Eckerson Group | White Paper

The Definitive Guide to Data Observability for Analytics and AI

BY KEVIN PETRIE AND WAYNE ECKERSON

FEBRUARY 2021

RESEARCH SPONSORED BY ACCELDATA

acceldata

About the Authors

For 25 years **Kevin Petrie** has deciphered what technology means to practitioners, as an industry analyst, writer, instructor, marketer and services leader. Kevin launched, built and led a profitable data services team for EMC Pivotal in the Americas and EMEA, and ran field training at the data integration software provider Attunity (now part of Qlik). A frequent public speaker and author of two books on data streaming, Kevin also is a data management instructor at eLearningCurve. He has a B.A. from Bowdoin College and M.B.A. from UC Berkeley.

Wayne W. Eckerson is an international thought leader in the data and analytics field since the early 1990s. He is a noted speaker, sought-after consultant, and widely read author. Eckerson has conducted many groundbreaking research studies, chaired numerous conferences, written two widely read books, and consulted with numerous organizations around the world. Eckerson is the president of Eckerson Group, a research and consulting firm that specializes in data analytics. He has degrees from Williams College and Wesleyan University.

About Eckerson Group

Eckerson Group is a global research and consulting firm that helps organizations get more value from data. Our experts think critically, write clearly, and present persuasively about data analytics. They specialize in data strategy, data architecture, self-service analytics, master data management, data governance, and data science. Organizations rely on them to demystify data and analytics and develop business-driven strategies that harness the power of data. **Learn what Eckerson Group can do for you!**

Table of Contents

Executive Summary

Exploding data supply and demand are pushing modern data pipelines to the breaking point. Enterprise data consumers want to use more data from a wider variety of sources, often on a real-time basis, to improve decision-making and optimize operations. But data teams struggle to architect, build, and operate the data systems that can meet these rapidly expanding business requirements.

New tools and platforms, combined with bigger investments in engineering and operations, only partly ease the pain. The reality is that most enterprise data teams still spend the bulk of their time fire-fighting daily operational issues. The problem is only getting worse as massive data volumes, data pipeline complexity, and new technologies conspire to overwhelm data team capabilities and undermine the business value of data systems.

The paradigm of data observability seeks to address this new world of unprecedented data complexity. Data observability offers a systematic approach by building on its predecessor technology, application performance monitoring (APM). It seeks to monitor and correlate data events across application, data, and infrastructure layers. By doing so, it enables business owners, DevOps engineers, data architects, data engineers and site reliability engineers to detect, predict, prevent, and resolve issues—sometimes in an automated fashion—that would otherwise break production analytics and AI.

To succeed with data observability, data analytics leaders must assemble and prioritize requirements, then select a comprehensive data observability product that minimizes custom integration work. They should tackle small, achievable observability projects first, enlisting a cross-functional team of contributors to focus on key pain points, such as performance and efficiency. Success on early projects can lead to more ambitious observability efforts—provided business and IT leaders continue to replace and retire duplicative older tools.

Evolution of Data Pipelines

For years, enterprise data pipelines served rigorous, but relatively stable, requirements for business analytics. Small teams of business intelligence (BI) analysts needed periodic historical measures of their sales pipeline, financial position, inventory levels, and other functional metrics. They relied on data engineers to build basic data pipelines, in which a handful of data extraction, transformation, and loading (ETL) and change data capture (CDC) tools ingested structured data from databases and applications. Data engineers used these tools to transform the data and load it into data warehouses in batch jobs, usually overnight. Analysts used conventional BI software to generate dashboards and reports.

