



What is Data Management and Why is it Important









In the 21st century, data is everything. With massive volumes of it generated every day, it stands to reason that we need to have better data management solutions available. Any business or organization that wants to succeed today needs to understand the what, why, and how of data management.

Fortunately, there are lots of resources available, from data management software to data management best practices, and everything in between. Let's begin with learning what data management is.

What is Data Management?

The Data Management Association or DAMA, defines data management as "the development of architectures, policies, practices, and procedures to manage the data lifecycle." To put it in simpler, everyday terms, data management is the process of collecting, keeping, and using data in a costeffective, secure, and efficient manner. Data management helps people, organizations, and connected things optimize data usage to make better-informed

decisions that yield maximum benefit.

Quantifying Data Management Principles

There are a handful of guiding principles involved in data management. Some of them may have higher weight than others, depending on the organization involved and the type of data they work with. The principles are:







- Creating, accessing, and regularly updating data across diverse data tiers
- Storing data both on-premises and across multiple clouds
- Providing both high availability and rapid disaster recovery
- Using data in an increasing number of algorithms, analytics, and applications
- Ensuring effective data privacy and data security
- Archiving and destroying data in compliance with established retention schedules and compliance guidelines

Data Management Best Practices

Data scientists face many challenges when setting up a successful, viable data management system. These best practices offer ways to address those obstacles and make it easier to implement an effective data management system.

- Identify your data by creating a discovery layer. Putting a discovery layer over your organization's data tiers enables data scientists and analysts to search and browse for useful datasets.
- Develop a data science environment to repurpose your data more efficiently. Data science environments automate a significant amount of activities. This practice brings in tools that remove the need for manual data transformation, making it easier to conduct testing.
- Maintain performance levels across your growing datasets by using autonomous technology. Bring in AI and machine learning methods to continuously monitor database queries and optimize indexes when those queries change. This practice maintains rapid performance and eliminates the need to perform time-consuming manual tasks.
- Stay ahead of compliance requirements by using discovery. Compliance demands are always increasing, so it's smart to use new data discovery tools to review data, including detecting, tracking, and monitoring your data wherever it resides.
- Manage and integrate multiple data storage platforms with a common query layer. By employing a standard query layer that spans the many kinds of data storage, you can access data centrally no matter where it resides or what format it is in.







Data Management Processes and Plans

We can also break down data management into five distinct processes. Not every organization uses each method. Like the principles, it depends on the business or organization in question:

Cloud Data Management

This process integrates data organization's from an of collection cloud applications. Cloud data management's defining characteristic is that all data storage, intake, and processing occurs within a cloud-based storage medium.

ETL and Data Integration

Extract, Transform and Load data from different sources into a centralized data warehouse.

DataAnalyticsandVisualization

Processes data from multiple data sources and data warehouses then perform advanced data analytics. This enables analysts and data scientists to present the data in visualizations and dashboards.

Master Data Management

Manageandstandardizeorganizationaldata(customers, employees, etc.) topreventredundancyand

duplication of effort across the organization.

Reference Data Management

Define a permissible value that can, in turn, be used by other data fields, such as postal codes, product serial numbers, and lists of cities, regions, and countries. Reference data can be created in-house or externally provided.







Alternately, data management can be understood as a combination of any of these disciplines:

- Business Intelligence and Analytics
- Data Architecture
- Data Governance and Data Stewardship
- Data Integration
- Data Modeling
- Data Quality
- Data Security
- Data Warehousing
- Data Storage and Big Data
- Document and Content Management
- Master and Reference Data Management
- Metadata Management

What is a Data Management Strategy?

Since data is so huge today, organizations need a sound data management strategy that works with the massive amounts being generated. Three critical components of a good data management strategy include:

Data Delivery

Making consistent а and accurate set of data or insights and conclusions drawn from analysis of that the data available stakeholders, to customers both within and outside of the organization.

Data Governance

Developing processes and best practices regarding the availability, integrity, security, and usability of the organization's data.

Data Operations

Also called DataOps, this involves implementing agile methods to design, deploy, and manage applications on a distributed architecture. Like DevOps, this also means removing the barriers between development and Its operations teams to improve the entire data lifecycle.

These three practices taken together will result in better data quality, more robust data security, and a better quality of data-driven insights for making more informed business decisions.







Data Management Platforms and Programs

There are several different data management systems available, including:

Document databases ER model databases Graph databases Hierarchical databases Network databases NoSQL databases Object-oriented databases Relational database

Data management platforms and data management programs are two indispensable management tools.

Data management platforms, also called DMP, are platforms that valuable store data like customer data (e.g., mobile identifiers, cookie IDs, etc.), and campaign data. DMPs help advertisers and marketing professionals build customer segments. The segments grow based on demographics, browsing history, geographical location, device type used, and other factors.

Here is a list of some popular DMPs:

Salesforce DMP Lotame Cloudera Nielsen SAS Data Management







And here are some of the better data management programs available today:

Matillion: Facilitates cloud data warehouse operations such as loading and transforming data





SolarWinds Backup: Ideal for backup and recovery

Panoply: A cloud data management tool that collects, sorts, combines, stores, and optimizes data without the need for data coding or modeling



Segment :

Segment: Collects data from the web and mobile apps and makes the information readily available to your teams

Tableau: Analyzes big dataand quickly translates it intoactionable insights. Ideal foranalytics and visualization







Collibra

Collibra: Automates workflows deliver userfriendly code, compares data from different parts of your business, and performs accurate data mapping.

Dell Boomi: A master data management tool that enables data stewarding defines models, governs data and deploy data models.



Data From: A SQL-based data transformation platform that manages cloud data warehouses' processes. It runs updating schedules to keep data current and ensure data reliability by creating data quality tests.

Stitch Data: A cloud-based ETL platform that's preintegrated with dozens of data sources, facilitating the movement of data. It features error handling and alerting, easy scheduling, and automatic scaling.

Stitch









Amazon Web Services: This well-known cloud provider offers a growing set of tools ideal for cloud data management.

Microsoft Azure: Another well-known cloud provider that offers cloud data management system tools and analytics.



Talend: An open-source data integration tool that enables users to cleanse, integrate, mask, and profile data, complete with MDM functionality and the ability to manage many source systems via a strong GUI.

talend







Data Management Platforms and Programs

modeling GData the İS determining practice of through extensive data analysis what is necessary to align business objectives with the information systems the business runs on. A data modeler documents complex software systems in easily understood diagrams for the of benefit non-technical These people. conceptual diagrams represent datasets and workflows in visual form and map them to the relevant line of business requirements and goals.

What Does Big Data Have to Do with Data Management?

In a word, everything! Big data, by its very nature, begs for a data robust management efficient data An system. management system takes big data and turns it into actionable items. It's а competitive world out there, and the businesses that stay ahead of the pack are the ones that make the best decisions, and the right information, in turn, creates the best decisions.

The above line of logic shows the importance of data in decision making, and the best way to achieve this is an alliance between big data and the right data management strategy.

modeling Common data techniques include entityrelationship diagrams, data mappings, and schemas. Note that data models must be updated whenever the organization brings in new data sources, or regular-ass updates occur, so this process is ongoing.

