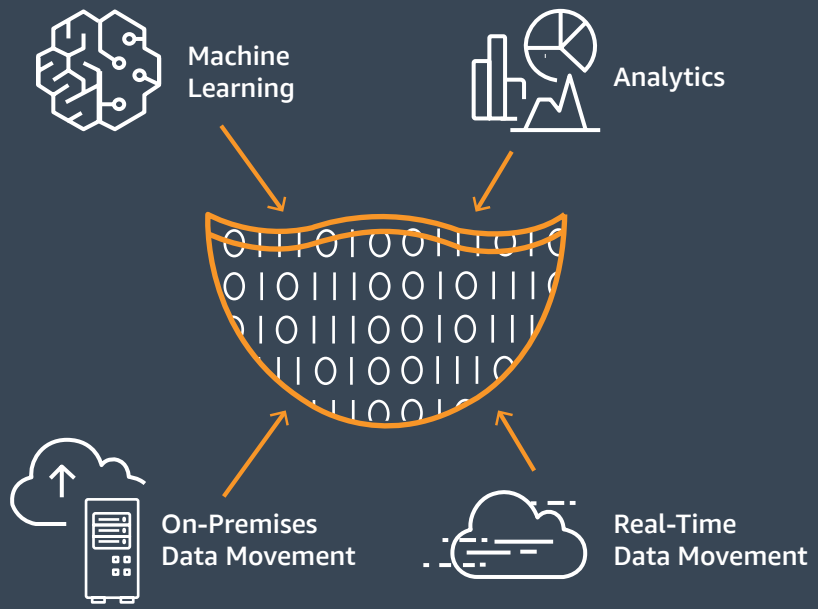
Harness the power of your data:

Why Financial Services institutions are building **data lakes** on AWS

aws financial services

What is a **data lake**?

A data lake is a centralized repository that allows you to store all structured and unstructured data at scale and run flexible analytics such as dashboards, visualizations, big data processing, real-time analytics, and machine learning, to guide better decisions.



Financial institutions are collecting and storing massive amounts of data

The Financial Services industry has relied on traditional data infrastructures for decades, but traditional data solutions can't keep up with the volumes and variety of data financial institutions are collecting today.

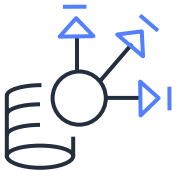
A cloud-based data lake helps financial institutions store all of their data in one central repository, making it easy to support compliance priorities, realize cost efficiencies, perform forecasts, execute risk assessments, better understand customer behavior, and drive innovation.

AWS delivers an integrated suite of services that provides the capabilities needed to quickly build and manage a secure data lake that is ready for analysis and the application of machine learning. In this overview, learn how financial institutions are unlocking the value of their data by building data lakes on AWS.

There are many benefits to building a **data lake on AWS**

Compliance & Security

Encrypt highly sensitive data and enable controls for data access, auditability, and lineage.



Scalability

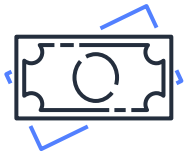
Amazon S3 data lakes allow any type of data to be stored at any scale, making it easy to meet variable data requirements.

Agility

Perform ad-hoc and cost-effective analytics on a per-query basis without moving data from the data lake.

Innovation

Aggregated and normalized data sets provide a foundation for advanced analytics and machine learning.



