unravel™

THE ROLE OF DATA OBSERVABILITY AND OPTIMIZATION IN ENABLING AI-DRIVEN INNOVATION

Data observability will play a key role in enabling organizations to drive innovation and deliver leading AI-driven applications and services.

Written by: Nancy Gohring, Research Director, Enterprise System Management Observability and AIOps Software

Harnessing Data Observability for AI-Driven Innovation

Organizations are now embarking on a journey to harness AI for significant business advancements, from new revenue streams to productivity gains. However, the complexity of delivering AI-powered software efficiently and reliably remains a challenge. With AI investments expected to surge beyond $520 billion by 2027, this brief underscores the necessity for a robust intelligence architecture, scalable digital operations, and specialized skills. Learn how AI-driven data observability can be leveraged as a strategic asset for businesses aiming to lead in innovation and operational excellence.

Continue reading to see the IDC Analyst Brief by Research Director Nancy Gohring.

databricks   snowflake   Google Big Query

Request a Health Check
Get free insights into your cloud data platform usage, costs, and projected savings.

Data observability will play a key role in enabling organizations to drive innovation and deliver leading AI-driven applications and services.
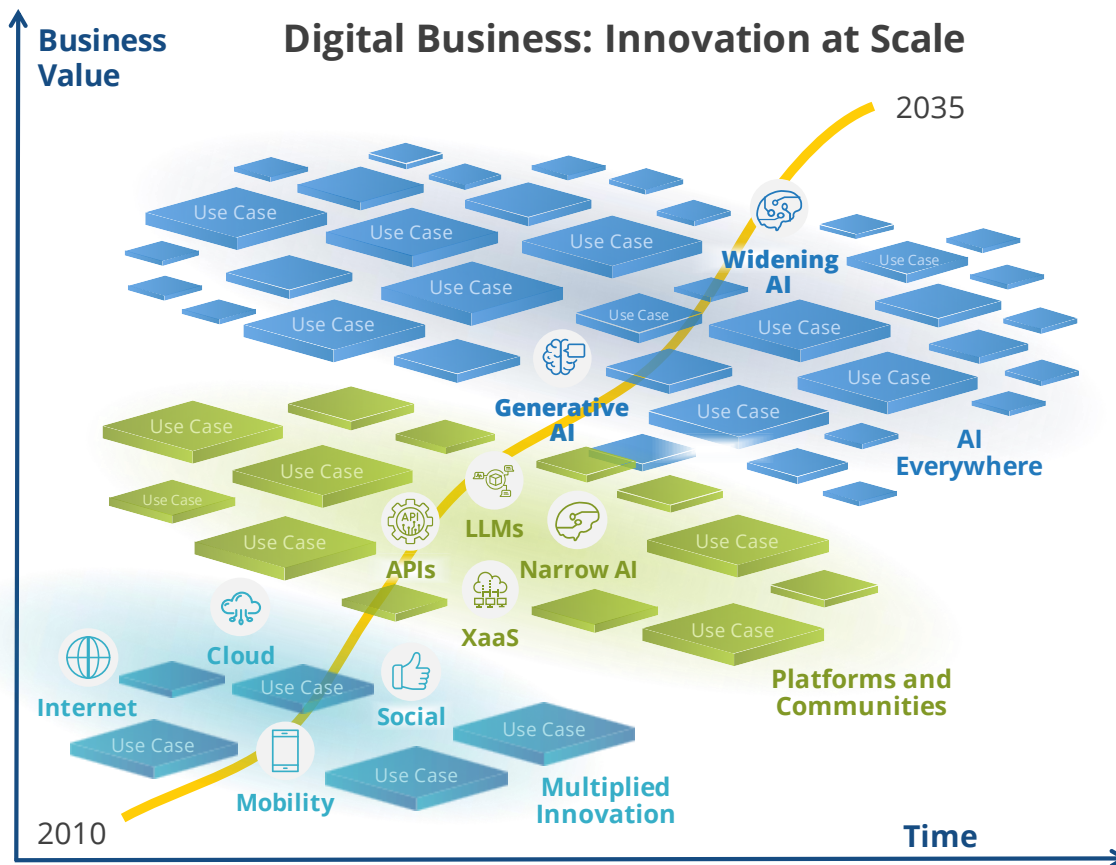
The Role of Data Observability and Optimization in Enabling AI-Driven Innovation

March 2024

Written by: Nancy Gohring, Research Director, Enterprise System Management Observability and AIOps Software

Executive Graphic

FIGURE 1: *AI Everywhere as the Next Digital Business Era*

Source: IDC, 2024

Introduction

AI is driving a dramatic change for enterprises. Not only are organizations adopting software that uses AI, but organizations are also recognizing that they can build AI-driven apps to deliver notable business outcomes, ranging from significant new revenue streams to large internal efficiency gains. However, most enterprises are currently struggling to put in place the requirements to enable the rapid, cost-effective, and reliable delivery of software that uses AI models.

IDC anticipates a multiplying of AI use cases driven by $521 billion in spending by 2027, resulting in a much wider application of AI than today.

