IRI

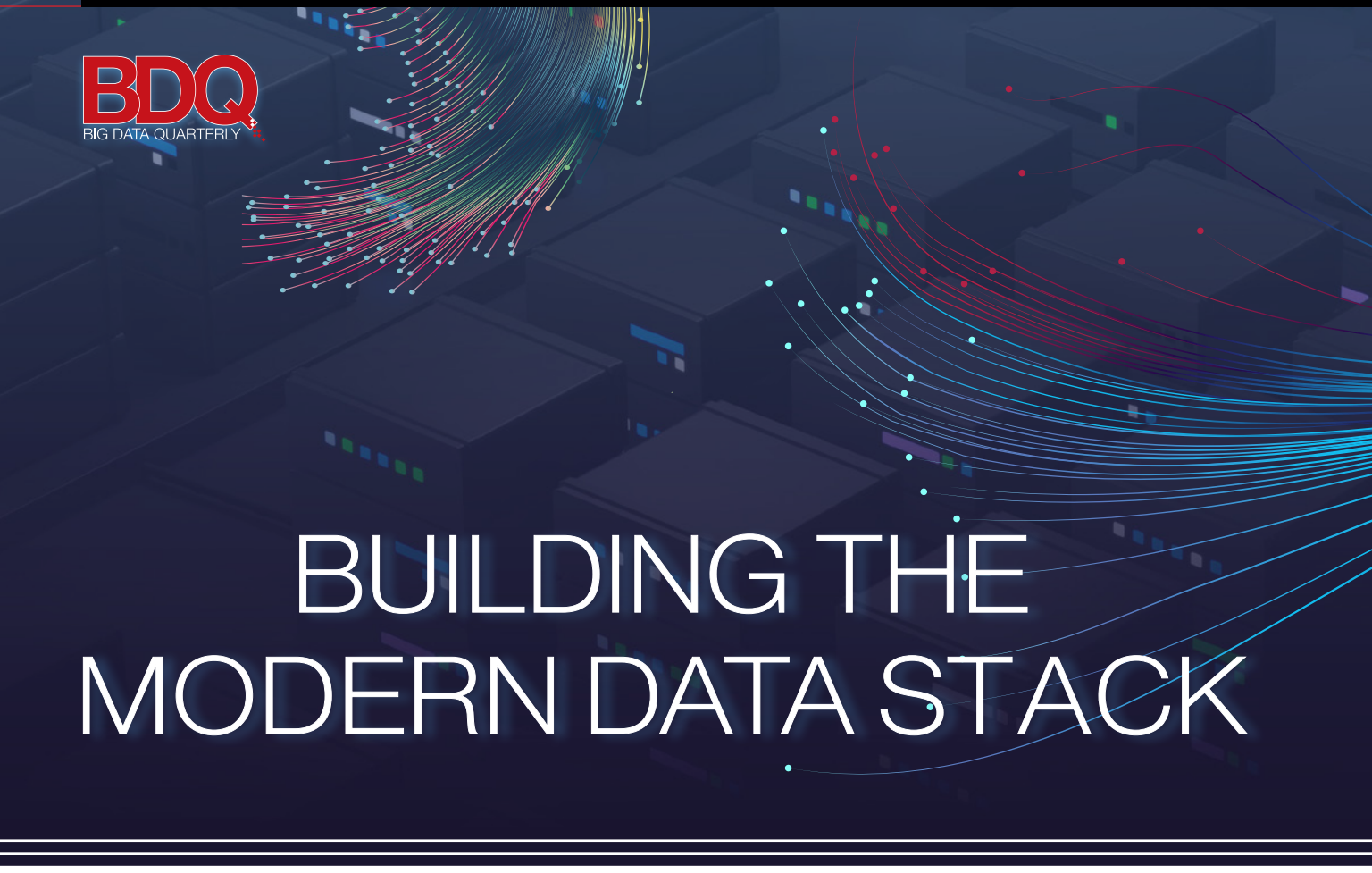
ADDRESSING 2025 DATA ARCHITECTURE AND GOVERNANCE GOALS

BDQ

BIG DATA QUARTERLY

THE MODERN DATA STACK IN 2025: New Tools, Platforms, and Strategies

Best Practices Series



BUILDING THE MODERN DATA STACK

THE MODERN DATA STACK IS INCREASINGLY BUILT UPON CLOUD services, boosted by AI and machine learning, flowing via DataOps-managed pipelines, and focused on delivering real-time data and insights to users at all levels of their organizations. This evolution that began with distributed computing many years ago continues to adapt to changing business needs, now dominated by the digital era.

