# What is data management?

Data management is the practice of collecting, organizing, managing, and accessing data to support productivity, efficiency, and decision-making.



# Data management overview

Data is essential to how a business operates and functions. Businesses must make sense of data and find relevance in the noise that's created by diverse systems and technologies supporting today's highly connected global digital economies. In this regard, data takes center stage. On its own, data is useless – companies need an effective strategy, governance, and data management model to leverage all forms of data for practical and efficient use across supply chains, employee networks, customer and partner ecosystems ... and much more.

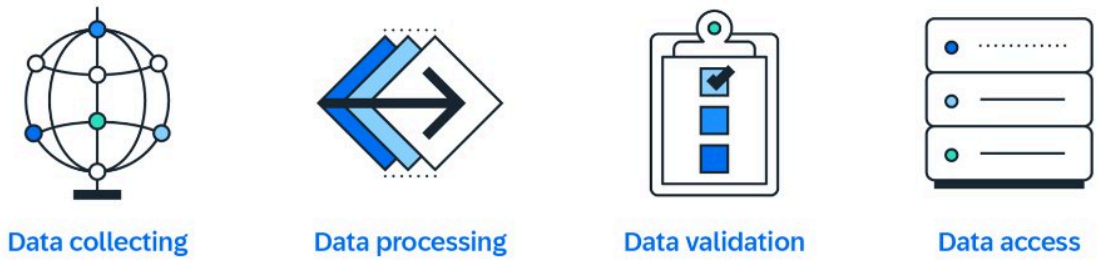
# Data management definition and process

Data management is the practice of collecting, organizing, managing, and accessing data to support productivity, efficiency, and decision-making. Given the pivotal role data plays in business today, an effective enterprise strategy, and a modern data management platform are essential for every company – regardless of size or industry. Data management is important for a variety of data-driven use cases including end-to-end business process execution, regulatory compliance, accurate analytics and AI, data migration, and digital transformation.

The data management process includes a wide range of tasks and pr

unstructured data

- Managing the quality of the data to adhere to business standards
- Providing self-service, collaboration, and access to data
- Protecting and securing data and ensuring data privacy
- Managing the lifecycle of data, from creation to deletion
- Ensuring high data availability and disaster recovery




The key elements of data management

# Why is data management important?

Every application, analytics solution, and algorithm used in a business (the rules and associated processes that allow technology to solve problems and complete tasks) depends on seamless access to high quality data. At its core, a data management system helps ensure data is secure, available, and accurate. But the benefits of data management don't end there.

Too much data can be overwhelming – and useless – if not managed properly. Add to this ever-growing levels of data diversity, distribution, and demands, and it's easy to see how organizations can struggle to leverage data as an asset to meet their digital business needs. But with the right tools, data can be harnessed to empower companies with deeper-than-ever insights, more accurate predictions and innovative business processes. It can give companies a better understanding of what customers want and help companies deliver exceptional customer experiences based on the learning data provides. It can also help drive new data-driven business models – such as service offerings based on generative AI – that wouldn't be accurate without a foundation of high quality data to base training models on.

***"Being data-driven means using data, whatever has to handle the complexity of state, store, access, quality and context to enable organizations to realize their data-driven aspirations, central to digital business success."***

Gartner "Data Management Solutions Primer for 2023." Adam Ronthal, Ehtisham Zaidi, February 14, 2023

It's no secret that data-driven organizations have a major competitive advantage. With advanced tools, companies can manage and access more data from more sources than ever before. They can also leverage many different types of data, structured and unstructured, in real time – including Internet of Things (IoT) device data, video and audio files, Internet clickstream data, and social media comments – opening more opportunities to monetize data and use it as an asset.

## Laying the data foundation for digital transformation

It's often said that data is the lifeblood of digital transformation – and it's true. Data and analytics (D&A) leaders must be able to meet the demands of digital business and the increasing complexity of the data landscape (including the impact of the cloud). AI, machine learning, Industry 4.0, advanced analytics, IoT, and intelligent automation all require high volumes of timely, accurate, and secure data to do what they do.

Machine learning and generative AI, for example, need very large and diverse data sets to "learn," identify complex patterns, solve problems, and keep models and algorithms

scenarios depend on a steady stream of machine and sensor data delivered at extremely high speeds.