Cost effective

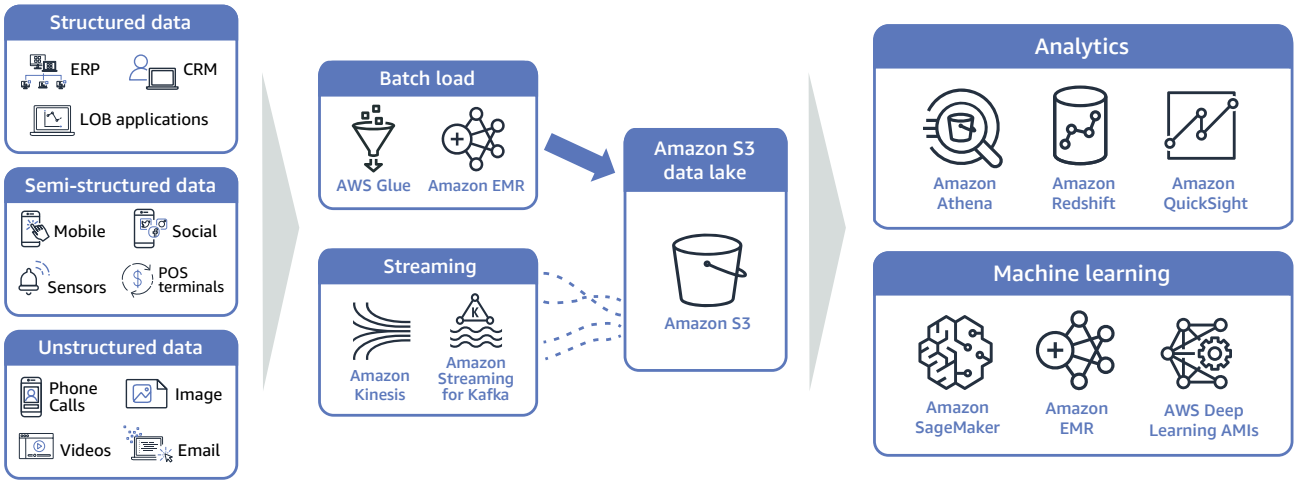
Pay-as-you-go pricing for compute, storage, and analytics.

Amazon S3 data lakes are built for security

Amazon S3 is the only cloud storage platform that allows you to apply access, log, and audit policies at the account and object level. S3 provides automatic server-side encryption, encryption with keys managed by **AWS Key Management Service (KMS)**, and encryption with keys that you manage. S3 also encrypts data in transit when replicating across regions.

Data lakes built on Amazon S3 provide the foundation for analytics and innovation

Unlike a data warehouse that can only store structured data, data lakes built on Amazon S3 are designed to collect and aggregate any type of data from multiple sources, eliminating the need for fragmented data silos and making it easy to share data. The data is stored as-is and is ready for any use case—from traditional analytics such as business intelligence to real-time streaming analytics and machine learning. Amazon S3 also integrates with AWS ingestion and analytics services, so you can easily move data into S3 and perform ad-hoc and highly cost-effective analytics on a per-query basis without moving data from the data lake.



