What is data architecture? A data management blueprint

June 2024

THE R



How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development? Data architecture is a discipline that documents an organization's data assets, maps how data flows through IT systems and provides a blueprint for managing data. This guide to data architecture further explains what it is, why it's important and the business benefits it provides. You'll also find information on data architecture frameworks, best practices and more. Throughout the guide, there are hyperlinks to related articles that cover the various topics in more depth.



How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development?

What is data architecture? A data management blueprint

CRAIG STEDMAN, INDUSTRY EDITOR

Data architecture is a discipline that documents an organization's data assets, maps how data flows through IT systems and provides a blueprint for managing data. Its goal is to ensure that data is managed properly and meets business requirements for information used to drive decision-making.

While data architecture can support operational applications, it most prominently defines the underlying data environment for business intelligence (BI) and advanced analytics initiatives. Its output includes a multilayer framework for data platforms and <u>data management</u> tools, as well as specifications and standards for collecting, integrating, transforming and storing data.

Ideally, data architecture design is the first step in the data management process. But that often isn't the case, which creates inconsistent environments that need to be harmonized as part of a data architecture. Also, despite their foundational nature, data architectures aren't set in stone and must be updated as data and business needs change. That makes data architecture work an ongoing task for data management teams.



How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development? Data architecture goes hand in hand with <u>data modeling</u>, which creates diagrams of data structures, business rules and relationships between data elements. They're separate data management disciplines, though. Explaining at a high level <u>how data</u> <u>modeling and data architecture differ</u>, practitioners distinguish between modeling's microfocus on individual data assets and data architecture's broader macro perspective on all those assets.

This guide to data architecture further explains what it is, why it's important and the business benefits it provides. You'll also find information on data architecture frameworks, best practices and more. Throughout the guide, there are hyperlinks to related articles that cover the various topics in more depth.

HOW HAVE DATA ARCHITECTURES EVOLVED?

In the past, most data architectures were less complicated than they are now. They mostly involved structured data from transaction processing systems that was stored in relational databases. Analytics environments consisted of a <u>data warehouse</u>, sometimes along with smaller data marts built for individual business units and an operational data store as a staging area. The transaction data was processed for analysis in batch jobs, using traditional extract, transform and load (ETL) processes for data integration.



How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development? Starting in the mid-2000s, the adoption of <u>big data technologies in businesses</u> added unstructured and semistructured forms of data to many architectures. That led to the deployment of <u>data lakes</u>, which often store raw data in its native format instead of filtering and transforming it for analysis upfront -- a big change from the data warehousing process. More recently, <u>data lakehouses</u> that combine elements of data lakes and warehouses have emerged as another analytics platform. These new approaches have driven wider use of ELT data integration, a variation on ETL that inverts the load and transform steps.

The increased use of <u>stream processing</u> systems has also brought real-time data into more data architectures. Many now support artificial intelligence (AI), machine learning and other data science applications, too, in addition to the basic BI and reporting driven by data warehouses. The widespread shift to cloud-based systems further adds to the complexity of data architectures.

Another emerging architecture concept is the <u>data fabric</u>, which aims to automate data integration and management tasks through reusable processes. It has a variety of potential <u>use cases in data environments</u>. Even newer is data mesh, a decentralized architecture that gives individual business domains responsibility for managing their own data. Federated governance processes are used to create organization-wide data standards and policies.



How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development?

Sample data architecture diagram





How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development?

WHY ARE DATA ARCHITECTURES IMPORTANT?

A well-designed data architecture is a crucial part of the data management process. It supports data integration and data quality improvement efforts, as well as data engineering and data preparation. It also enables effective <u>data governance</u> and the development of internal data standards. Those two things, in turn, help organizations ensure that their data is accurate and consistent.