The common denominator in any digital transformation project is data. Before businesses can transform processes, take advantage of new technologies, and become intelligent enterprises, they need a solid data foundation. In short, they need a modern data management system.

***"The continued survival of any business will depend upon an agile, data-centric architecture that responds to the constant rate of change."***

Donald Feinberg, Vice President at Gartner

## Ensuring compliance with data privacy laws

Good data management is also essential for ensuring compliance with national and international data privacy laws – like the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act in the United States – as well as industry-specific privacy and security requirements. And when those protections are to be proven or audited, having solid data management policies and procedures in place is essential.

# Data management approaches

Data fabric and data mesh have both become popular pieces of data management terminology in recent years, with many organizations adopting them for their data management architecture. The data fabric style of architecture is meant to create a connecting layer for various data sources, facilitating self-service, data access and data delivery across the enterprise. A data fabric architecture is meant to abstract away the different locations in which you store your data from an end user perspective, presenting

## Data mesh

Data mesh is an approach to data management that uses a distributed architectural framework. In other words, it spreads ownership and responsibility for specific data sets across the business, to those users who have the specialist expertise to understand what that data means and how to make the best use of it. Data mesh architecture connects and draws data from various sources like data lakes and warehouses and distributes the relevant data sets to the appropriate human experts and domain teams across the business. Essentially, a voluminous jumble of data in a central data lake is sorted and distributed into manageable chunks to those best suited to understand and leverage it.

## Data fabric

A data fabric is a combination of data architecture and dedicated software solutions that centralize, connect, manage, and govern data across different systems and applications. Data fabric solutions allow you to connect and manage data in real time, across different systems and applications. This makes it possible to create a single source of truth, and to use and access that data whenever and wherever you need it – democratizing and automating data management processes. A data fabric also streamlines data, especially in complex distributed architectures, making it ready for use in analytics, AI, and machine learning applications by unifying, cleansing, enriching, and securing it. Data fabric architecture and solutions allow businesses to leverage their data and scale their systems, while adapting to rapidly changing markets.

## Master data management (MDM)

Master data management is the discipline of creating one trusted master reference (a single version of the truth) for all important business data, such as product data, customer data, asset data, finance data, and more. MDM helps ensure businesses don't use multiple, potentially inconsistent versions of data in different parts of the business, including processes, operations, analytics, AI, and reporting. The three key pillars to effective MDM include: data consolidation, data governance, and data quality management.

*"A technology-enabled discipline in which business and the IT organization work together to ensure the uniformity, accuracy, stewardship, semantic consistency, and accountability of the enterprise's official, shared master data assets."*

Gartner definition of MDM

## Data integration

Data integration is the practice of ingesting, transforming, combining, and provisioning data, where and when it's needed. This integration takes place in the enterprise and beyond – across partners as well as third-party data sources and use cases – to meet the data consumption requirements of all applications and business processes. Techniques include bulk/batch data movement, extract, transform, load (ETL), change data capture, data replication, data virtualization, streaming data integration, data orchestration, and more.

Data discovery and data cataloguing enable you to find what data you have and determine how it is related. Discovery is often provided as part of data profiling (which itself exists to generate a bird's-eye view of your data in terms of its structure, content and so on), and indicates the whereabouts of and the relationships that exist between different datasets within and across multiple (heterogeneous) data sources. More broadly, it is a fundamental tool for understanding your data landscape. Sensitive data discovery is a notable subcategory that is particularly concerned with locating and classifying personal or otherwise sensitive data within your organization so that it can be appropriately protected for the sake of data privacy, security, and regulatory compliance.

Data discovery is also used to build data catalogues, along with other more advanced automation techniques like AI and machine learning. Data catalogs provide a repository of information (known as metadata) about your data assets: what data is held, where it is located, what format it is in, and within which domains it is relevant. As much of this information as possible should be collected automatically, and it may be classified further by geography, time, access control, and so on. Catalogues are indexed and searchable, and support self-service and collaboration. More comprehensive catalogues will ingest metadata from various derived sources, such as analytical reports and dashboards, in addition to the physical sources of your data. Catalogues are commonly used in conjunction with data preparation tools, and are important for supporting data governance and collaborative, self-service-based data access.