Amazon S3

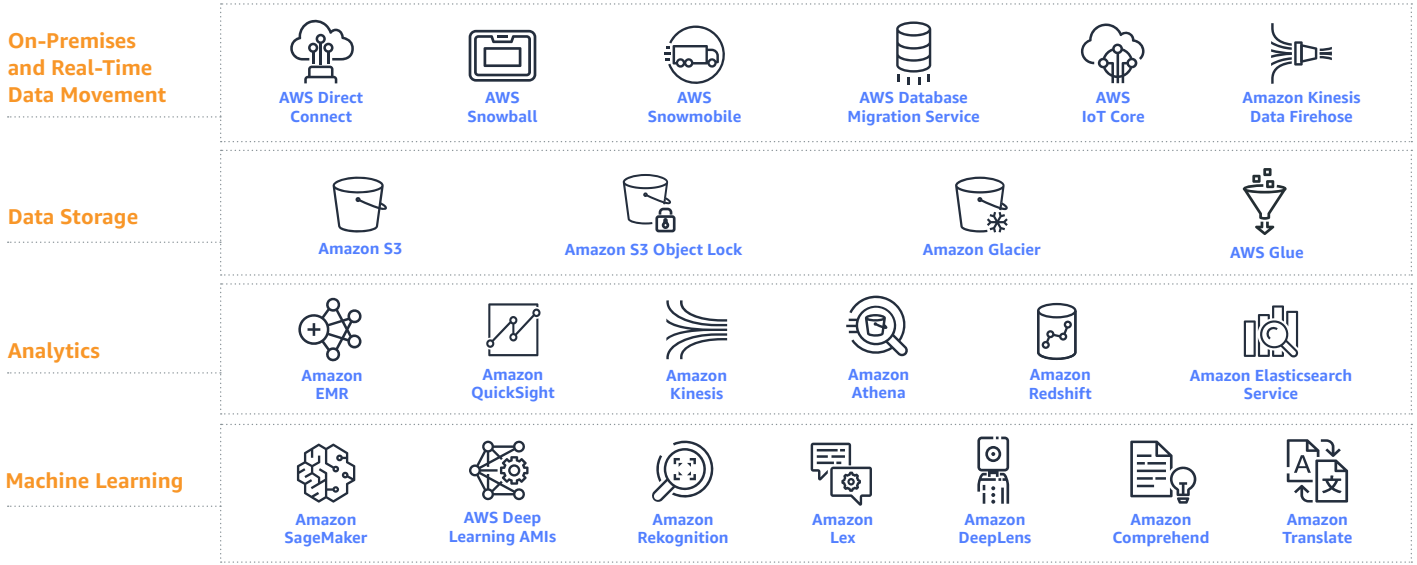
The foundation of an AWS data lake

- Designed for 99.99999999999% (11 9's) durability
- Store objects in a WORM-compliant environment to meet regulatory requirements such as SEC Rule 17a-4(f), FINRA Rule 4511, and CFTC Regulation 1.31
- Retrieve only a subset of data from an object with S3 Select, leading to performance increases of most applications that access data from S3 by up to 400%
- Glacier Select allows queries to run directly on data stored in Amazon S3 Glacier without having to retrieve the entire archive

AWS services integrate to help you realize the true value of your data

AWS provides the most comprehensive, secure, scalable, and cost-effective portfolio of services to enable you to quickly and easily build and manage a data lake for analytics.

AWS-powered data lakes can handle the scale of today's data volumes while providing the agility and flexibility required to combine different types of data while meeting data lineage and auditability requirements.



AWS Lake Formation

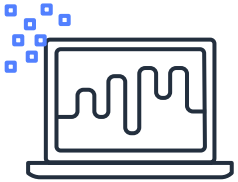
Enables you to build a secure data lake in days

- Collects and catalogs data from databases and object storage
- Moves the data into your new Amazon S3 data lake
- Cleans and classifies data using machine learning algorithms
- Secures access to your sensitive data
- Gain new insights into your data with AWS analytics and machine learning services

Data lakes are helping solve challenges across the industry

Compliance & Regulatory Reporting

Using an AWS data lake makes it easy to collect, store, and analyze data, while providing data lineage and auditability and enabling compliance with regulations such as Anti-Money Laundering (AML) and Consolidated Audit Trail (CAT).



Business Analytics

Aggregating and analyzing data across business lines enables you to gain a more holistic view of the business, identify market trends and opportunities, and detect fraud.

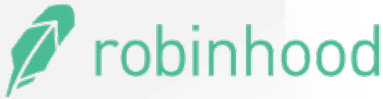
Customer Analytics

Storing and analyzing customer data in an AWS data lake enables you to mine deeper customer insights, recommend tailored products and services, and create a personalized and enhanced customer experience.

Markets Surveillance & Trading

The scale and agility of AWS data lakes makes it easy to aggregate data from multiple sources and conduct large-scale data analytics such as backtesting thousands of trading strategies and monitoring the markets to ensure market integrity.

Industry-leading financial institutions are building data lakes on AWS



Select customer stories

Capital One wanted to leverage machine learning capabilities to provide better fraud detection services for its customers. The bank chose to build a data lake on Amazon S3, enabling it to store and analyze large volumes of data.

Using Amazon S3 means the bank is better able to detect and prevent fraud in real time. When suspicious activity occurs, Capital One automatically alerts customers and walks them through how to report instances of fraud.

National Australia Bank (NAB) built its Data Hub data lake to power "Discovery Cloud," a laboratory for the bank's data scientists.

By building its data lake on AWS, NAB is able to provide full data lineage, access the data in real-time via APIs, and analyze the data using a wide range of AWS or third-party services.

Select customer stories

Guardian Life wanted to build a platform to expand its digital experience and gain new insights into its customers. The insurer built a data lake on AWS using Amazon S3 and Amazon EMR to support its anticipated data growth and analytics strategy.

By moving to AWS, Guardian launched an all-digital platform, Guardian Direct, that allows consumers to research and purchase both Guardian products and third-party products in the Insurance sector.

Mastercard acquired **NuData Security** to improve its fraud prevention techniques by using passive biometrics to authenticate account holders' identities. NuData uses an Amazon S3 data lake to store customer data that it collects and analyzes in real time.

