

# DATA ARCHITECTURE VS INFORMATION ARCHITECTURE: WHAT'S THE DIFFERENCE?



There's a well-known argument around [data architecture](#) versus information architecture. And the question often asked is: *Are they the same thing?*

Enterprise architect and Microsoft blog contributor, Nick Malik, recognized the inherent confusion when he was part of a group working to clean up the Wikipedia entries on the subjects. His team believed the entries should be combined. However, in 2014, when he [polled the IT community](#) he soon discovered a split audience, where about half of all survey participants believed the two should remain separate.

Let's take a look at the differences between data and information and the key considerations your enterprise organization needs to understand.

## Data Architecture vs Information Architecture

This author agrees that information architecture and data architecture represent two distinctly different entities. There are a couple of reasons for this as described below:

### Distinction in Data vs Information

Simply put, data refers to raw, unorganized facts. Think of [data as bundles of bulk entries](#) gathered

and stored without context. Once context has been attributed to the data by stringing two or more pieces together in a meaningful way, it becomes information.

Similarly, it's also important to understand the difference as it regards infrastructure:

- Information architecture refers to the development of programs designed to input, store, and analyze meaningful information.
- Data architecture is the development of programs that interpret and store data.

## Distinction in Architecture

Since we've established that data and information are not the same, it stands to reason that they can't be treated the same way in their architecture platforms.

Data architecture is foundational. It looks at incoming data and determines how it's captured, stored, and integrated into other platforms. One such platform is likely a piece of information architecture, like a CRM, that uses raw customer data to draw meaningful connections about sales and sales processes.

The CRM is the information architecture in this example because it specializes in taking raw data and transforming it into something useful.

That's the clear distinction between data architecture and information architecture. Data architecture defines the collection, storage and movement of data across an organization while information architecture interprets the individual data points into meaningful, useable information.

An "[information asset](#)" is the name given to data that has been converted into information. And creating information assets is the driving purpose of information architecture. Information assets can exist in one of several categories:

- Catalogues
- Dashboards
- Documents
- Ontologies
- Schedules
- Taxonomies
- Templates
- Terminologies

Each category suggests the conversion of data into something that is helpful for business initiatives, whether it be a grouping of like data or a visual representation that can offer a meaningful snapshot of data to stakeholders.

## Data and Information Lifecycle Management

Another distinction relates to requirements from a lifecycle management perspective. Besides the obvious difference between data and information, each has a unique lifecycle and best practices for managing it within an organization. Similar to how data infrastructure is at the foundation of solid information infrastructure, proper data lifecycle management will be a key driver of the information lifecycle management process.

Now, let's dive into some more definitions.

[Data lifecycle management](#) refers to the automated processes that push data from one stage to the next throughout its useful life until it ultimately becomes obsolete and is deleted from a database. On the other hand, information lifecycle management looks at questions like whether or not a piece of data is useful, and if yes, how? In a nutshell, information lifecycle management seeks to take raw data and implement it in a relevant way to form information assets.

In addition, information assets have their own lifecycle and value, which are determined by the quality and usefulness of data involved as well as the type of asset as described above. Part of the information lifecycle process requires developers to consider future state implementations.

For instance, making recommendations that a piece of data could be better implemented as a dashboard or document attachment. This may be required to improve overall consumption of knowledge throughout an organization, democratize information or create more meaningful insights.

## Data-Driven Business Models

More and more, IT departments are becoming an integral part of the enterprise business model. Gone are the days when IT departments were ancillary to process. Now, the vast majority of departments and processes are powered by IT innovation.

A study by the [University of Cambridge](#) suggests that increasingly businesses are creating new models to accommodate a commitment to data and information. And results show that this approach is paying off, offering increases in productivity over competitors.

The report suggests that when coming up with a new business model, enterprise business leaders ask themselves these questions:

- What is our target outcome for a data-driven business model?
- What would we like to offer our target market?
- What software, hardware and services do we require to deliver on this model?
- Where are we going to acquire these resources?
- How will collected data be used?
- How can this be monetized to support a revenue model?
- What challenges will we face in accomplishing these goals?

