

Balancing Speed and Risk:

# Next-Gen Data Governance Meets Real-Time Data Delivery



# The imperative for speed versus the threat of risk.

In the digital economy, real-time agility is essential. The faster your organization can respond to events as they evolve, the better it can compete. That means arming more people and systems with up-to-date, real-time information and analysis to optimize every business decision.

But there's a conflict. As you accelerate the delivery of data and make it more immediately available to more users, the security risks increase. And no data leader wants to end up in the news as a result of a breach. As CIO or CDO, you live with this tension every day. How can you balance the imperative for speed with the threat of risk?

## THE TWO-SIDED EQUATION

### When naming their top digital ambitions:

**35%** are speeding up automation to cut costs

**31%** are modernizing with new capabilities

**29%** of executives cite speed and efficiency

### With more ambitious digital initiatives, businesses are investing in security:

**55%** are increasing security budget

**51%** are increasing security headcount

**50%** report that security and privacy will be baked into every decision or plan

SOURCE: PwC, 2021 Global Digital Trust Insights<sup>1</sup>

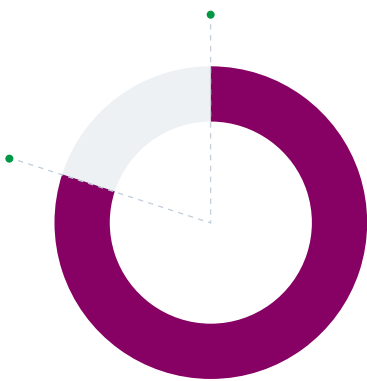


# A breakthrough in data delivery.

For decades, the speed of data delivery was a roadblock in accelerating the analytics cycle. But in recent years, data delivery has been revolutionized by new data architecture options (including, of course, the cloud); automation technology; and the novel approach of DataOps.

Today's data integration solutions have the power to collect and interpret multiple data sets; to eliminate information silos; to democratize data access; to replicate data in real time without impacting source systems; and to provide a consistent view of governed, real-time data to every user across the business. At the same time, the industry trend of consolidating data management and analytics functions into streamlined, end-to-end platforms is making it possible for businesses to advance the speed and the accuracy of data delivery. That, in turn, is advancing the speed and accuracy of insights – especially when we layer in assists from Artificial Intelligence and machine learning, which augment data literacy so that users at every level can interact with data.

Additionally, the powerful discipline of DataOps – modeled on its predecessor, DevOps – brings together people, processes, and technologies to optimize data pipelines for meeting today's considerable demands. Through a combination of agile development methodology, rapid responses to user feedback, and continuous data integration, DataOps makes the data supply chain faster, more reliable, and more flexible. As a result, modern data and analytics initiatives become truly scalable – and businesses can take even greater advantage of the data revolution to pull ahead.



## Revisiting the 80/20 rule.

For years, organizations have struggled with an unfortunate ratio governing the pursuit of data-driven insights: **80% data preparation time to 20% data analytics time.** The new data delivery technology begins to reverse that equation by dramatically reducing the labor in data preparation.

# Enabling DataOps with a four-pronged approach.

The principles of iterative development, sprints, and quick pivots can apply to analytics just as they do to app development – but not if the data isn't available. After all, how useful is failing fast if analysts have to wait weeks for their next data set?

In order to enable DataOps, it's critical to know when data will be required and what analytics cycle it will generate. The data pipeline should be handled like any operational process, with supply chain, resource, and distribution planning. And four key items should be aligned:

## 1. BUSINESS STRATEGY

At the heart of any analytics problem is a business need. While this sounds obvious, it's not always easy to get the C-suite on board. Convinced by popular culture that modern analytics are incomprehensibly complex, executives often delegate strategy to IT. Instead, they should focus on connecting strategy to business events that can be altered through data.

## 2. DATA AVAILABILITY

If we know what strategies are driving analytics, we can shift our attention to the needed data. Customer attrition? Grab web logs, customer demographics, and transactions. Property risk? Try weather, geospatial, and loss data.