A data architecture is also the foundation of a <u>comprehensive data strategy</u> that supports business goals and priorities. Business strategies increasingly depend on data. As a result, data management and usage are too important to leave to individuals, according to Donald Farmer, principal of consultancy TreeHive Strategy. In addition to the data itself, he listed data catalogs, data management tools, various analytics techniques, collaboration capabilities and documented goals as <u>key data</u> <u>strategy components</u>. But that should all be underpinned by a strong data architecture.

Along with building an architecture, the main aspects of developing a data strategy include the following, as outlined by Donna Burbank, managing director of consulting firm Global Data Strategy:

• Identifying business goals that an organization's data assets must support.



How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development?

- Assessing the current state of data management processes and technologies.
- Proposing upgrades to the data management environment to meet business needs.
- Planning and communicating a roadmap for both the data architecture and data strategy.

Key stages in creating a data strategy

Consultant Donna Burbank broke down the data strategy development process into four major parts.



1. Identify strategic business goals and align planned data initiatives with them

2. Assess the current state and maturity of your data management environment



3. Propose new capabilities, processes and technologies to meet business needs



4. Plot out an implementation roadmap and an internal communication plan

ROFT ALL RIGHTS RESU

ILLUSTRATION: ALEXDNDZ, ADOBE STOCK



7

How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development?

CHARACTERISTICS AND COMPONENTS OF A MODERN DATA ARCHITECTURE

The <u>principles of modern data architectures</u>, also as cited by Farmer, include alignment with data governance and regulatory compliance processes; support for multi-cloud environments; and efficient deployments that avoid unneeded data platforms. A data architecture also needs to ensure that data is available for planned analytics uses. Otherwise, the data's potential business value will be wasted.

Other common characteristics of well-designed data architectures include the following:

- A business-driven focus that's aligned with organizational strategies and data requirements.
- Flexibility and scalability to enable various applications and meet new business needs for data.
- Strong security protections to prevent unauthorized data access and improper use of data.

From a purist's point of view, data architecture components don't include platforms, tools and other technologies. Instead, a data architecture is a conceptual infrastructure that's described by a set of diagrams and documents, which data management teams then use to guide technology deployments and how data is managed.

How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development? The following are some examples of those components, commonly referred to as *artifacts*:

- Data models, data definitions and common vocabularies for data elements.
- <u>Data flow diagrams</u> that illustrate how data flows through systems and applications.
- Documents that map data usage to business processes, such as a CRUD matrix -- short for create, read, update and delete.
- Other documents that describe business goals, concepts and functions to help align data management initiatives with them.
- Policies and standards that govern how data is collected, integrated, transformed and stored.
- A high-level architectural blueprint, with different layers for processes such as data ingestion, data integration and data storage.

If the technology elements are incorporated, a modern data architecture includes the ones previously mentioned in the evolution section plus some others, as listed here:

- Data warehouses, data lakes and data lakehouses.
- Cloud systems, storage and applications.
- Al and machine learning tools.
- Data streaming and real-time analytics systems.



How have data architectures evolved?

Why are data architectures important?

Characteristics and components of a modern data architecture

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development?

- Various data integration methods.
- API connectors to streamline data sharing between applications.
- Data pipelines that deliver needed data to users.
- Containerized and microservices applications.

Key data architecture principles to keep in mind

A modern data

architecture likely

environments.



Manage data storage systems carefully, despite the reduced cost of storage devices.

Deploy data repositories and analytics applications efficiently, without unneeded platforms.



Data governance and compliance will include multiple with privacy laws cloud platforms and are separate but related processes.



Data architectures should be optimized for analytics initiatives to generate business value.

ICONS: ALEXDNDZ/ADOBE STOCK



TechTarg

How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development?

WHAT ARE THE BENEFITS OF A DATA ARCHITECTURE?

A well-designed data architecture helps organizations <u>develop effective data</u> <u>analytics platforms</u> that deliver useful information and insights. Those insights improve strategic planning and operational decision-making, potentially leading to better business performance and competitive advantages. They also aid in other types of applications, such as scientific research, government programs and the diagnosis and treatment of medical conditions.