But even after a data-driven model has been created, [some companies fail](#) because they don't understand the importance of a workflow that pushes data through the lifecycle and through the process of becoming an information asset.

Establishing best practices and a workflow in your data and information life cycles provides the following benefits:

- Improves overall speed to market
- Greatly reduces the complexity between all cloud environments
- Readily scalable
- Helps mitigate risk
- Improves integration

In order to achieve this, companies should look at how they can [integrate, automate, and orchestrate these workflows](#). Application Workflow Orchestration solutions such as [Control-M](#), help organizations to abstract the complexity involved with the numerous data sources, multiple

applications and diverse infrastructure. It help organizations to focus on creating new information assets and delivering insights to the business, rather than spending precious time and efforts on fixing broken workflows.

Still, with all things considered, enterprise businesses must have the right IT employees in place to create a functional business model. Below is an employee snapshot created for both information architecture and data architecture.

## Employee Snapshots: Information vs Data

At the heart of a well-functioning enterprise business is an IT department with the right people in place to manage their information and data architectures. In the following text, we will look at positions that may be necessary for data architecture, information architecture or both.

### Chief Information Officer (CIO)

The CIO of an enterprise organization makes important decisions about technology and innovation, and is central to any [digital transformation](#) or shift toward IT in enterprise business model.

Some responsibilities in this role include innovating, integrating cloud environments, motivating the IT department and establishing an IT budget based on projected needs. The CIO will make decisions regarding both data and information architecture. As it regards data architecture, one of the big considerations will be deciding between a data lake and a data warehouse. More on these points later.

(Compare [CIOs to CTOs](#).)

### Information Architect

The information architect is integral to information architecture and automated lifecycle management processes. He or she will implement information structure, features, functionality, UI and more. The primary role of the information architect is to focus on structural design and implementation of an infrastructure for processing information assets.

### Data Architect

Like an information architect, data architects work on the structural design of an infrastructure but in this case it's specific to collecting data, pulling it through a lifecycle and pushing it into other meaningful systems.

### Data Analyst

The [data analyst's](#) typical day involves the gathering, retrieval and organization of data from various sources to create valuable information assets. This is someone who likely works in both systems comprised of data architecture and information architecture.

More and more, some functions of the data analyst are being automated, but even with automation, analysts remain important to the creation of future information states.

# Information Analyst

Information analysts specialize in the extraction and analysis of information assets.

## A quick note: Data lakes vs data warehouses

Data lakes have been rising in popularity these days but are still confused with data warehouse. However, it's important to realize that these two have unique differences and are used in different ways. A data warehouse refers to a large store of data accumulated from a wide range of sources within an organization.

- A warehouse is used to guide management decisions.
- A data lake is a storage repository or a storage bank that holds a huge amount of [raw \(unstructured\) data](#) in its original form until it's needed.

*(Read more about the differences in [data lakes & warehouses](#).)*

## Final Thoughts: Data Architecture vs Information Architecture

Hopefully by now, it's clear why information and data architecture are two different things. If not, here's a quick recap.

Data and information architecture have distinctly different qualities:

- They work with different assets: data assets vs information assets
- They yield different results
- They have distinctly unique life cycles
- They require different things from an architecture perspective
- They require roles with different specialties to be part of an enterprise organization

Although data and information architecture are unique, an important takeaway is that they rely on each other in order for enterprise organizations to gain the insights they need to make the most informed business decisions.

## Related reading

- [BMC Machine Learning & Big Data Blog](#)
- [Data Ethics for Companies](#)
- [3 Keys to Building Resilient Data Pipelines](#)
- [Data Management vs Data Governance: An Introduction](#)
- [What Is Goodhart's Law? Balancing Authenticity & Measurement](#)
- [What Is Data Gravity?](#)