## 3. ANALYTICS EMPOWERMENT

In a data-democratized world, we should expect every business user to be able to read, work with, analyze, and argue with data. That requires improving data literacy skills at every position. Everyone should have the encouragement and tools to experiment. And every data initiative should be evaluated for potential business impact, including risk.

## 4. ANALYTICS OPERATIONALIZATION

Now it's time to deploy the app in a production context – and measure performance and impact. Naturally, the impact must come full circle to the business strategy. This closed-loop process will provide a critical link back to the non-technical community and allow for clear measurement.

### Let's be real: This is tough.

There are significant cultural, procedural, political, and technical challenges in successfully implementing DataOps. But **having a clear vision with language that's business-focused** rather than technology-focused will accelerate the journey.

# Consumers versus providers: What's behind the tension?

For years, data consumers and providers have been at odds. Consumers want data for their projects immediately. Providers are working their hardest to provision data responsibly, to everyone, without putting the business at risk.

“

We all know that the faster we put data to use, the better we can compete. But I have to wait weeks and jump through hoops to get access – and sometimes, it's not even the data I need.”



**DATA ANALYST/SCIENTIST/BUSINESS LEAD**

“

We're always just one hacker away from a breach. So when a user works around the process, I take it personally. When they do follow process, they want their data immediately – but so does everyone else.”



**CDO/DATA ENGINEER**

There's a logical reason for this tension. Data governance was developed to create and certify a common understanding of enterprise truth. But modern analytics, especially in predictive modeling and AI, is about the discovery of insight and probability.

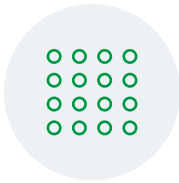
# Data in the production line.

How can you make peace between these two forces? Let's start with a metaphor: data as material within a factory supply chain. Every manufacturing company has raw materials, semi-finished goods, and finished goods. Data is like that, too:



## RAW DATA

Raw data exists within applications to record information, including transactions (customer orders) and other records (employee expenses).



## SEMI-FINISHED DATA

Semi-finished data is created for a discrete purpose – a spreadsheet with a sales forecast, for example, or a BI application with service incidents.



## FINISHED DATA

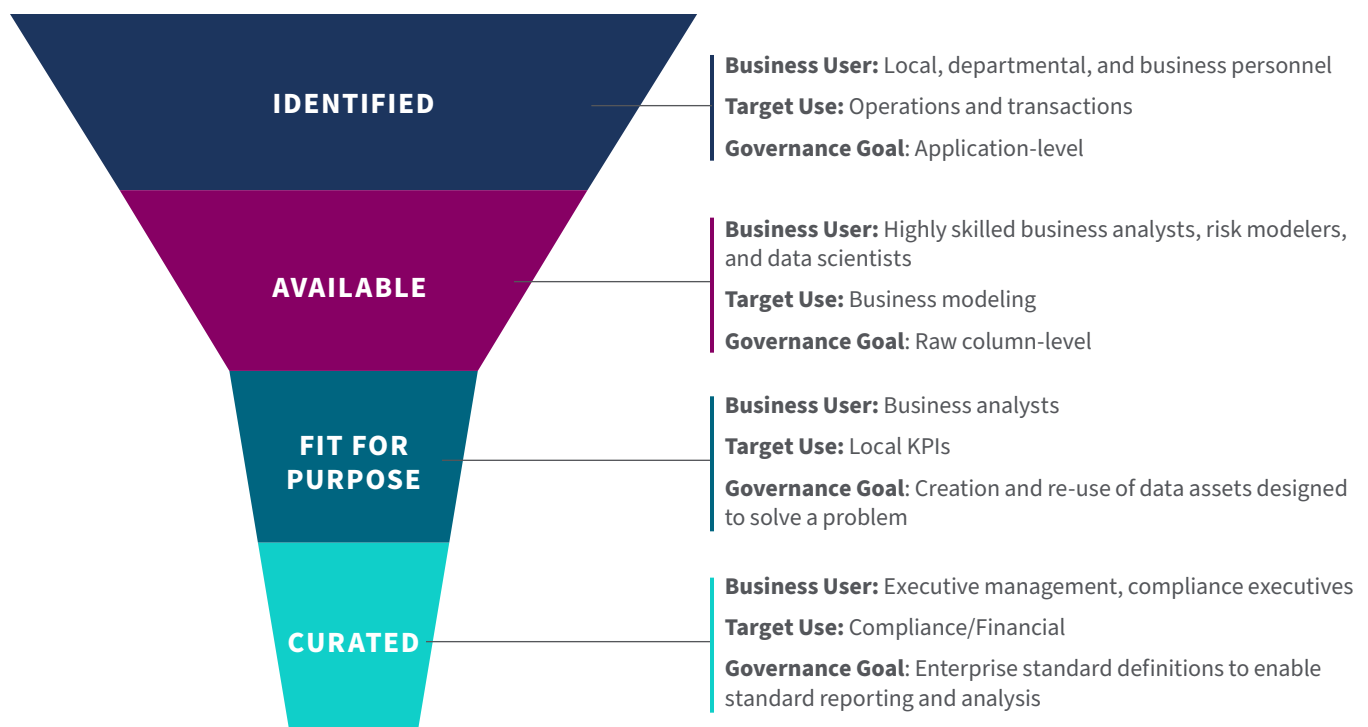
Finished data is analytics gold. Financial close metrics, regulatory submissions, and critical KPIs require one formal definition with high data quality.