## Data governance, security, and compliance

Data governance is a collection of rules and responsibilities for ensuring data availability, quality, compliance, and security across the organization. Data governance establishes the infrastructure and names the individuals (or positions) within an organization that have both the authority and the responsibility for the handling and safeguarding of specific kinds and types of data. Data governance is a key part of compliance. The systems will take care of the mechanics of security, storage, and access as well as proper deletion and retention. Data governance also helps to ensure that data is accurate to begin with and that it meets business standards before being entered into the system, while being used, and when retrieved from the system for use or storage elsewhere. Organizationally governance specifies how responsible individuals use processes and technologies to manage and protect data.

Data security is a major concern in today's world of hackers, viruses, cyberattacks, and data breaches. While security is built into systems and applications, data governance is there to ensure that those systems are properly set up and administered to protect the

# What is an enterprise data strategy and why should you have one?

Today, business strategies depend on data to automate processes, customize customer and employee experiences, drive growth through new markets or acquisition, and to innovate. Thus, business success increasingly depends on aligning your data strategy to your business strategy. A data strategy should resonate across all levels of the organization. It needs to have meaning and context to the business.

A company needs a data strategy to prioritize its work. We all know that the amount of data a company generates and uses is growing significantly. There will always be more data issues and requirements than resources. Companies need a way to prioritize data activities based on what will realize the most value through the data strategy. The strategy needs to be “living and breathing” and fully aligned with business priorities yet flexible enough to shift as the business transforms and matures. It can’t just be words in a document but must take “life” within the organization.

A data strategy outlines all the data capabilities that have to be built to achieve the business outcome. That includes not just data management capabilities and tools, but business capabilities such as organizational structure, data acquisition and data network strategy, compliance and ethics capabilities. It lays out a roadmap to develop capabilities over multiple years, setting expectations on what can be delivered, in what time frame, for what cost and executive support that is required.

# The evolution of data management

Effective data management has been critical to business success for well over 50 years – from helping companies improve the accuracy of information reporting, spot trends, and make better decisions to fueling digital transformation and powering new technologies and business models today. Data has become a new kind of capital, and forward-thinking organizations are always on the lookout for new and better ways to use data to their advantage. Here are the latest trends in modern data management that are important to keep an eye on and explore their relevance to your business and industry:

- **Data fabric:** Most organizations today have a variety of types of data deployed on premise and in the cloud – and they use multiple database management systems, processing technologies, and tools. A data fabric, which is a custom combination of architecture and technology, uses metadata, dynamic data integration and

- **Data management in the cloud:** Many companies are moving some or all of their data management platform to the cloud. Cloud data management takes advantage of all the benefits cloud has to offer – including scalability, advanced data security, improved data access, automated backups and disaster recovery, cost savings, and more. Cloud databases and database-as-a-service (DBaaS) solutions, cloud data warehouses, and cloud data lakes are all growing in popularity.
- **Data as a product:** Data as a product refers to the practice of treating your internal data like a first-order product, with the job of your data team(s) – and by extension, your Chief Data Officer or equivalent executive – being to provide the rest of your organization with the right data that it needs at the right time and at the right level of quality. The goal is to enable greater utilization of your data in general - such as more timely and more accurate analytical insights, for instance.
- **Augmented data management:** One of the newer trends is called "augmented data management." Augmented data management uses AI and machine learning to make data management processes self-configuring and self-tuning. Augmented data management automates everything from data quality and master data management to data integration – freeing up skilled technical staff to focus on higher value activities.
- **Augmented analytics:** Augmented analytics uses artificial intelligence (AI) technologies, machine learning, and natural language processing (NLP) to not only find the most important insights automatically, but to democratize access to advanced analytics so everyone, not just data scientists, can ask questions of their data and get answers in a natural, conversational way.

Explore additional data management terms and trends.

# Summary

We know that information is derived from data. And if information is power, then effectively managing and capitalizing on your data could very well be your company's superpower. As such, data management responsibilities and the role of the Chief Data (and Analytics) Officer are evolving to become key change agents in the organization – in driving cloud adoption, leveraging new trends and technologies, and delivering strategic value to the business.

## Database and data management solutions

SAP can help you manage, govern, and integrate your enterprise data.

Learn more >

## Ideas you won't find anywhere else

Sign up for a dose of business intelligence delivered straight to your inbox.