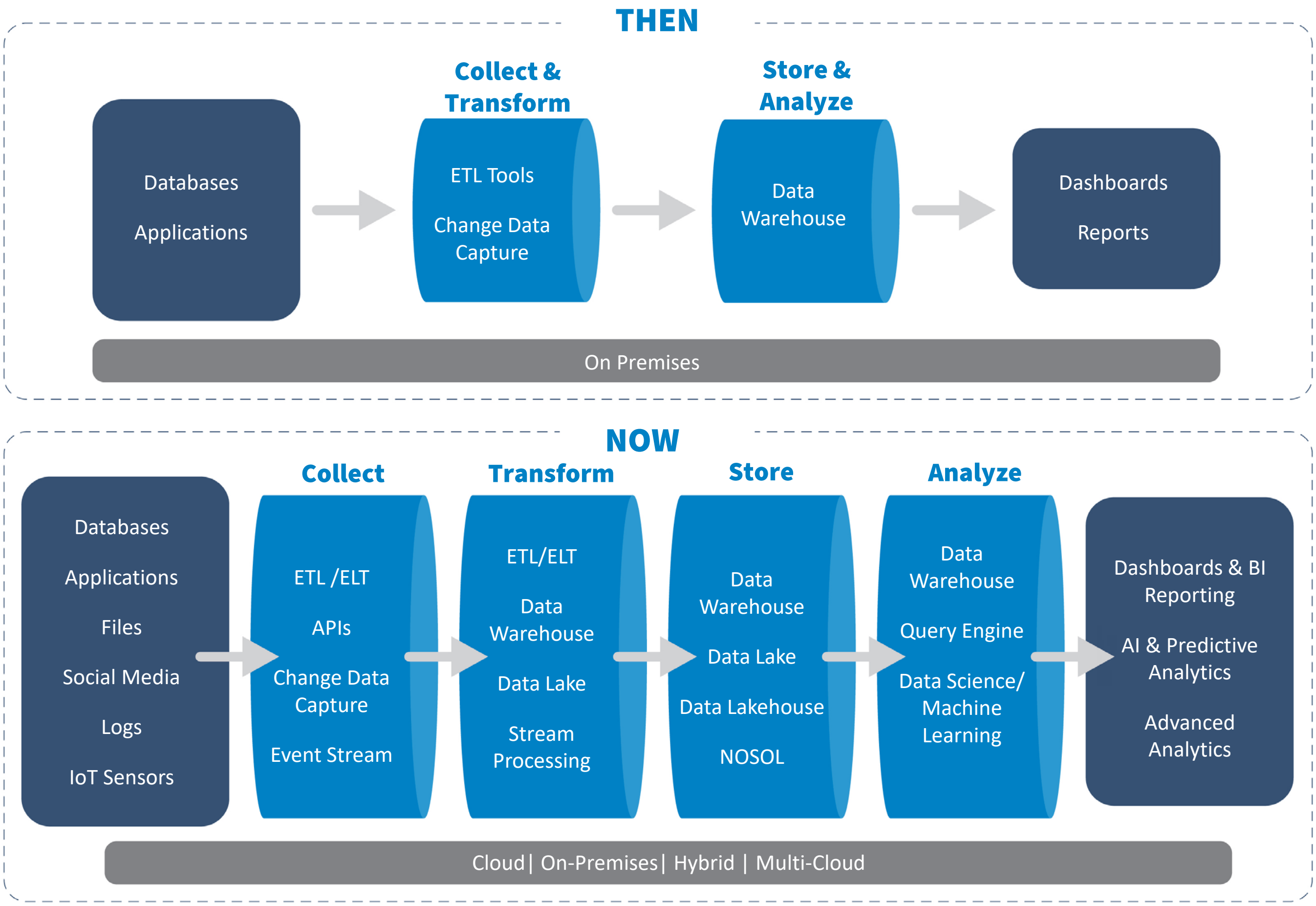
Growing Complexity. About 15 years ago the supply and demand of data exploded, and data pipelines have grown steadily more complicated ever since. A growing population of enterprise data consumers, ranging from operational managers to data analysts to data scientists, started using new algorithms to generate new intelligence and analysis from new data sources, often on a real-time basis. This prompted data teams to adopt new tools and platforms, ranging from Snowflake for BI to Databricks and Amazon Sagemaker for data science. These tools and platforms leverage cloud-native storage and compute infrastructure that offer unprecedented elasticity and scalability. All the while, data volumes tick upward.

Architects and data engineers now build data pipelines with a plethora of tools: extraction, loading, and transformation (ELT) tools, CDC, application programming interfaces (APIs) and event streaming systems such as Apache Kafka. They ingest structured, semi-structured, and unstructured data from social media, IT logs, and Internet of Things (IoT) sensors. They transform and store that data in data warehouses, data lakes, NoSQL, and even streaming platforms. Cloud object stores and compute engines must integrate with legacy on-premises systems, increasing complexity. Amidst this hybrid- and multi-cloud mess, companies must build fragile data pipelines that feed data to a multiplying set of targets, including BI tools, artificial intelligence (AI), and embedded analytics workflows.

Figure 1 illustrates the impact of the data supply and demand explosion on enterprise data pipelines.

This complexity and rising tide of data can overwhelm the enterprise teams that manage the infrastructure, applications, and networks underpinning modern analytics pipelines. They struggle with slow, unwieldy Hadoop data lakes, which persist on-premises thanks to data gravity. They struggle to control the performance and cost of cloud data platforms, while maintaining integrated views across hybrid, multi-cloud environments. It's become increasingly difficult to deliver data and analytics pipelines with sufficient levels of performance, availability, and reliability. In short, the complexity of data pipelines undermines the value they deliver to the business.

Figure 1. The Growing Complexity of Enterprise Data Pipelines



Rising users and use cases; data volumes and sources; hybrid and multi-cloud platforms

The complexity of data pipelines undermines the value they deliver to the business.