In addition, data architecture helps provide the following benefits in managing data:

- Improved data quality.
- Streamlined data integration.
- Reduced data storage costs.
- Increased data consistency across systems.
- More effective data governance.
- Better collaboration on data management and governance.

It does so by taking an enterprise view compared to domain-specific data modeling or focusing on architecture at the database level, according to Peter Aiken, a data management consultant and associate professor of information systems at Virginia Commonwealth University.



How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development? Data architecture brings "purposefulness" to efforts to ensure that data assets support business strategies, Aiken said during a Dataversity webinar in October 2023. "The architecture is the thing that you use as your playbook in trying to figure out how you're going to be using these [assets]."

WHAT ARE THE RISKS OF BAD DATA ARCHITECTURE DESIGN?

One data architecture pitfall is too much complexity. The dreaded "spaghetti architecture" is evidence of that, with a tangle of lines representing different data flows and point-to-point connections. The result is a ramshackle data environment with incompatible <u>data silos</u> that are hard to integrate for analytics uses. Ironically, data architecture projects often aim to bring order to messy environments that developed organically. But if not managed carefully, they can create similar problems.

Another challenge is getting universal agreement on standardized data definitions, formats and requirements. Without that, it's hard to create an effective data architecture. Putting data in a business context can also be challenging. Done well, data architecture captures the business meaning of data. But failing to do so can create a disconnect between the architecture and the strategic data requirements it's supposed to meet. As a result, data architecture work should have a practical focus,



How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development? Aiken said in the Dataversity webinar. "If it doesn't move the needle somewhere, it's probably not worth investing in."

DATA ARCHITECTURE VS. DATA MODELING

Data modeling focuses on the details of specific data assets. It creates a visual representation of data entities, their attributes and how different entities relate to each other. That helps in scoping the data requirements for applications and systems and then designing database structures for the data -- a process that's done through a progression of conceptual, logical and physical data models.

Data architecture looks at all of an organization's data to create a framework for managing and using it. But, as mentioned previously, data modeling and data architecture complement each other. Data models are a crucial element in data architectures, and an established data architecture simplifies data modeling, according to David Loshin, president and principal consultant at Knowledge Integrity Inc.

There are <u>various techniques for modeling data</u>. The most-used ones now are entityrelationship, dimensional and graph modeling approaches. The first two are variations of the relational data model that underpins relational databases, but they



How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development? can also be used to model other types of data. Graph data modeling is mainly used to map out relationships in graph databases that store data in graph-like structures.

Phases of data modeling



How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development? The following are some data modeling best practices:

- Gather both business and data requirements upfront, before building models.
- Develop data models iteratively and incrementally to make the process manageable.
- Use data models as a tool to communicate with business users about their needs.
- Manage data models just like any other type of application code.

DATA ARCHITECTURE VS. INFORMATION ARCHITECTURE AND ENTERPRISE ARCHITECTURE

While they sound somewhat similar, there's a <u>difference between data architecture</u> <u>and information architecture</u> in enterprise applications. It comes down to the basic difference between data and information: The latter is what data provides. As such, an information architecture defines the context for managing business operations and making decisions, including information workflows both inside an organization and with customers and business partners. A data architecture that delivers high-quality, reliable data is the foundation for the information architecture.

How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development? Meanwhile, data architecture is commonly viewed as a subset of <u>enterprise</u> <u>architecture</u> (EA), which aims to create an organizational blueprint in four domains. In addition to data, EA encompasses the following three areas:

- Business architecture, which involves business strategy and key business processes.
- Application architecture, which focuses on individual applications and their relationships to business processes.
- Technology architecture, which includes IT systems, networks and additional technologies that support the other three domains.

WHAT DATA ARCHITECTURE FRAMEWORKS ARE AVAILABLE?

