

File systems are typically used to store arbitrary data as files on disks, with some form of metadata (such as a filename) used to identify and locate these files. File systems offer rudimentary search and indexing capabilities, and searching the contents of files to find the information you need is often an involved and laborious process. When the data is amenable to structuring, it is typical to organize it in a **database**, using a well-defined model (also known as a **schema**). The following video provides an overview of databases:

A database consists of an organized collection of data for one or more uses, typically in digital form. Databases are one of the most popular information storage and retrieval methods for many types of applications, including financial, manufacturing, business, and personnel data.

Databases have evolved over the years, starting with the **navigational model** in the 1960s, which stored data as records with pointers that linked to the next and/or previous record in the database. Examples include the [CODASYL](#) approach, which eventually evolved into the Common Business Oriented Language ([COBOL](#)). The CODASYL approach was based on the "manual" navigation of a linked data set, which was formed into a large "network" of linked data, much like a circular linked list. When the database was first opened by a program, the program was handed back a link to the first record in the database, which also contained pointers to other pieces of data. To find any particular record, the programmer had to step through these pointers one at a time until the required record was returned. Simple queries that aggregate information across records required a complete scan of the entire database—concepts such as **search** and **indexing** did not exist then. The CODASYL approach was appropriate in an era when magnetic tapes were the primary mode of nonvolatile storage, which could be read only sequentially and had no random access properties.

The emergence of the **relational database model** and the subsequent development of IBM's System R were a major milestone in database system development. For many decades, relational database systems were the only databases in wide use, and were extremely successful. With the growth of the internet and the need for massively scalable systems that can service millions of simultaneous requests, **NoSQL** database systems emerged. However, of late, a class of relational database systems has emerged, re-architected from scratch to try to combine the strengths of the two models.

Types of database systems

Current database systems can be classified into the following types:

Relational databases are the classical database systems that have been around for decades. They follow the relational model of representing data as relations between entities. Relational database systems typically have a rigid schema, use the Structured Query Language (SQL) as the interface to interact with the database, and typically have strong ACID properties.

NoSQL databases are a new breed of databases that are markedly different from relational database systems. NoSQL systems typically do not enforce a strong schema or relational model of the data. They provide a different abstraction (like a key-value store), instead of a SQL-like interface. NoSQL databases are focused on scalability and performance, and typically do not have strong transactional guarantees for database operations.

NewSQL is a class of modern relational database systems that seek to provide the same level of performance and scalability as NoSQL systems while maintaining some degree of ACID properties that relational database systems offer.

Next unit: Database design: Schema and scalability

[Continue](#)

Need help? See our [troubleshooting guide](#) or provide specific feedback by [reporting an issue](#).