
THE ULTIMATE GUIDE TO
BUILDING AN
END-TO-END
DATA
STRATEGY

CONTENTS

Introduction	03
Key challenges and considerations	04
Unlock more value from data	05
Comprehensive	06
Integrated	09
Governed	12
Making security more strategic	15
About TESCHGlobal	12
Benefits of building a data strategy with AWS and TESCHGLOBAL	16
Case Study	17
Learn More	18

INTRODUCTION

Now more than ever, data is at the center of every application, process, and business decision. It's the genesis for modern invention, and in today's fast changing and complicated landscape, how you put your organization's data to work can be the golden ticket to accelerating innovation and accomplishing your organizational goals. The stakes are high. According to Forrester Research, organizations that have a system to promote data-driven insights are 140 percent more likely to create sustainable competitive advantage and 78 percent more likely to fuel a revenue growth environment. With a pressing need to empower the entire organization to use data to make better, faster decisions that fuel new ideas and drive business agility, leaders are embracing a fundamental truth: The journey to innovation begins with data, and successfully becoming a data-driven organization begins by implementing an end-to-end data strategy. The good news is that you don't have to reinvent the wheel. Organizations are already capturing the benefits of an end-to-end data strategy built on Amazon Web Services (AWS). While the achievements are limitless, the central challenge is this: many organizations are sitting on a treasure trove of data but don't know how to gain value from it. In this eBook, you will learn the fundamentals of building an end-to-end data strategy to keep up with your data needs now and in the future—enabling a sustainable advantage that comes from unlocking the value of your data.

¹ "Creating a data-driven culture," CIO.com, March 31, 2022.

KEY CHALLENGES AND CONSIDERATIONS

More data than ever is being generated and stored

On-premises tools and legacy data stores can't meet today's demands, organizations need new data stores that can scale and grow as business needs change—whether from the gigabytes and terabytes handled today or the petabytes and exabytes that will be managed in the future.

Data siloed across multiple sources creates productivity and cost inefficiencies

Organizations need to easily access and analyze diverse types of data. However, wide-ranging data types are typically stored in silos across multiple data stores. To extract intelligence, organizations must break down these silos to unify all types of data. This important optimization of costs and operations is transforming the infrastructure from a source of complexity and expense to an engine of value creation.

Current state of decision-making is unsustainable

65 percent of decisions made today are more complex (involving more stakeholders or choices) than they were five years ago.² To make better and faster decisions, organizations need the ability to perform analytics and machine learning (ML) operations in an agile, cost-effective way—using optimal tools and performance to scale for each use case.

Analytics and machine learning adoption is still impeded by a lack of skills and inertia

Many businesses are struggling to make progress with scaling analytics and ML tools. Gartner finds that organizations investing in AI moved just 54 percent of their AI proof-of-concept pilots into production.³ A continued lack of data and ML skills and quantity or quality of data to train on are just some of the issues slowing progress in this important area.

Trying to maintain data governance is a full-time job

Traditional data architectures require risky, complicated management procedures because data is accessed from so many places. Granting, tracking, auditing, and removing employee access—while simultaneously remaining in compliance with a growing number of regulations—is a full-time job. Automating these mandatory data governance tasks frees teams to shift their focus back to innovation.

Data is increasingly difficult to secure

There was a time when IT teams chose between making their architectures fast or making them secure. Now, they need to deliver both. Meanwhile, security attacks increased by 31 percent from 2020 to 2021, while average attacks per organization grew from 206 to 270 year over year.⁴

2 "HOW TO MAKE BETTER BUSINESS DECISIONS," GARTNER, OCTOBER 20, 2021, 3 "HALF OF AI MODELS NEVER MAKE IT TO PRODUCTION," ENTERPRISEAI, AUGUST 23, 2022, 4 "HOW ALIGNING SECURITY AND THE BUSINESS CREATES CYBER RESILIENCE," ACCENTURE, 2021

UNLOCK MORE VALUE FROM DATA

According to a PwC survey of more than a thousand senior executives, highly data-driven organizations are three times more likely to report significant improvements in decision-making compared to those that rely less on data.⁵