By using AWS, NuData is able to aggregate, anonymize, and analyze petabytes of customer data to detect anomalous behavior patterns and protect customers from fraud.

Select customer stories

Robinhood needed a centralized data platform to aggregate information from various data stores across accounting, compliance, brokerage, and the business.

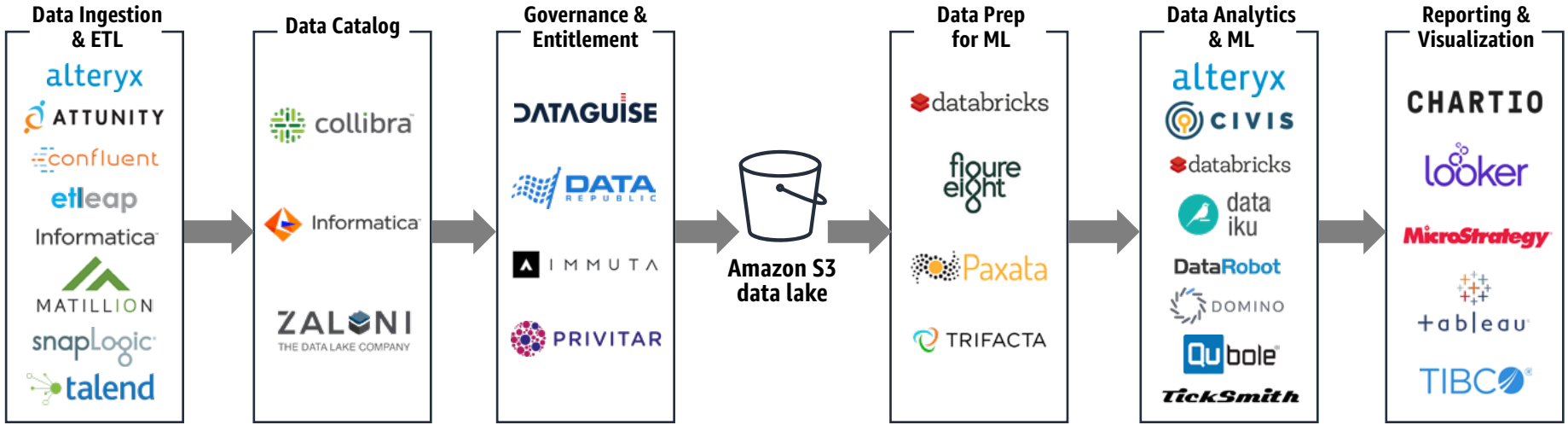
Robinhood was able to build its data lake on Amazon S3 with only three engineers, allowing other team members to focus on developing new products. Using AWS also made it easy for the company to scale its compute and storage and manage user access and governance.

Nasdaq needed to provide greater accessibility to data for both internal users and regulators.

By building a data lake on AWS, Nasdaq is able to move an average of 30 billion rows into the cloud everyday (with 60 billion on a peak day), while fulfilling security and regulatory requirements and realizing cost efficiencies.

AWS Competency Partners are here to help

The AWS Partner Network (APN) is a global community of Technology and Consulting partners. APN Competency Partners have demonstrated success helping customers collect, store, govern, and analyze data at any scale.



Ready to get started?

AWS Partner Network (APN) and AWS Marketplace

APN Partners are focused on your success, helping customers take full advantage of all the business benefits that AWS has to offer. Solutions in the AWS Marketplace can provide financial institutions with a competitive edge, and help to improve and transform the customer experience, manage and mitigate compliance risks, and drive innovation by modernizing infrastructure.

For more information, visit https://aws.amazon.com/solutionspace/financial-services and https://aws.amazon.com/marketplace/solutions/financial-services.

AWS Professional Services

AWS Professional Services is a global team of experts that can guide you through the steps to building your data lake. Work with the Professional Services team to set up an AWS Data Lake Workshop, AWS Data Lake Assessment, or AWS Data Lake Accelerator.

For more information, visit https://aws.amazon.com/professional-services.

Helpful Resources

Data lake and analytics solutions for Financial Services »

Data lake storage on AWS »

Contact AWS sales to talk about your Financial Services needs »

Learn more about AWS and try AWS for free »