Traditionally, data governance has focused on the management of that third category, finished goods. But advanced data analytics use raw and semi-finished goods. And that's where much of the tension between producers and consumers comes from.

# The solution: a tiered system of data governance.

There's another way to govern data – with a tiered system that addresses different user needs with different types of data, applying increasing scrutiny and quality standards as the data works its way through the system.

## ENTERPRISE DATA GOVERNANCE FUNNEL



**Important:** Users of data at each step of the curation process should have an understanding of the quality and reliability of the data. Want to put a KPI on the CEO's desk? It has to be curated, with no argument about the data. Want to create a predictive model of how weather will affect sales? Available data may be good enough. Once the user figures out which sources they're using to make their model useful, they can insist on a higher degree of curation.

This system helps the enterprise governance function focus on a breadth of understanding across the enterprise, including enabling restrictions to sensitive data, as well as a depth of understanding for a smaller number of critical data assets.

# Enter the data catalog.

How can you make different kinds of data available to different kinds of users quickly, without compromising risk? Establish a clear, use case-based catalog with the ability to onboard, profile, describe, secure, and even prepare and obfuscate data quickly in anticipation of analytics sprints.

Data governance in this context is not about torturing data into truth but rather clearly defining, classifying, and provisioning data assets – at speed and scale – to people who are authorized to see it.

## THE CORE CAPABILITIES OF AN EFFECTIVE DATA CATALOG



**Onboard data:** A data catalog should automatically profile and document the exact content, structure, and quality of enterprise data as it enters the marketplace from any and all sources. Metadata should be used to capture, enforce, and monitor data policies and usage from the moment data enters the marketplace.



**Enrich the catalog:** The catalog should document every aspect of the data and management process. As users search and explore, technical, business, and operational metadata should make each data element transparent, trustworthy, and actionable.



**Prepare the data:** The data catalog should make data business-ready by preparing and enhancing it with data standardization, cleansing, transformation, and protection measures.



**Shop and publish:** The catalog should make data available for easy, secure consumption by all types of enterprise users. It should support one-time exports and the recurring, automated publishing of bespoke data sets to downstream consumers, including data science or analytics platforms, applications, and cloud data stores. And it should obfuscate sensitive fields automatically so data security is enforced.



# The ever-increasing importance of the cloud.

The world is changing in ways that make two enterprise imperatives even more urgent: 1) the need to establish real-time agility based on the very latest data, and 2) the demand to shift from traditional, on-premises solutions to Software as a Service (SaaS) and Infrastructure as a Service (IaaS). In particular, the hybrid model of work that is so quickly becoming the norm has made the cloud more relevant than ever before. And we can see that in the latest metrics.