AWS can help your organization implement an end-to-end strategy that makes data management easier at every step of the journey—from ingesting, storing, and querying data to analyzing, visualizing, and running ML models. Regardless of your business challenges, your data strategy should be:

1 Comprehensive

Equipped with the right tools, with the optimal price performance for any user, type of data, and use case

2 Integrated

The ability to integrate data that is stored and analyzed in different tools and systems to gain a better understanding of your business and predict what will happen

3 Governed

Governance of all your data to securely give data access when and where your users need it to speed innovation

A data-driven mindset may also require a broader cultural change in which both goals and decisions are supported by the data strategy.



⁵ "DATA AND INTUITION: GOOD DECISIONS NEED BOTH," HARVARD BUSINESS PUBLISHING, JANUARY 6, 2023

1 COMPREHENSIVE

Equipped with the right tools, with optimal price performance for any user, use case, and data type

Businesses need to build future-proof data strategies that can meet their needs now and in the future. It takes more than just a single data lake, data warehouse, or business intelligence (BI) tools to harness data effectively. It requires an end-to-end data strategy with a comprehensive set of tools that accounts for the scale and variety of data and the many purposes for which you want to use it. In fact, 94 percent of the top one thousand AWS customers use more than 10 AWS Database and Analytics services.⁶

Building with a cloud provider that innovates to continuously bring you all the data tools you'll need and more with the right price performance for your use case ensures that you have a data strategy that grows with you. AWS has a broad and deep set of data capabilities to support any data workload or use case. From databases for applications to storage for data lakes to analytics to ML and end-user tools, AWS provides the right capability in each area, so you don't have to compromise on performance, cost, or results.

Scaling data-driven applications

Build future-proof applications on a modern data infrastructure for the best price and performance for your use case at scale. AWS databases include Amazon Aurora, which provides the performance and availability of commercial-grade databases at one-tenth the cost.

Providing analytics for all use cases

True agility helps organizations adapt quickly to changing business needs. To power these rapid actions, AWS analytics services enable your organization's teams to ingest, combine, and run historical, real-time, and predictive analytics on all of your data. This includes services for SQL querying, log analytics, streaming, and Apache Spark. For big-data querying, Amazon EMR supports more big-data frameworks than any other provider and gets you up to two times faster time-to insights. To make decisions in real-time, you'll need streaming data services such as Amazon Kinesis Data Streams (Amazon KDS), which allow you to build applications for high-frequency event data such as clickstream data, and gain access to insights in seconds. Amazon Kinesis Data Firehose simply and reliably loads data streams into data lakes, warehouses, and analytics services—no extract, transfer, and load (ETL) or cumbersome data preparation required.

⁶ "Data: The genesis for modern invention." AWS Big Data Blog, November 30, 2022

1 COMPREHENSIVE

Deploying data science and machine learning

ML adds intelligence to existing processes, automates time-intensive manual tasks, and accelerates innovation with the creation of new products and services. With AWS, you have access to the most comprehensive set of AI and ML services. With Amazon SageMaker, used by tens of thousands of customers, you can build, train, and deploy ML models for any use case with fully managed infrastructure, tools, and workflows.

Enabling data insights throughout the organization

It's no longer just data-savvy individuals who can rapidly extract valuable, relevant insights from data to help inform decision making. ML-powered BI solutions such as Amazon QuickSight, enable easy connectivity to data sources. Business analysts can utilize this data to showcase fresh trends and predictive insights on interactive BI visualizations and dashboards. QuickSight Q uses ML, allowing users to query their data in plain language without writing a single line of code. A visual point-and-click interface enables business analysts to generate accurate ML predictions without prior experience. In just a few clicks, analysts can import data from various sources, automatically prepare data, and build and analyze ML models.

Boosting data proficiency

Having employees who can use data effectively will help your organization achieve its data objectives. Invest in educating and upskilling your workforce in data, analytics, and ML with AWS Training.

