

The demand for data engineers and data architects is higher than ever. A report by [Opinium, in collaboration with the UK's Department for Digital, Culture, Media and Sport, shows](#) UK companies alone are currently recruiting for 178,000 to 234,000 roles that require hard data skills. And as we continue to generate and use data to enhance critical business decisions, the demand for data professionals will continue to increase.

The data architect and data engineer roles are sometimes used interchangeably. Although there is some overlap, and they share some specific duties, understanding how both roles operate in the organization can benefit engineering managers looking to build an engineering team and students and professionals looking to develop a career in either field.

- [What is a Data Architect?](#)
- [What is a Data Engineer?](#)
- [Data Architect vs Data Engineer: What are the Differences?](#)
- [Data Architect vs Data Engineer: Which One is Right for You?](#)

## What Is a Data Architect? The Role and Its Responsibilities



A data architect is responsible for formulating the organizational data strategy and defining the data management standards and principles on which the organization operates. Data architects design the “data blueprint” that other data consumers follow and implement.

They create the organization’s logical and physical data assets and set data policies based on company requirements. Data architects are often veterans in the industry who have had experience in many data roles and have gained experience navigating complex business scenarios and designing solutions that data teams can implement.

### Data architect responsibilities

The data architect’s primary responsibility revolves around providing deep technical expertise for designing, creating, managing, and deploying large-scale data systems in the organization.

A data architect’s responsibilities include:

- Designing, developing, implementing, and translating business requirements and the overall organizational data strategy, including standards, principles, data sources, storage, pipelines, data flow, and data security policies
- Collaborating with data engineers, data scientists, and other stakeholders to execute the data strategy
- Communicating and defining data architecture patterns in the organization that guide the data framework
- Leading data teams to develop secure, scalable, high-performance, and reliable big data and analytics software and services

### What skills do data architects have?

Data architects need a combination of technical and soft skills to thrive. Typically, data architects start in other data roles, such as data scientist, data analyst, or data engineer, and work their way up to becoming data architects after years of experience with data modeling, data design, and data management. Some of the skills data architects have include:

- **Data modeling, integration, design, and data management:** Data architects understand the concepts, principles, and

implementation of data modeling, design, and data management. They can produce relevant data models and design the organization's data flow.

- **Databases and operating systems:** Data architects are experienced in the various SQL and NoSQL databases. They understand the advantages and drawbacks of each and how they can be set up effectively and securely across different operating systems (Linux, Windows) and environments (development and production).
- **Data architecture:** Data architects know the best practices on data architecture for enterprise data warehouse development. They have a solid understanding of the organization's data infrastructure and how different systems interact.
- **Data security and governance:** Data architects understand the processes, roles, policies, standards, and metrics that ensure the effective and efficient use of data/information. Data architects are skilled in data governance strategies with a good understanding of the risks and mitigations of each. They ensure data governance is in line with the overall organizational strategy.
- **Communication and leadership:** Data architects are usually leaders of the data management team; as such, they must communicate clear technical solutions to complex data problems to both technical and non-technical audiences.

Although every organization has slightly different requirements, you'll notice similar skills and common themes (like the ones outlined above) throughout job descriptions for this role. For example, [this data solutions architect role](#) with Lightspeed asks that candidates have experience with data management for software as a service (SaaS) tools and build data solutions and models for this team.

## What Is a Data Engineer? The Role and Its Responsibilities



A data engineer is responsible for designing, maintaining, and optimizing data infrastructure for data collection, management, transformation, and access. Data engineers create pipelines that convert raw data into usable formats for data scientists and other data consumers to utilize.

The data engineer role evolved to handle the core data aspects of software engineering and data science; they use software engineering principles to develop algorithms that automate the data flow process. They also collaborate with data scientists to build machine learning and analytics infrastructure, from testing to deployment.

### Data engineer responsibilities

The primary responsibility of a data engineer is ensuring that data is readily available, secure, and accessible to stakeholders when they need it.