## Why cloud? Innovation and security.

More than a third (35%) of executives strongly agree that moving to the cloud is foundational for the next generation of business solutions for their organization. And 36% strongly agree that new solutions exist to secure cloud infrastructures better than ever before.

PWC, 2021 Global Digital Trust Insights<sup>7</sup>

In 2020, the move to remote work kicked public cloud growth into overdrive. Amazon Web Services (AWS) grew by 33% in the first quarter,<sup>3</sup> Google Cloud experienced a 52% year over year growth,<sup>4</sup> and Microsoft Azure grew a massive 59%.<sup>5</sup> Those numbers were fueled in part by the strong adoption of cloud data warehouses. And cloud data warehousing is here to stay. A recent IDG survey<sup>6</sup> found that:



Of course, the benefits of the cloud go well beyond data storage and elastic compute functions. But to fully harness these IaaS capabilities requires a new approach to data integration that puts a premium on real-time data delivery and the automation of analytics-ready data.

# How Qlik® can help.

## Rapid, secure data delivery for next-generation analytics.

With the open, end-to-end Qlik Data Integration Platform, you can vastly accelerate the availability of real-time, analytics-ready data to the cloud of your choice by automating data streaming, refinement, cataloging, and publishing. This infrastructure sets the stage for a DataOps approach. And it gives you a head start on the journey to Active Intelligence – a state of continuous intelligence that uses real-time data pipelines to trigger immediate action.

Built on a platform of hardened data security and featuring governance capabilities, Qlik Data Integration builds an enterprise-scale repository of all the data your business has available for analytics. It provides consumers with a single, go-to catalog where they can find, understand, and gain insights from any underlying enterprise data source. The solution's data preparation and metadata tools streamline the transformation of raw data into analytics-ready assets, while the Smart Data Catalog and graphical user interface help people easily discover and select whatever data they need.

**Learn more about the Qlik Data Integration Platform. →**

1 PwC, 2021 Global Digital Trust Insights, <https://www.pwc.com/us/en/services/consulting/cybersecurity-privacy-forensics/library/global-digital-trust-insights.html>

2 "Security vs. Innovation: IT's trickiest balancing act," CIO Magazine, February 2020, <https://www.cio.com/article/3521009/security-vs-innovation-its-trickiest-balancing-act.html>

3 "Amazon reports \$75.5 billion in Q1 2020 revenue: AWS up 33%, subscriptions up 28%, and 'other' up 44%," VentureBeat, April 2020, <https://venturebeat.com/2020/04/30/amazon-earnings-q1-2020/>

4 "Google Outpaces Microsoft, Amazon in Cloud-Revenue Growth at 52%," Cloud Wars, April 2020, <https://cloudwars.co/google-cloud/google-cloud-q1-revenue-outpaces-microsoft-amazon/>

5 Microsoft, April 2020, <https://www.microsoft.com/en-us/Investor/earnings/FY-2020-Q3/press-release-webcast>

6 "Cloud Is the New Center of Gravity for Data Warehousing," datanami, January 2021, <https://www.datanami.com/2021/01/08/cloud-is-the-new-center-of-gravity-for-data-warehousing/>

7 PwC, 2021 Global Digital Trust Insights, <https://www.pwc.com/us/en/services/consulting/cybersecurity-privacy-forensics/library/global-digital-trust-insights.html>

## ABOUT QLIK

Qlik's vision is a data-literate world, where everyone can use data and analytics to improve decision-making and solve their most challenging problems. Qlik provides an end-to-end, real-time data integration and analytics cloud platform to close the gaps between data, insights, and action. By transforming data into Active Intelligence, businesses can drive better decisions, improve revenue and profitability, and optimize customer relationships. Qlik does business in more than 100 countries and serves over 50,000 customers around the world.

[qlik.com](https://qlik.com)