7 "Data: The genesis for modern invention," AWS Big Data Blog, November 30, 2022

94%

With a broad and deep set of data services, 94 percent of the top 1,000 AWS customers each use 10+ AWS Database and Analytics services.⁷

Scale data-driven decision making throughout your organization

Amazon QuickSight

Meet varying analytic needs from the same source of truth through modern interactive dashboards, paginated reports, embedded analytics and natural language queries

Amazon SageMaker Canvas

AWS no-code interface that enables business analysts to generate accurate ML predictions without prior experience

Amazon DataZone

Simplifies governed access to data for business users
AWS Training and Certification: More than 150 professional development courses related to data, analytics, and ML

2 INTEGRATED

Break down silos so data can be put to work effectively

Opportunities to transform your business with data exist all along the value chain. But making such a transformation requires you to see the full picture of your customer and business. With data spread across multiple departments, services, on-premises databases, and third-party applications, you need to be able to easily integrate data across silos to get the best insights. Companies have various approaches to how they are unifying data—data mesh, lake house, data fabric, and so on—but typically, it involves a data lake as a foundational element. Data lakes allow you to collect, store, organize, and process valuable data from your data silos and make it available to analytics, visualization, and ML tools in a governed way.

Zero-ETL

Many organizations have multiple data lakes in addition to data warehouses, analytics tools, ML tools, and SaaS applications. Integrating data across silos requires complex ETL pipelines, which can take hours, if not days. That's just not fast enough for modern decision making. Organizations should adopt technologies that automate or eliminate ETL where possible.

AWS is investing in a Zero-ETL future, allowing organizations to automatically integrate all of their data. This includes bringing ML to the data source with SageMaker integration into Amazon Redshift, Amazon Aurora, Amazon Athena, and Amazon Neptune, integrating Amazon Aurora and Amazon Redshift for real-time analytics; and providing a direct integration between Amazon Simple Storage Service (Amazon S3) and Amazon Redshift for real-time data streams. In addition, you can run queries across data stored in operational databases, data warehouses, and data lakes to provide insights across multiple data sources with no data movement using Amazon Athena and Amazon Redshift.

Analyzing all of your data and third-party data

To break down data silos, you can't have connections to only some of your data sources—you need to be able to seamlessly connect to all of them, whether they live in AWS or in external third-party applications, on premises, or even in another cloud environment. No matter where they live, with AWS, you can automatically integrate hundreds of data sources across AWS and third parties.

Increasingly, organizations are also harnessing third-party data to deepen insights by joining this third-party data with their own data. AWS Data Exchange enables AWS customers to access third-party data through files, tables, and APIs from more than 300 data providers and more than 3,500 data products, all from one place. Third-party data from partners and customers are also being used, which increases the need for comprehensive governance policies to protect the data. Data clean rooms—protected environments where multiple parties can analyze combined data without ever exposing the raw datasets—have emerged as a solution. AWS Clean Rooms helps companies and their business partners securely analyze and collaborate on their datasets—without sharing or revealing the underlying data.

Connect with hundreds of data sources

Amazon AppFlow

Integrate data lakes and data warehouses with 50+ sources of data

AWS Data Exchange

Access 350+ third-party providers and 3,500+ public data products

Amazon SageMaker Data Wrangler

Build ML models with 40+ data sources with a single click

Data strategy tip



Pursue Zero-ETL

Avoid building and maintaining complex data pipelines to perform ETL operations. Instead, Amazon Aurora now supports zero-ETL integration with Amazon Redshift, to enable near real-time analytics and ML. Analyze data from multiple Aurora database clusters in the same new or existing Amazon Redshift instance and derive holistic insights across many applications.

3 GOVERNED

Free your teams to move faster with governed data access across the data lifecycle

Beyond being comprehensive and integrated, it's equally important to ensure that your users can access data where and when it is needed with the right level of control. With the right data governance strategy in place, you can move faster to empower users with the data access they need—when they need it.

