**What is a data lake?**

A data lake is a virtual or physical centralized repository that allows you to migrate and store all structured and unstructured data at unlimited scale, as well as gain insights through big data analytics and machine learning tools.

**Unprecedented data growth from new and diverse sources**

- Data creation will grow to 163 zettabytes (ZB) by 2025.

**Challenges to gaining the full value of your data**

- Unprecedented data growth from new and diverse sources.
- Break down data silos.

**Build your Data Lake on Amazon S3**

- Amazon S3 is the largest and most performant storage service for structured and unstructured data, allowing you to cost-effectively build and scale a data lake of any size in a secure environment.

**Amazon S3 data lake lifecycle**

- Ingest and store data (migrate data from a variety of data sources, real-time data movement, remove silos with one data lake for structured and unstructured data, unmatched scale, durability, security, and performance).
- Catalog and transform data (know your data with better management and higher quality data, AWS Glue crawls, catalogs, and indexes data for searchability, AWS Glue automates the effort in building, maintaining and running ETL jobs, Amazon Athena, Amazon Redshift, Amazon SageMaker, Amazon Rekognition, Amazon FSx for Lustre, Amazon EMR).
- Analyze (run AWS analytics and machine learning services to gain insights).
- Extract value from data (improve customer interactions, guide R&D innovation choices, maximize operational efficiencies).

**Why Amazon S3 is the best place to build a data lake**

- Cost-effective storage that scales on demand.
- Data management features with object-level controls.
- Best security, compliance and audit capabilities.
- Most options to migrate data to the cloud.
- Connect to Amazon FSx for Lustre to quickly process file system data.
- Automated data lake setup with AWS Lake Formation.
- Most durable, performant cloud storage service.

- Amazon S3 is designed for 99.99999999999% (11 9s) of durability.
- Unlimited scalability and high performance.
- Works seamlessly with analytics and machine learning services.

- Why Amazon S3 is the storage service of choice to build a data lake.

- Get started building your data lake on Amazon S3.

---

**Key components of a data lake**

- 10,000+ data lakes on Amazon S3.
- Ingests structured and unstructured data.
- Stores, secures and protects data at unlimited scale.
- Catalogs and indexes for analysis without data movement.
- Connects data with analytics and machine learning tools.

---

**Break down data silos. Build your data lake on Amazon S3.**

- 10,000+ data lakes on Amazon S3.
- Ingests structured and unstructured data.
- Stores, secures and protects data at unlimited scale.
- Catalogs and indexes for analysis without data movement.
- Connects data with analytics and machine learning tools.

---

**Why Amazon S3 is the best place to build a data lake**

- Cost-effective storage that scales on demand.
- Data management features with object-level controls.
- Best security, compliance and audit capabilities.
- Most options to migrate data to the cloud.
- Connect to Amazon FSx for Lustre to quickly process file system data.
- Automated data lake setup with AWS Lake Formation.
- Most durable, performant cloud storage service.

- Amazon S3 is designed for 99.99999999999% (11 9s) of durability.
- Unlimited scalability and high performance.
- Works seamlessly with analytics and machine learning services.

---

**Get started building your data lake on Amazon S3**

[Link to Amazon S3 website](https://aws.amazon.com/products/storage/data-lake-storage/)

---

**Amazon S3 lake formation**

- Amazon S3 Lake Formation allows you to automate the effort in building, maintaining and running ETL jobs.
- Amazon S3 Lake Formation connects data with analytics and machine learning tools.

---

**What?**

- Break down data silos.
- Build your data lake on Amazon S3.

---

**Why?**

- Amazon S3 Lake Formation allows you to automate the effort in building, maintaining and running ETL jobs.
- Amazon S3 Lake Formation connects data with analytics and machine learning tools.

---

**Get started building your data lake on Amazon S3**

[Link to Amazon S3 website](https://aws.amazon.com/products/storage/data-lake-storage/)

---

**Key components of a data lake**

- Ingest and store data.
- Catalog and transform data.
- Analyze.
- Extract value from data.

---

**Build your Data Lake on Amazon S3**

- 10,000+ data lakes on Amazon S3.
- Ingests structured and unstructured data.
- Stores, secures and protects data at unlimited scale.
- Catalogs and indexes for analysis without data movement.
- Connects data with analytics and machine learning tools.

---

**Unprecedented data growth from new and diverse sources**

- Data creation will grow to 163 zettabytes (ZB) by 2025.

---

**Challenges to gaining the full value of your data**

- Unprecedented data growth from new and diverse sources.
- Break down data silos.

---

**Amazon S3 lake formation**

- Amazon S3 Lake Formation allows you to automate the effort in building, maintaining and running ETL jobs.
- Amazon S3 Lake Formation connects data with analytics and machine learning tools.

---

**Get started building your data lake on Amazon S3**

[Link to Amazon S3 website](https://aws.amazon.com/products/storage/data-lake-storage/)