The adoption of cloud data warehouses and data lakehouses will continue to grow alongside real-time data platforms and data fabric deployments, according to a survey of data managers by DBTA. Areas of focus include data warehousing (45%), data lakes (41%), data lakehouses (36%), real-time unified platforms (32%), data fabric (32%), and data mesh (27%) ("Market Study: 2024 Modern Data Management & Data Fabric," Unisphere Research).

The benefits of a modern data stack include greater agility and flexibility, along with faster insights, improved data governance and observability, and cost savings. The move to modern data stacks is essential at a time when businesses require hyper-personalization to reach customers, the ability to predict market shifts, and greater automation. Data ingestion, transformation, and orchestration tools will all increase in prominence to meet these requirements.

Today's business environments also need to support initiatives, from real-time and advanced analytics to self-service capabilities. As a result, they can meet the need for reliable analytics tools that seamlessly connect and accommodate the large variety of analytics requirements at organizations.

Today's enterprises are wrestling with data complexity, compliance laws, and mandates; demanding digital users; and new configurations, such as data residing and being processed within edge devices. These are challenges that were multiplying before the recent proliferation of AI.

Data-related issues stand in the way of effective AI development. Many business leaders—and even data managers—lack visibility into the type and scope of data assets they have within their enterprises. In addition, there is a wide, and often bewildering, array, of tools and platforms that serve data management requirements, exacerbating the growth of data silos.

Modern data stacks need to address all of these challenges with targeted investments in the right technologies and initiatives. Data can no longer take a backseat in enterprise technology priorities. A new generation of tools is enabling greater observability and, ultimately, governance that positions data where and when it is needed to support decision making and customer engagement.

Here are some key elements being ingrained into the modern data stack to deliver data-driven advantages:

AI and machine learning add intelligence to business operations and also support modern data stacks. Enterprises are moving forward with foundation large language models

Best Practices Series

The move to modern data stacks is essential at a time when businesses require hyper-personalization to reach customers, the ability to predict market shifts, and greater automation.

(LLMs), as well as their own LLMs, as the core data engines for analytics. These operate in conjunction with retrieval-augmented generation (RAG) environments to serve as the connective tissue between corporate databases and AI. RAG, built on natural language processing for query capabilities, is seen as an important component of data environments built on legacy investments. Investments in AI and machine learning are required to ensure the availability of high-performance hardware and storage capacity.

While modern data stacks are essential for AI, AI may also serve to ensure performance across the data stack along with security and provisioning of data resources.

Cloud computing underpins the modern data stack. Increasingly, data managers and their enterprises are taking a cloud-first approach to supporting and managing their data assets. This provides a far greater depth of capacity, keeping up with technology, and, ultimately, security to support data-driven initiatives such as AI and edge computing.

The scalability of cloud services is a compelling value proposition. Another emerging trend, serverless architecture, a byproduct of cloud computing offered by leading platform providers, frees data managers and developers from the low-level plumbing tasks associated with server provisioning and management.

Real-time technology is now a key goal of modern data stacks. Modern data stacks need to be designed to support the real-time flow of information from their sources, as well as refinement and delivery to end users. The need for real-time insights and awareness of business environments is a competitive necessity. The batch computing era, which may have held back critical decision-making data for hours or even days, is effectively behind us. Decision makers need to understand customer requests or issues when they happen—or even before they happen.

Intelligent data pipelines and intelligent storage define the architectures of modern data stacks. Intelligent data pipelines and storage are essential to the growth of the real-time enterprise. The push is on to harness the firehose of raw data flowing into organizations and convert it into insights for users at all levels. These capabilities support real-time data delivery, enabling enterprises to react to customer requests, as well as manage glitches or slowdowns in the flow of data to various points in the enterprise.