As more data migrates to the cloud, driven by the cloud's near-infinite scale and horsepower, it's imperative that enterprise data governance models evolve in lockstep. IT and business leaders need up-to-date policies to protect data as it moves back and forth among different repositories and to accommodate changing privacy and data security regulations about where data can be stored.⁸

“

The key to good governance is figuring out how to define access, then getting out of the way. By that I mean creating exception processes, rather than taking an approach that anytime you need data, you have to ask someone in a central organization for it.

Rahul Pathak, VP for Analytics, AWS

⁸ "A unified approach to data governance," CIO Magazine, August 2021

Simplifying data access permissions

Implementing a successful governance strategy continues to present a unique set of challenges. It's time-consuming and challenging for organizations to provide internal or external consumers with their data with the right level of access to specific datasets. They often engage in heavy lifting, such as manual scripts or investigating individual data clusters, to figure out which consumers have access to what data.

Manual work can also lead to costly data quality issues across different teams and departments. Without centralized governance tools, data gets locked down in siloes, which means you won't be able to access and analyze all the data you may need to solve problems or identify large areas of opportunity.

Govern holistically with AWS

AWS Lake Formation

Makes it easy to govern and audit the actions taken with data in your data lake on Amazon S3

Amazon DataZone

A data management service to catalog, discover, share, and govern data

Developing a data governance strategy

A new AWS/MIT survey of more than 350 data professionals shows that data governance is the top priority of chief data officers (CDOs), with more than 50 percent of CDOs noting “establishing clear and effective data governance” as their leading responsibility. Governance is also an area CDOs spend much of their time on, as more than 66 percent of survey respondents said data governance initiatives are a top focus.⁹

Without a governance approach that supports innovation, organizations will find it hard to be data-driven and, ultimately, to remain competitive. After all, the more time workers spend grappling with data, the less time they spend innovating with it.

AWS is investing across the data journey to enable end-to-end data governance with less effort. AWS Lake Formation makes it easy to govern and audit the actions taken with data in your data lake on AWS S3 and AWS Lake Formation can also be used to govern data sharing in Amazon Redshift. Amazon DataZone is a new data management service to catalog, discover, share, and govern data so that everyone in the organization can act on data. And for your ML models, Amazon SageMaker has features to help you govern and audit the end-to-end ML development cycle.

9 “The Best Offense Is a Great Defense,” TechCrunch Brand Studio, by AWS, 2022

MAKING SECURITY MORE STRATEGIC

AWS has prioritized security since day one—with continuously protected, high-performing, resilient, and efficient infrastructure for your workloads and applications. World-class security experts who monitor the AWS infrastructure also build and maintain a broad selection of innovative security services—which can help simplify the complexities of your own security and regulatory requirements. AWS Security services and solutions can enable:

Getting to insights faster

Provide the right level of access to your resources at all times.

Reducing downtime

Tougher, more modern cloud security to keep your enterprise moving, so you don't have to stop analyzing data to perform security processes.

Staying within budget

AWS keeps security cost-effective and scales with the evolving needs of your security risks and requirements.

Keeping your focus

From infrastructure to services, AWS is secure by taking security into account at every step along the way, so you can spend more time transforming data into better decisions that drive business results and less time worrying about security and governance.

ABOUT TESCHGLOBAL

TESCHGlobal is dedicated to meeting your support needs.

Managing enterprise technology is time-consuming, requires deep technical knowledge, and depends on specialized staff who are not always available to organizations. TESCHGlobal's US-based team offers managed services to give organizations a comprehensive and operational managed IT service. We call it Flare Support.

TG's Support provides:

1. Access to Online Support Portal for ticket management
2. Dedicated Customer Support Personnel
3. Defined Protocols for Response Time, Update Frequency, and Support Mode for all support levels (Mission Critical, Urgent, High, Medium, Low)
4. Clear and Concise Support Guidelines

What is the benefit of working with TESCHGlobal?

Through our partnership with AWS, we are able to deliver value to our customers that is attentive, responsive, thorough, and timely. Committed to excellence, TESCHGlobal strives to provide a better understanding to our customers so they can feel confidence in their data configurations and beyond.