This is where observability technology comes in. It enables data teams to monitor data infrastructure underlying modern analytics and AI applications and detect issues that might constrain performance or cause outages. Effective data observability tools must address requirements at each layer of the stack.

- **Infrastructure layer.** Platform engineers, DevOps engineers and site reliability engineers (SREs) need to monitor storage and compute availability, utilization, performance, and their impact on data flows. Without sufficient visibility into these systems and the ability to correlate activity with data and analytics pipeline issues, engineers hit operational issues, performance bottlenecks and system outages.

> **Data layer.** Data architects and data engineers need to monitor database and networking applications, such as Apache Spark and Apache Kafka, that power modern data and analytics pipelines. To improve processing throughput and minimize network latency, they need tools that quickly find and remediate issues. Otherwise, data timeliness and quality fall short of service-level agreements (SLAs).

> **Application layer.** BI analysts, data scientists, and business managers need to understand the root causes of performance issues that make it challenging to build and use analytics tools and data. Traditional application performance monitoring (APM) tools identify application issues, but they cannot answer data-related questions such as why an ETL job hung up.

Without observability tools, these professionals play a game of Whac-A-Mole, knocking down one problem after another, only to have the next problem arise almost instantaneously. It's a constant struggle. Data teams patch together monitoring views by customizing individual tools. They apply fast new technologies like Dremio SQL query engines, Amazon Kinesis stream processors, or the Apache Druid distributed data store. But in the end, they still must manually hunt down performance issues and determine root causes of outages that adversely impact system reliability and the data consumers and business users at the end of an analytics pipeline.

Data Observability Defined

Data observability proposes a systematic solution. It seeks to improve control of all the elements that handle AI and analytics workloads. Observability enables business and IT organizations to monitor, detect, predict, prevent, and resolve issues from source to consumption across the enterprise. Observability uses automation and AI to correlate thousands of alerts across multiple servers, nodes, clusters, containers, and applications. It pinpoints issues and helps solve problems, so systems can scale and start to meet expectations again. Observability also includes data quality controls to help mitigate the rising risks of inaccurate AI and analytics output. It's not a panacea, but it's a good start.

Data observability means that business and IT can monitor, detect, predict, prevent, and resolve issues from source to consumption across the enterprise data pipelines that power analytics and AI workloads.

With data observability:

> DevOps, platform, and site reliability engineers ensure infrastructure performance, efficiency, and capacity

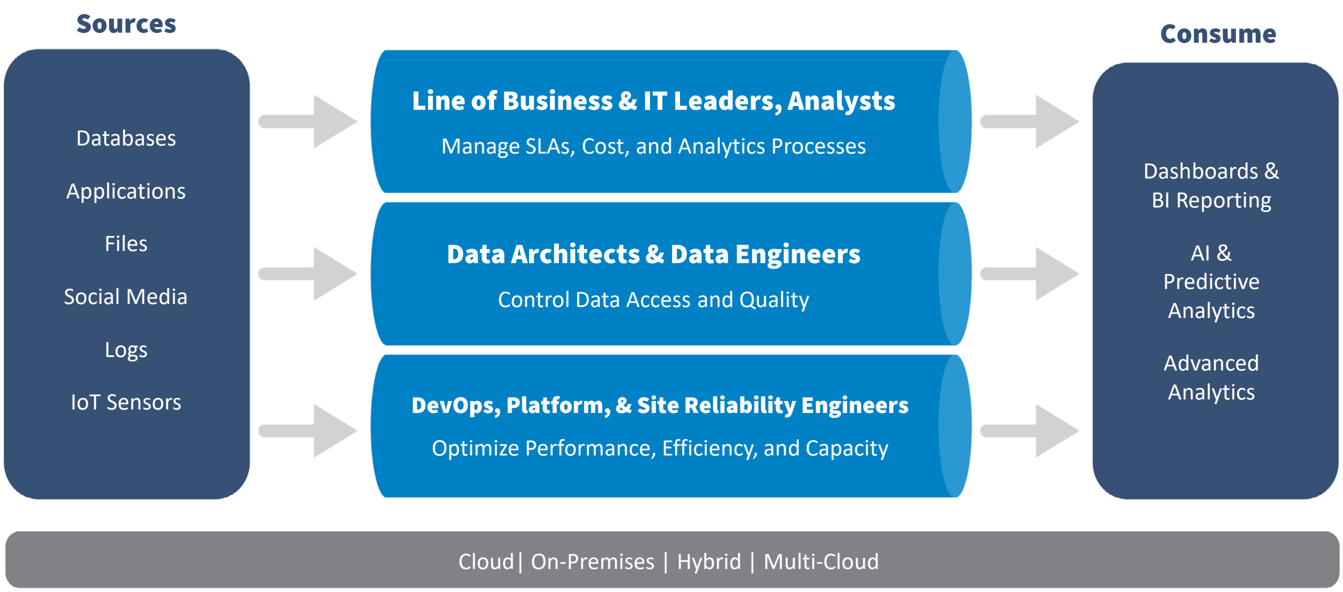
> Data architects and data engineers improve data access and quality

> Data teams can meet or exceed SLAs

> Line of business leaders, IT Leaders, and analysts improve decision-making, analytics processes planning, and cost control

Figure 2 illustrates how key stakeholders use data observability.

