, and it increases the rate of return when ultimately implementing a [data migration program](#) on Google Cloud.

From foundational data integrity and quality comes opportunities to leverage data science practices on Google's unified ML Platform, [Vertex AI](#). You can begin with cloud-native notebooks on [Vertex AI Workbench](#) to manage data modeling and experimentation. As capabilities grow, we can help you standardize workflows with reusable components on [Vertex AI Pipelines](#). This provides the orchestration for [MLOps](#) on Google Cloud for continuous model integration, deployment and retraining. The ability to centrally manage artifacts, uphold data-model parity with an integrated Feature Store, and target endpoints to serve real-time online predictions all through a unified API in a single pane of glass is remarkable. The automation and efficiency gains through Vertex AI can augment transparent, explainable and responsible AI practices to uphold far-reaching enterprise mandates.

Lift & Rehome

- Conservative approach
- Fast migration from existing services such as TD Vantage, Databricks on to GCP
- No modernization or improving existing solutions apart from running them over GCP as a tactical intermediate decision

Lift & Replatform

- Optimal phased approach. loww disruption, low risk and high impact
- Migrate data into BQ from legacy EDW
- Migrate data into Dataproc from on-premise Hadoop cluster
- Optimize queries and data pipe-lines for performance
- Up to 57% lower TCO than on-prem

Modernize

- All in on cloud-native, clean break from the past
- Built natively on GCP
- Can be slower as it requires rewriting jobs
- Greatest development velocity and agility
- 60-88% lower TCO than on-prem, plus value from Google AI on unstructured data

---

Customized

Automation & Efficiency

AI & Machine Learning
- Real-time automation for business processes
- Intelligent data products power decisions and CX at scale
- Predictive/Prescriptive and AutoML continuous learning requires fewer human decisions

Data Decisioning
- Self-serve data products for an Insights-driven organisation, better and faster decisions
- Customer 360/Single view of the customer
- Insights democratization across all employees

Discovery, Analysis & Insights
- One-stop-shop for data discoverability, expertise & collaboration
- Monetisation of data
- Federated Data / API access by domain
- Single platform for analytics and operational use-cases

Management, Governance & Architecture
- Centralized data inventory
- Operationalized policies and workflow
- Controlled access, standards, quality, lineage & traceability for all data
- Semantic model connects Data Silos

Data Collection & Storage
- Move priority data workloads. Unlock elastic scale & cloud native services
- Enable federated & domain-driven access to Data-as-a-Product
- Low-latency data integration

Commoditized

The combined abilities of Google's Data Cloud create an interoperable and secure framework of open source standards and extensible APIs that can harness data in hybrid-cloud environments and accelerate the momentum to turn data into insights at varying stages of data-led transformation initiatives.

Google Cloud users benefit from an integrated Data → Value journey

Any user

Data engineers	ML engineers	Data scientists	Developers	Data analysts	Business users	Consumers
Clean, useful data	Integrated intelligence	Models that work	Intelligent apps	Query and analyze	Insights Everywhere	Value

Any Data

Corporate and 3rd Systems	E-commerce	SaaS Applications	IoT Devices	Social Media	Geospatial

Google's Data Cloud

Any workload

Data Engineering	Data Analysis	Model Development	ML Engineering	Insights activation
Dataproc	BigQuery	Vertex AI	Vertex AI	Vertex AI
Dataflow				Looker / App Engine

Cloud Composer   Vertex AI Pipelines

Any innovation

Decision Making	App Development	SaaS Applications	Exception Management	Operational Intelligence	Data Monetization

KIN + CARTA

Ready to make data work harder for your organization?

Kin + Carta's solutions leveraging Google Cloud across the data value lifecycle can smooth the cloud adoption curve and accelerate time to value by building data products and intelligent experiences at scale.

Learn how we can help you harness the power of Google Cloud.

(→) **Get in touch with our experts**

Find us on:

in 𝕏 ⊙ f