To support the growing deployment of AI, enterprises will need to put in place:

» An intelligence architecture that leverages a data-centric platform

» Digital operations at scale that enable the ability to cost-effectively consume digital infrastructure

» Targeted skills, which mean attracting and reskilling talent, as well as adopting tools that are easy to use

A host of new tools, processes, and skilled resources are required to build these three foundational elements that enable the rapid delivery of cost-effective, reliable AI applications. Data observability, a key component that underpins and enables all three, describes the monitoring and management of data, data pipelines, and data platforms, including the intelligent identification of problems and remediation recommendations. Data observability that employs AI becomes particularly important, using AI to intelligently determine the root cause of problems in the data stack, recommend resolutions, and empower engineers of all skill levels.

An effective AI-driven data observability implementation helps govern and optimize the following aspects of a data analytics platform.

Data

Since data fuels businesses and the applications they build and use, organizations must manage data quality or risk not only the loss of trust but the ability to achieve critical business outcomes. Companies must have the ability to identify data quality issues including duplicate, stale, incomplete, and inaccurate data, and then quickly determine the cause of the problem to quickly resolve it.

Infrastructure and Platforms

To deliver reliable and trustworthy data and analytics, organizations require a high-performing, modern data stack. That stack comprises a data pipeline running either on cloud or on-premises infrastructure. In a generative AI app, for instance, data from enterprise apps may need to be transformed and combined with information from a data warehouse and data lake. To deliver high-quality data, businesses must ensure that the infrastructure is optimized to support the data analytics platform and that the data platform is optimized for efficiency.

Cost

In AI-driven applications, performance and cost are closely tied. Inefficient data jobs and poorly optimized back-end infrastructure slow performance, leading to missed service-level agreements (SLAs). They may also cause unintended

cost spikes, as cloud data platforms scale up to meet the requirements of an inefficient job. Insight into cost and intelligent optimization techniques are the keys to the successful delivery of a data analytics platform in the enterprise.

The Role of Data Observability

Data observability, including AI-driven data observability, supports the governance and optimization of data, infrastructure, and cost so that businesses can harness data to drive innovation. A data observability implementation should have the key capabilities detailed in the sections that follow.

Visibility into the Heterogenous Data Environment

To feed AI-driven applications, enterprises often require multiple data analytics platforms running on a variety of cloud compute, storage, and network infrastructure. Visibility can't be limited to the infrastructure though; businesses require a deep level of insight into the data pipeline. In this environment, when an organization adopts too many tools, operations pros are hampered by tool management overhead, context switching, data silos, and visibility gaps. A data observability tool that offers intelligence on multiple data platforms can mitigate these challenges. It can also assess whether workloads should run on premises or in the cloud, based on performance as well as infrastructure and cost efficiencies.

AI-Driven Automation

A data observability tool that uses AI tuned to the qualities of a particular data platform can automate problem and resolution identification. Reliable intelligence specific to the back end can be applied at the beginning of the application development process so engineers can optimize apps before their deployment. Once applications are running in production, AI-driven data observability surfaces issues that are causing or may eventually cause performance problems and identifies actions to resolve them.

Flexibility

To operate a data platform that feeds an AI application, organizations require data platform owners, data engineers, data scientists, FinOps teams, and product managers. Each of these users has different requirements, drivers, and challenges. The most valuable data observability tools will have the flexibility to be customized to serve multiple types of users. In addition, data observability tools that use AI to pinpoint the root cause of a problem or recommend a remediation empower engineers and other users that may not have much experience using the tools.

Benefits

Data is the lifeblood of AI. However, applications that don't efficiently access and manage data will not succeed. Slow queries and jobs are costly and create a poor user experience, impeding the application's potential to deliver intended business outcomes. In many industry sectors, the ability to innovate via AI-powered applications and services may mark the difference between companies destined to succeed and those that will fall behind. Data observability will play a key role in enabling organizations to drive innovation and deliver leading AI-driven applications and services.

Considerations

AI-driven data observability best practices are evolving, as enterprises are still gaining experience in developing applications that use AI. Many organizations have spent years developing their first AI-powered applications and are just

beginning to put in place the people, processes, and tools necessary to operate them in production. Enterprises are hearing about the operations of data applications and infrastructure from a range of suppliers that collect and deliver telemetry. Determining which tools to use and which enterprise roles should be responsible for optimizing data workloads is challenging.

Conclusion

AI-driven data observability for the modern data stack does more than just monitor components of the data platform. An effective data observability implementation allows enterprises to efficiently and effectively deliver AI-enabled applications in a repeatable fashion. Data observability underpins the efficient delivery of an intelligence architecture. It enables the intelligence architecture's cost-effective and performant operations to support the scale required for growing use cases. Finally, AI-driven data observability lowers the barriers to expertise in this fast-moving field, surfacing intelligence that otherwise only experts with decades of experience might deliver. The outcome of an AI-driven data observability implementation is a dependable intelligence architecture that helps achieve business outcomes.

About the Analyst

Nancy Gohring, *Research Director, Enterprise System Management Observability and AIOps Software*

Nancy Gohring is research director for IDC's Enterprise System Management Observability and AIOps Software service. She publishes research describing trends impacting the application, infrastructure, and log monitoring and observability sectors as well as AIOps. Ms. Gohring offers competitive intelligence and guidance to sector-leading vendors, advises enterprises about the tools and tactics required to drive the top performance and user experience for their most important applications, and offers trend insights to the investor community.

IDC Custom Solutions

The content in this paper was adapted from existing IDC research published on www.idc.com.

≡IDC