Figure 2. How Stakeholders Use Data Observability



Comprehensive views explain causes and inter-dependencies

Use Cases

Data observability empowers data teams and enables use cases that often span multiple roles and layers of the stack. Key stakeholders become significantly more productive because routine tasks can be automated, processes streamlined and decision-making improved.

Data observability improves staff productivity by automating tasks, streamlining processes, and enabling better decisions.

DevOps, platform and site reliability engineers

Data observability tools assist performance management, infrastructure streamlining, and capacity planning.

Infrastructure performance management. DevOps, platform and site reliability engineers can configure observability monitors for metrics such as memory availability, CPU/storage consumption, and cluster/node status. They can define alert thresholds and notifications, all sortable by entity type—user, application, CPU, disk, distributed file systems such as Hadoop File System (HDFS), processors such as Spark, schedulers such as YARN, container systems such as Kubernetes, etc. This level of granularity helps identify data flow congestion, outages, and runaway users or applications. Engineers can then troubleshoot and drill into jobs and components, which helps remediate issues. For example, they can automatically adjust service queues, tasks, and capacity levels of containers.

Infrastructure streamlining. Datasets often have skew, meaning that most of the input/output (I/O) focuses on a small portion of data. Observability sheds light on skew to help reduce storage cost. For example, you can filter or sort your files by size to identify large files that have not been accessed recently. Archiving those files to "cold storage" such as Amazon S3 Glacier saves costs and frees up capacity to accommodate growth. You also can rebalance existing datasets in this infrastructure layer to support faster performance at the data and application layers.

Capacity planning. DevOps, platform and site reliability engineers use data observability to measure and predict the resources that are required to meet SLAs with the business. They monitor data workloads to pinpoint constrained resources, or identify spare CPUs and re-assign tasks to them. AI-driven features help them calculate future capacity requirements based on available capacity, necessary buffer, and expected workload growth. They set capacity ranges by forecasting when starved resources or runaway workloads will create performance redlines.

Data architects and data engineers

Data observability helps data teams manage pipeline performance and data quality, improving architectural efficiency and effectiveness over time.

Data pipeline performance management. Data architects and data engineers must automatically collect thousands of pipeline events, correlate them, identify and inspect anomalies or spikes, and then use those findings to predict, measure, prevent, troubleshoot, and fix all kinds of issues. For example, they must closely track and adjust how distributed schedulers assign jobs to cluster nodes in an on-premises data lake. They need to monitor read and write I/O for an Apache Spark or Hive server and correlate those metrics to memory or CPU utilization and execution time, as another example. Data observability tools provide such views to help and recommend ways to tune performance, such as by boosting CPU or memory.

Data quality. Leading observability tools take a page from the DataOps playbook and automatically inspect data transfers for accuracy, completeness, and consistency. They create rules to compare sources to target tables—or incremental updates—then flag mismatches for alerting, review, drilldown, and reconciliation. Violations might include null values, duplicate records, altered schemas, or mismatched value ranges. Data observability tools also track lineage end to end, integrating with both sources and target BI tools. Such operational data quality checks ensure data pipelines meet expectations. However, they do not eliminate the need for dedicated data quality solutions, which can address issues like industry-specific compliance regulations.

Architectural design. Data architects and data engineers also must step back from daily firefights to design better architectures. They can apply newfound insights into pipeline performance and utilization trends to connect the dots between what they have today and what they need tomorrow. They can select, deploy, and configure new platforms such as the high-performance Apache Druid with confidence, based on observed workload behavior, scenario modeling, and impact analysis. While there is no silver bullet in a fast-changing world, data observability reduces architectural planning risk.

Line of business and IT leaders, and analysts

Data observability also makes life easier for the line of business leaders, IT leaders and analysts that both consume data and oversee its usage.

