

Pyramid of life and computation



- **Data** – Data can be facts related to any object in consideration. It refers simply to raw facts and figures. Alone it tells you nothing. Data can be of two types – qualitative and quantitative. The real goal is to turn data into information.
- **Information** - Data becomes information when it's presented in a context so that it can answer a question or support decision making. To turn data into information we use various tools and software, this includes – database management systems, spreadsheets.
- **Knowledge** - Knowledge is information that has been retained with an understanding about the significance of that information. Knowledge includes something gained by experience, study, familiarity, association, awareness, and/or comprehension. Only relevant information can become knowledge.
- **Wisdom** – Wisdom is specialized knowledge that places constraints on knowledge. Not that it limits how knowledge performs but selects which knowledge applies to a situation or set of data. Wisdom is a selection process. Wisdom like all forms of data, fails over time as new data changes context in the models and in the world, it needs to evolve to remain relevant.

In conclusion these four elements make up the cognitive portions of our technological and biological existence.

Database

A database is a collection of inter-related data which is used to retrieve, insert and delete data efficiently. It is used to organise data in the form of a table, schema, views and reports.



Database Properties

- Managing large amount of data is a hassle, a database provides data integrity through systematic collection of data.
- Database assures accuracy and consistency.
- Databases allow restricted access.
- Databases ensure the security of data.
- Database allow for flexibility of updating.

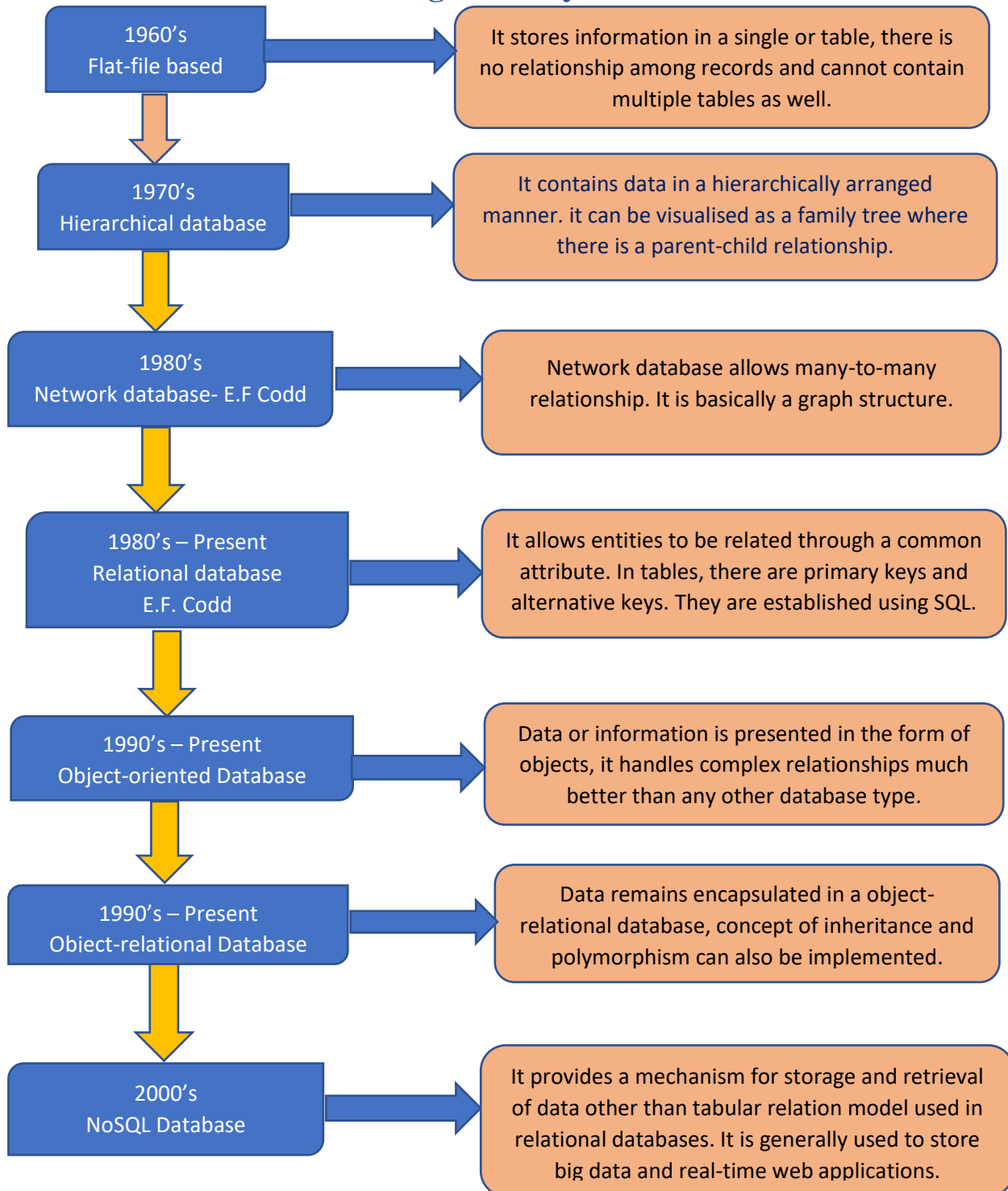
Database Management System (DBMS)

A DBMS is a software which is used to manage the database. It provides an interface to perform various operations like database creation, storing data, updating data. It provides protection and security to the database and also maintains data consistency. It also provides backup, multiple user interface and controls database redundancy(duplicity).