Organizations can use standardized frameworks to design and implement data architectures instead of starting completely from scratch. These are three well-known framework options:

• DAMA-DMBOK2. DAMA-DMBOK: Data Management Body of Knowledge, as it's formally named, is a data management framework and reference guide created by DAMA International, a professional association for data managers. Now in its second edition and commonly known as DAMA-DMBOK2, the framework addresses data architecture along with other data management disciplines. The first edition was published in 2009. The second one became available in 2017 and was revised in 2024.



How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development?

- TOGAF. Created in 1995 and updated several times since then -- most recently in 2022 -- TOGAF is an enterprise architecture framework and methodology that includes a section on data architecture design and roadmap development. It was <u>developed by</u> The Open Group, and TOGAF initially stood for The Open Group Architecture Framework. But it's now referred to simply as the TOGAF Standard.
- The Zachman Framework. This is an ontology framework that uses a six-bysix matrix of rows and columns to describe an enterprise architecture, including data elements. It doesn't include an implementation methodology; instead, it's meant to serve as the basis for an architecture. The framework was originally developed in 1987 by John Zachman, an IBM executive who retired from the company in 1990 and founded a consulting firm named Zachman International that continues to oversee it.

KEY STEPS FOR CREATING A DATA ARCHITECTURE

Data management teams must work closely with business executives and other end users to develop a data architecture. If they don't, it might not be in tune with business strategies and data requirements. Engaging with senior execs to get their support and meeting with users to understand their data needs are two key <u>data</u> architecture planning steps.

How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development? Other steps to take include the following:

- Evaluate data risks based on data governance directives.
- Classify data sets according to their usage and how sensitive the data is.
- Track data flows, as well as data lifecycle and data lineage info.
- Document and appraise the existing data management technology infrastructure.
- Scope out a roadmap for the data architecture deployment projects.

The same steps apply in <u>building a cloud-based architecture</u> for data management and analytics, as organizations increasingly are doing. But data management teams face new potential challenges on architecture design in the cloud, including data security requirements, regulatory compliance mandates and data gravity issues that can complicate migrations of data sets from on-premises systems.

WHAT ARE THE DIFFERENT ROLES IN DATA ARCHITECTURE DESIGN AND DEVELOPMENT?

Not surprisingly, <u>data architects</u> typically play the lead role in data architecture initiatives. They need a variety of technical skills plus the ability to interact and communicate with business users. A data architect spends a lot of time working with



How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development? end users to document business processes and existing data usage, as well as new data requirements.

On the technical side, data architects create data models themselves and supervise modeling work by others. They also build data architecture blueprints, data flow diagrams and other artifacts. Other duties can involve outlining data integration processes and overseeing the development of data definitions, business glossaries and data catalogs. In some organizations, data architects also are responsible for designing data platforms and evaluating and selecting data management technologies.

Read more about the data architecture profession

Top data architect and data engineer certifications

Data architect skills required, responsibilities and salaries

Other <u>data management professionals</u> who often are involved in the data architecture process include the following:

• **Data modelers.** They also work with business users to assess data needs and review business processes. Then, they use the information they've gathered to create data models.



How have data architectures evolved?

Why are data architectures important?

<u>Characteristics and</u> <u>components of a modern</u> <u>data architecture</u>

What are the benefits of a data architecture?

What are the risks of bad data architecture design?

Data architecture vs. data modeling

Data architecture vs. information architecture and enterprise architecture

What data architecture frameworks are available?

Key steps for creating a data architecture

What are the different roles in data architecture design and development?

- **Data integration developers.** Once the architecture is implemented, they're tasked with creating ETL and ELT jobs to integrate data sets.
- **Data engineers.** They build the pipelines that funnel data to data scientists and other analysts. They also help data science teams with the <u>data preparation</u> process.

Craig Stedman is an industry editor who creates in-depth packages of content on analytics, data management, cybersecurity and other technology areas for TechTarget Editorial.

CONTINUED READING

- Principles of a well-designed data architecture
- Data modeling vs. data architecture: What's the difference?
- Data architecture vs. information architecture: How they differ