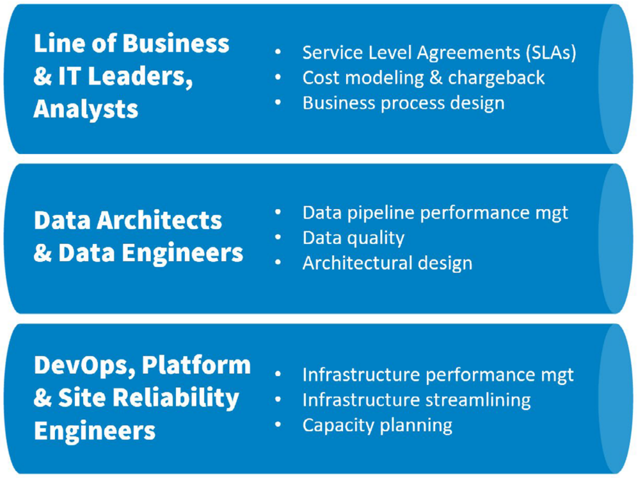
Service Level Agreements (SLAs). Voracious data consumers continue to demand rigorous latency, throughput, and uptime service commitments from the data architects and data engineers that make it happen. While this tension will not ease anytime soon, data observability helps both sides enter more responsible SLAs. Site reliability and platform engineers create more accurate capacity estimates. Data architects and data engineers design faster, more reliable data pipelines. And line of business leaders, IT leaders and analysts can get better guidance about what SLAs are really feasible.

Cost modeling and chargeback. Line of business and IT leaders use data observability to manage the business of data analytics in a more accurate, granular way. They can collaborate with BI analysts, data architects, and data engineers to model the implied operational cost of their SLAs, based on estimates such as aggregate compute, storage, memory, and data transfer requirements. They can slice these estimates by time period, user group, geography, etc., to help with granular budgeting and chargeback decisions. While no single data observability tool covers every possible component, these offerings enable more informed financial decisions than were previously possible.

Business process design. Enterprises continue to embed analytics and AI into more aspects of their operations for a variety of reasons, including to enable real-time fraud prevention, customer recommendations, or IoT preventive maintenance. They need to enhance rather than disrupt operations, which means they need to have high confidence in their data workload SLAs. Data observability helps. Line of business leaders, IT leaders, data architects, and data engineers can collaborate to design creative analytics-operational workflows with an acceptable level of risk, and ideally a reasonable ROI.

Figure 3 summarizes these use cases by role.

Figure 3. Data Observability Use Cases



Market Landscape

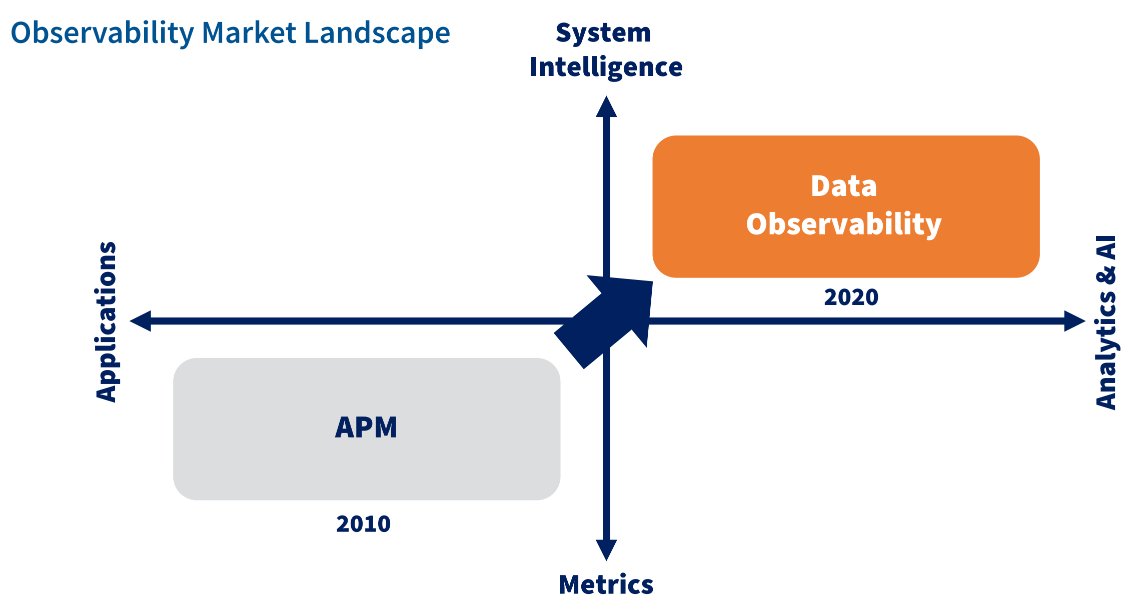
Data observability extends the core capabilities of APM tools that enterprises have used for years to optimize applications. Traditional APM tools discover applications, monitor their workings, and diagnose issues, often with the assistance of AI and machine learning (ML). They trace requests and correlate events across microservices, servers, containers and other resources. They identify risks, slowdowns, blockages, and failures, then help remediate them to improve and maintain operational workloads.