A data engineer's responsibilities typically include:

- Building and maintaining data infrastructure for the optimal extraction, transformation, and loading of data from a wide variety of sources such as Amazon Web Services (AWS) and Google Cloud big data platforms
- Ensuring data accessibility at all times and implementing company data policies with respect to data privacy and confidentiality
- Cleaning and wrangling data from primary and secondary sources into formats that can be efficiently utilized by data scientists and other data consumers
- Collaborating with engineering teams, data scientists, and other stakeholders to understand how data can be leveraged to meet business needs

### What skills do data engineers have?

Data engineering is a synthesis of software engineering and data science, so knowledge of both fields is advantageous. Because data engineering is heavily reliant on programming, most data engineers begin their careers as software engineers and then pivot to data engineering.

Some of the skills required of data engineers include:

- **Database systems (SQL and NoSQL):** Data engineers have a good knowledge of SQL and NoSQL databases and are skilled

in writing queries to manipulate and retrieve data.

- **Data migration and integration:** Data engineers are often tasked with aggregating data from multiple sources and migrating data from one platform to another based on business needs. Data engineers understand data migration and integration techniques (Big-Bang, trickle, lift and shift )and the tools required to perform them. [Striim](#) is a popular tool used by data engineers for data integration and migration; it provides modern, reliable data integration and migration across the public and private cloud.
- **Data wrangling:** Data wrangling is the process of gathering, cleaning, enriching, and transforming data into the desired format to incorporate better decision-making in less time. A data engineer is skilled in various data wrangling techniques and tools, such as extraction, transformation, and loading (ETL).
- **Data processing techniques and tools:** Data engineers are experienced in various data processing techniques, such as real-time processing and batch processing; they are comfortable working with data processing tools such as [Apache Kafka](#) and [Apache Spark](#).
- **Programming languages (Python):** Data engineering relies heavily on programming; data engineers are typically fluent in at least one programming language, with [Python](#) being regarded as the most popular and widely used programming language in the data engineering community.
- **Cloud computing and distributed systems:** With more companies relying on cloud providers for data infrastructure needs, companies rely on data engineers to create data solutions using popular cloud providers, such as Amazon Web Services, Google Cloud, and Azure. Data engineers have experience working with tools such as [Hadoop](#) for the distributed processing of large datasets.

Although each organization's requirements are slightly different, you'll notice similar skills and common themes (such as the ones outlined above) throughout the job descriptions for data engineer roles. For example, the candidate [for this data engineer, infrastructure engineering](#) role at TikTok will be responsible for collaborating with software engineers and data scientists to build big data solutions.

## Data Architect vs. Data Engineer: What Are the Differences?

The data architect and data engineer titles are closely related and, as such, frequently confused. The difference in both roles lies in their primary responsibilities.

- Data architects design the vision and blueprint of the organization's data framework, while the data engineer is responsible for creating that vision.
- Data architects provide technical expertise and guide data teams on bringing business requirements to life; data engineers ensure data is readily available, secure, and accessible to stakeholders (data scientists, data analysts) when they need it.
- Data architects have substantial experience in data modeling, data integration, and data design and are often experienced in other data roles; data engineers have a strong foundation in programming with software engineering experience.
- The data architect and the data engineer work together to build the organization's data system.

Here's a table to briefly summarize the key differences and help you visualize the contrast in responsibilities:

Data Architect	Data Engineer
Visualizes the blueprint for the organizational data framework, defining how the data will be stored, consumed, integrated, and managed by different data entities and IT systems	Builds and maintains the data systems and information specified by data architects in the data framework
Deep expertise in databases, data modeling, data architecture, and operating systems	Strong background in software engineering, algorithms, and application development
Focused on leadership and high-level data strategy	Focused on the day-to-day tasks of cleaning, wrangling, and preparing data for other data consumers, such as data scientists

## Data Architect vs. Data Engineer: Which Is Right for You?

When choosing a career or hiring, note that the data architecture role requires years of experience in a previous data-related role; both roles require a deep understanding of database systems, data processing tools, and experience working with big data. To put together an effective data management team, you must first understand the differences between the roles.

When interviewing for data engineers, consider whether the candidate has experience building software and APIs, as well as a solid understanding of various databases, data wrangling, and data processing techniques.

For data architects, on the other hand, be sure to ask what data projects they've led in the past and get a sense of their "data philosophy." Remember that a data architect will be the leader of your data management team to whom you should feel confident delegating authority.