Data democratization is delivered—finally—through modern data stacks. One of the challenges—and goals—of data platforms for the past few decades has been providing the ability to access, through self-service, data any time it is needed, within proper security constraints. This provides analytical power that extends beyond the domains of data scientists and quants. The natural outgrowth of this capability is the rise of a data-driven culture, which can set an organization light-years ahead of competitors.

DataOps brings people and resources together within the modern data stack. DataOps is the glue that keeps data organized and flowing through the business. This is essential for real-time delivery, as well as maintaining the quality and timeliness of data. DataOps, in practice, opens up communication between data teams and IT operations, as well as developers. DataOps practices are especially important for enterprises that rely on multiple clouds and need to sync the flow and availability of data between clouds, as well as any on-prem environments.

Organizations are relying more than ever on the performance and capabilities of their data teams, environments, and assets. Developing and deploying modern data stacks are crucial steps toward achieving success as agile, data-driven enterprises with superior intelligence, enhanced operational efficiency, and improved decision making.

—Joe McKendrick

IRI Voracity
An Insatiable Appetite for Data

Addressing 2025 Data Architecture and Governance Goals

THOUGH ALWAYS EVOLVING, the data management trends that DBTA identifies in 2025 prioritize cloud-based data, real-time analytics, and enhanced data governance. This article explains what those trends mean and how you can meet them with the IRI Voracity data management platform.

HARNESSING THE POWER OF CLOUD DATA WAREHOUSES AND DATA LAKES

These architectures combine traditional and modern data capabilities, providing organizations with scalable and flexible solutions to handle vast amounts of data.

IRI Voracity can run data manipulation and masking jobs on cloud VMs or bare metal and manage data directly in cloud databases (relational or NoSQL), and file stores. This allows businesses to leverage the power of cloud infrastructure to scale data processing and enhance performance. Robust ETL functionality in Voracity—powered by IRI CoSort and front-ended in Eclipse—can ensure that data is accurately and efficiently ingested, transformed, and loaded into cloud data warehouses and lakes to generate reliable insights faster.

EMBRACING REAL-TIME PLATFORMS AND DATA FABRICS

These technologies enable organizations to access and analyze data ad hoc, providing a competitive edge in today's fast-paced business environment. Real-time data integration in Voracity empowers data architects to build dynamic data pipelines that deliver up-to-the-minute insights.

Data fabric capabilities in Voracity support a unified and cohesive approach to data management. By connecting, integrating, and federating data in disparate sources, Voracity users can achieve greater visibility and control over their data.

AUTOMATING DATA INGESTION, TRANSFORMATION, AND ORCHESTRATION

Intuitive job design and automation allow data architects to create and run complex data processes with ease. The IRI Workbench GUI for Voracity supports multiple data sources, data profiling and quality features, and diverse automation options. These capabilities help data stay clean, accurate, and ready for analysis.

ENHANCING DATA SECURITY AND GOVERNANCE

With the growing prevalence of data breaches and stringent regulatory requirements, businesses must adopt comprehensive

data governance strategies to protect sensitive information. Powerful data discovery, re-ID risk scoring, PII masking, subsetting and synthesis wizards in Voracity allow organizations to identify and mitigate data risk in both production and lower (test) environments.

EMPOWERING ADVANCED ANALYTICS AND SELF-SERVICE CAPABILITIES

The demand for reliable analytics tools that support advanced use cases continues to grow. Requirements include sophisticated modeling and knowledge graphs, self-service data exploration and real-time recommendations.

IRI Voracity supports embedded BI (direct reporting), data wrangling for third-party analytic tools, and data democratization through integration with DataSwitch to derive insight from their data. These features also foster a data-driven culture and accelerate more informed decisions.

CONCLUSION

The IRI Voracity platform—with its affordable, accessible data discovery, integration, migration, governance, and analytics capabilities—assists data architects and governance teams as they address these modern data management challenges.

IRI Voracity