Data observability builds on APM in two ways. First, it applies these familiar APM functions to data analytics workloads rather than operational applications. Second, data observability provides deeper correlation of events across the data, applications, and infrastructure layers. This integrates metrics into system intelligence. For example, data engineers can use observability to isolate slow ETL jobs at the data layer, then work with platform or site reliability engineers to isolate the storage configuration errors that are causing issues. They can prioritize Spark jobs by execution time, then drill down into the problematic areas for inspection and debugging. Data observability also helps analysts and business owners control the efficiency and cost of their enterprise analytics and AI stacks.

Data observability applies traditional APM functions to data analytics workloads.

Figure 4 compares APM and data observability.

Figure 4. Comparison of APM and Data Observability



Data observability shares capabilities with two additional market segments. First is DataOps, which applies the principles of DevOps and agile software development to the creation and management of data pipelines. Data observability addresses the monitoring, data validation, and lineage pieces of the DataOps puzzle, helping ensuring data quality and data delivery performance. Second, ITOps tools (and their AIOps sub-segment) provision, manage, monitor, and tune IT infrastructure resources. Data observability addresses the monitoring, diagnosis and remediation aspects of ITOps and AIOps.

Putting Data Observability to Work

Before getting started on an observability initiative, data teams should understand the role of data modernization as an adoption driver. They also should consider the challenges of data observability and its intended benefits.

Data Modernization

Enterprises are pushing hard to monetize their data assets, which drives adoption of large, shared platforms that need data observability. There are several dimensions to this. Many business-oriented analysts now use automated and pre-packaged ML code, which expands the ML user base beyond data scientists. These projects consume high volumes of diverse data, including social media text, IoT sensor feeds, and even satellite imagery. They drive up the need for speed and scale, which in turn prompts data teams to modernize their architectures with streaming, data lake and data warehousing platforms based on elastic cloud infrastructure. While cloud-native resources typically improve performance and ease of use, their linkages back to legacy on-premises systems create new risks and complexity that need attention.

Challenges

Complexity, paradoxically, is also the primary challenge facing observability initiatives. No single commercial solution can cover all the permutations of a modern environment, which means data teams must either customize their observability software or accept less visibility than they need. Complexity also bedevils the task of monitoring: data teams must carefully configure and filter their logs, traces, and metrics to improve the signal to noise ratio. Without sufficient attention up front to these details, observability solutions might create confusion, hurt productivity, and raise performance risk. Early adopters of data observability such as GE Digital have overcome these challenges, for example by consolidating several tools onto a single monitoring platform that is easier to manage (see Case Study below.)

Benefits

When planned, implemented, and managed well, observability yields several benefits. Lower latency, higher throughput, and more accurate data all help analysts and data scientists capture new value from their analytics and AI projects. Improved data pipeline reliability, uptime, and issue resolution reduce operational risk. Consolidation of monitoring tools reduces administrative overhead, improves data team productivity, and creates a "talent bridge" that reduces the need for training or hiring. Data architectures also become more flexible, scalable, and efficient, able to bend with the needs of the business.

Common Environments

Data observability provides the greatest value for three overlapping types of environments: on-premises Hadoop data lakes, hybrid cloud, and multi-cloud. Each brings a different mix of challenges and requirements.

On-premises Hadoop

Data observability helps data teams manage Hadoop data lakes that persist in large on-premises environments. Enterprises planted some analytics data and workloads in Hadoop five to ten years ago, and many have been slow to abandon the investment altogether even as more manageable, cost-effective alternatives arise on the cloud. As a result, some data teams still run analytics on Hadoop and need help maintaining performance levels across myriad Apache open source components. Spark accelerates batch processing on Hadoop compared with MapReduce, but often needs careful monitoring, troubleshooting, and debugging to meet production analytics latency and throughput requirements. Data observability can be used to improve performance, reliability, and scalability.

Hybrid cloud

Data warehouses and data lakes are converging on a common set of functions in the cloud, pairing high-performance SQL query structures with efficient, elastic object storage. Enterprises adopt these cloud data platforms, such as Azure Synapse, Databricks, and Snowflake, to reduce administrative hassle and consolidate data workloads. But data teams still need observability to keep a close eye on cloud platform performance, for example, to meet BI query latency requirements. They also need to reduce the risk of compute cost overruns—and maintain all the necessary linkages back to legacy on-premises systems.

Multi-cloud

As enterprise data teams gain experience on the cloud, they seek to optimize workloads and meet specialized requirements by shopping around for new AI tools, cloud data platforms, etc. As a result, many enterprise environments now include two or even three Cloud Service Providers in addition to legacy on-premises systems. Observability helps them oversee these distributed topologies and maintain efficient, effective data pipelines.