Charles Bachman developed the first DBMS named Integrated Data Store (IDS).



Evolution of Database Management Systems



Database terms

- **Query** – a type of command that retrieves data from the server.
- **Command** – a command is a string that you send to the server in hopes of having the server do something useful.
- **Column (Attribute, field)** – A column is the smallest unit of storage in a relational database. It represents one piece of information about an object. Columns are grouped into rows and rows are grouped into tables.
- **Row (Record, Tuple)** – A row is a collection of column values. Every row in a table has the same shape i.e., every row is composed of the same set of columns. Row represents a real-world object.
- **Table (Relation, file, class, entity, instance)** – A table is a collection of rows.
- **Relational database** – they were first proposed by E.F. Codd, it is a database which employs relational model in which data is arranged into sets of rows and columns and they are organised into relations(tables). It stores and provides access to data points that are related to one another.

This relational model imposes structure on its contents, in contrast to semi-structured or unstructured data of NoSQL architectures. It is a simple, intuitive and straightforward way of representing data in tables.

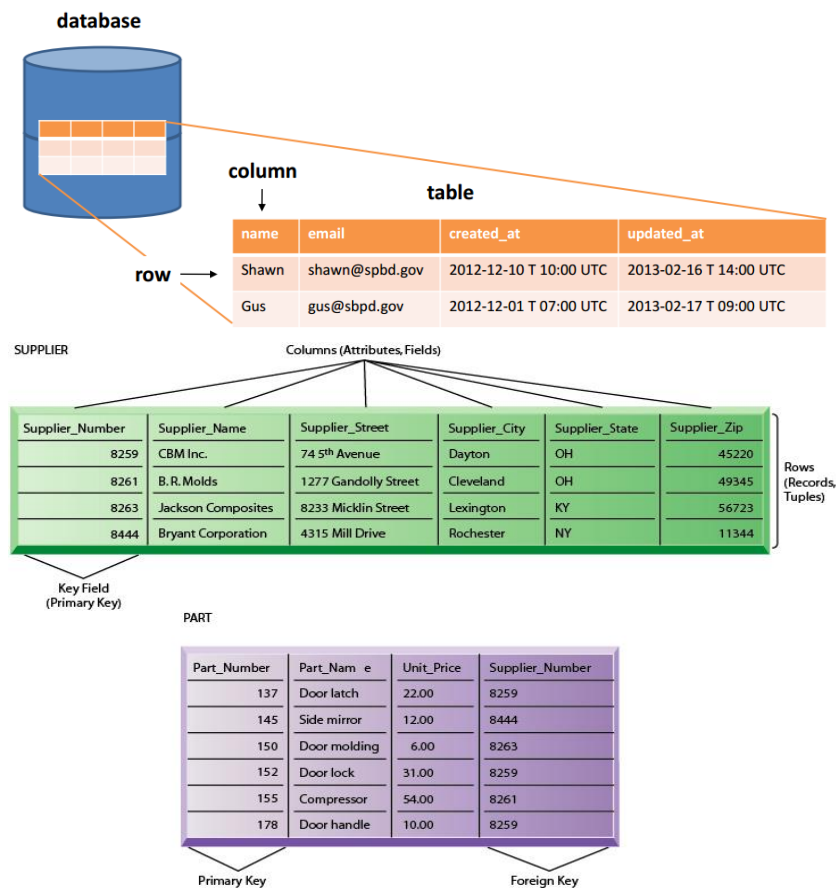
- **Primary key** – in relational model, a primary key is a single attribute or a combination of attributes that can be used to uniquely identify a row of data in a given table. A table cannot have more than one primary key.
- **Foreign key** – a reference to a primary key of another existing table.
- **Set** – A set is a collection of objects that are different.
- **Cartesian product** – A set that is constructed from two given sets and comprises all pairs of elements such that the first elements of the pair is from the first set and the second is from a second set.
- **Graph database** – a graph database is premised on edges acting as relationships, directly relating data to one another. They are useful for data mining and pattern recognition.
- **Constraints** – they are used to specify rules for data in table.
- **SQL (Structured Query Language)** - SQL is Structured Query Language, which is a computer language for storing, manipulating and retrieving data stored in a relational database.

Dremio

The only data lake-house platform built for SQL and built on open-source technology that provides a self-service experience that makes data consumable and collaborative.



Relational model



SQL

SQL is Structured Query Language, which is a computer language for storing, manipulating and retrieving data stored in a relational database.

SQL is the standard language for Relational Database System. All the Relational Database Management Systems (RDMS) like MySQL, Microsoft SQL server, Oracle, Sybase, Informix, Postgres and SQL Server use SQL as their standard database language.

SQL functions as a

- DDL – Data Definition Language (**CREATE, ALTER, DROP**)
- DML – Data Manipulation Language (**INSERT, UPDATE, DELETE**)
- DCL – Data Control Language (**GRANT, REVOKE**).
- DQL – Data Query Language (**SELECT**)
- TCL – Transaction Control Language (**SAVE POINT, ROLL BACK, COMMIT**)

Database Keys

A key refers to an attribute or a set of attributes that help us identify a row (tuple) uniquely in a table (relation). A key is also used when we want to establish relationships between different columns and tables of a relational database.