Case Study

GE Digital. GE Digital runs the fifth-largest Oracle Enterprise Resource Planning (ERP) and third-largest SAP ERP implementations in the world, leveraging 1,000s of Amazon Web Services (AWS) cores. Before embracing observability, GE's data team could not understand and control the interactions of the myriad data pipeline components that process six million financial transactions per minute for both operational and analytical workloads. This undermined GE's ability to meet performance targets. "There were too many handshakes," says chief data officer Diwakar Goel, and "the cost of a mistake is very high."

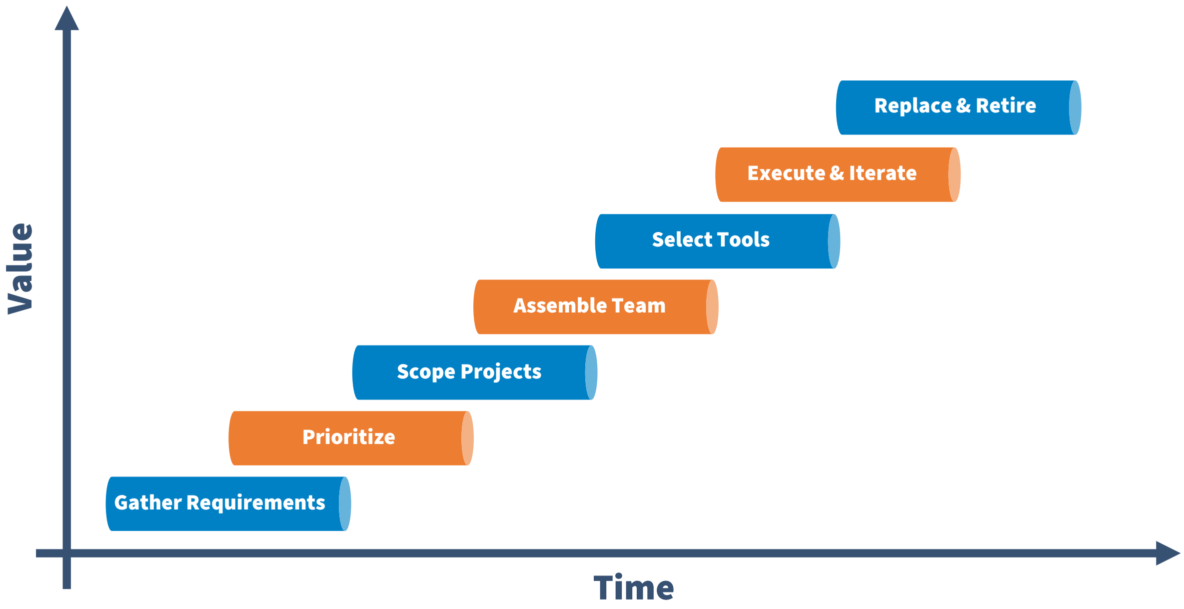
GE deployed an observability solution from Acceldata to regain control of their data pipelines. Their data team centralized monitoring activities from several tools, including open-source Apache Ambari for managing Hadoop and Logstash for log analytics, onto Acceldata. They tamed longstanding Apache Spark processing performance issues, freeing the engineering team to migrate analytics workloads to SingleStore (previously known as MemSQL) on AWS from a problematic Hadoop data lake. Goel estimates GE's operating costs were reduced by $30 million annually through the finance data lake effort. Observability is mandatory for their mission-critical systems.

Before embracing observability, GE's data team could not understand and control the interactions of the myriad data pipeline components that process six million financial transactions per minute for both operational and analytical workloads.

Start Observing

Data observability has legs because it addresses an indisputable technology problem—namely, complexity—that will only increase in severity in coming years. Enterprises must adopt a sound observability program that decreases rather than increases complexity. Achieving this requires a methodical, step by step approach, with each successive step delivering new value to your organization. Figure 5 illustrates the phases and relative value of a data observability program.

Figure 5. Phases in a Data Observability Program



Let's consider how to make fast, then sustainable, progress with each of these steps.

> **Gather requirements.** The value of data observability depends on its holistic approach. Build a comprehensive inventory of requirements that covers the breadth of your users, use cases, technology components, and pipelines across your infrastructure, data, and application layers. For instance, you might need to resolve long-standing Spark bottlenecks that impact "real-time" customer offer generation and conversion. You might be struggling to anticipate data science workload spikes that cause runaway cloud compute costs. At another level, you might need to create or improve chargeback estimates based on actual infrastructure resource consumption. Document as many of these requirements as you can, including potential future needs.

- **Prioritize.** Identify the most painful and "fixable" requirements. You can assess pain according to business metrics—i.e., revenue or cost impact due to transaction processing performance, customer wait times, etc. "Fixable" requirements are those most easily remediated, at the lowest risk, with the highest business benefit. While needs vary widely, the higher priority requirements often derive from on-premises data lake implementations, hybrid cloud data migrations and advanced analytics initiatives.
- **Scope projects.** Start scoping your first project or projects to address these highest priority requirements. Define the capabilities you need to change, and how you should change them. What people, process and technology changes does this project entail? You might need to hire or re-train employees, reconfigure data pipelines, change out applications or implement new tools. Define all the associated tasks, as well as their dependencies and risks, then plot achievable project milestones to improve and measure progress. By starting with a small and more achievable project, you increase the odds of a quick win that demonstrates traction to the larger organization.
- **Assemble your team.** Next you need to identify and recruit the necessary cross-functional team members to drive your program. You need executive sponsorship from both a business leader and IT leader, as they can allocate the necessary resources to ensure success. Consider forming a Center of Excellence, which IT trains various stakeholders across the business units to manage data observability with common policies, practices and tools. Carefully select a small team for your first project in particular, working with managers to identify the right technical and business subject matter experts among the ranks of analysts, data architects, and data engineers, as well as DevOps, platform and/or site reliability engineers. Collectively the project team should have all the necessary knowledge for your focus requirements and use case, which might center on machine learning, data warehousing, ETL, streaming, etc.
- **Select tool(s).** In all likelihood, your requirements phase identified gaps in your monitoring capabilities that only a comprehensive data observability product can fill. You can evaluate potential products based on a few key criteria. Does it provide sufficient visibility into the current and near-term future components of your application, data and infrastructure layers? What level of customization does it require? What level of training does it require, and do the expected benefits of the tool justify the cost and ramp-up time? Can this tool replace other tools in your environment?

- **Execute and iterate.** Once you have selected and implemented your observability tool, you can start project execution. Be sure to measure the success of each project—whether it is resolving those Spark bottlenecks, forecasting data science workloads, improving chargeback, etc.—according to business-oriented benchmarks. What is the revenue impact of faster customer recommendations? How much are you able to reduce overall costs with better workload forecasts or chargeback? Based on your answers, you can refine your existing projects and scope more successful future projects. Once you demonstrate success, you can secure executive sponsorship and funding to tackle more strategic projects.

- **Replace and retire.** A classic IT mistake is to adopt a fancy, easy to use technology, but keep its clunky predecessor, and thereby make life harder than ever. Don't let this happen with data observability. Build and stick to a clear phased plan for your new observability product to replace those functions in other APM or ITOps tools. Otherwise, your team will get more distracted and less efficient than ever, with the expected impact on performance, data quality, etc. By ruthlessly streamlining and phasing out old processes, you can squeeze additional cost and efficiency out of the system.

Data observability offers enterprises the potential to give exploding data workloads the attention they deserve at all levels of the IT stack and organization. Data analytics leaders should get smart on their options to ease the pain and risk of data access, performance, quality, efficiency and cost. By doing so, they can resolve existing problems, while significantly increasing the upside of what's achievable in future analytics and AI projects.

About Eckerson Group

Wayne Eckerson, a globally-known author, speaker, and consultant, formed **Eckerson Group** to help organizations get more value from their data. His goal was to provide organizations with expert guidance during every stage of their data and analytics journey.

Today, Eckerson Group helps organizations in three ways:

> **Our thought leaders** publish practical, compelling content that keeps data analytics leaders abreast of the latest trends, techniques, and tools in the field.

> **Our consultants** listen carefully, think deeply, and craft tailored solutions that translate business requirements into compelling strategies and solutions.

> **Our advisors** provide one-on-one coaching and mentoring to data leaders and help software vendors develop go-to-market strategies.

Eckerson Group is a global research and consulting firm that focuses on data and analytics. Our experts specialize in data governance, self-service analytics, data architecture, data science, data management, and business intelligence.

Our clients say we are hard-working, insightful, and humble. It all stems from our love of data and desire to help organizations harness the power of data. We are a family of continuous learners, interpreting the world of data and analytics for you.

Get more value from your data. Put an expert on your side. **Learn what Eckerson Group can do for you!**

Eckerson Group

RESEARCH CONSULTING ADVISORY

GET • MORE • VALUE • FROM • YOUR • DATA

About Acceldata

Acceldata is the industry's first Unified Data Observability platform for analytics & AI systems, enabling enterprises globally to thrive in a data-driven world. With the ability to observe, optimize, and scale complex data pipelines, enterprises leverage Acceldata's Observability platform to deliver seamless analytics and AI success in the cloud and on-premises. GE, True Digital, Walmart PhonePe, Michelin, PubMatic, DBS, and many other global enterprises are Acceldata customers that optimize their data success. We invite you to visit us at Acceldata.io and follow us on LinkedIn and Twitter.