

: *What is a database?*

by Dan Power

Editor, DSSResources.com

Database is a fundamental concept for information technology professionals. While some lay people have an intuitive understanding of the term, many would have difficulty defining and explaining the concept of a database. At one time "data base" was two words used to describe a computerized "place" for storing data. Over the past 25 years, the single word "database" has become the common term to refer to computerized software for storing and retrieving data. A specific database is usually an organized collection of "related" data. One hopes all of the stored data is logically related, relevant to the data storage purpose and accurate.

In a broad sense, any structured collection of data can be called a database. From this perspective, a card catalog, a website, a printed catalog or a telephone directory are each examples of databases. More commonly a database refers to storing data using computing technology (cf., Gahan, 2000).

Data refers to the "encoding" that represents events or occurrences that can be captured or recorded electronically. We especially try to capture needed data for the who, what, when, and where facts about people, places, things, and events. Physical data is stored using an encoding system called ASCII, for American Standard Code for Information Interchange. ASCII code is the numerical representation of a character, cf., <http://www.asciitable.com/>. Data in a database is manipulated by a computer program called a database management system (DBMS) to select and display specific data. A database is like an electronic filing cabinet with multiple component files organized to contain data. "Data" is both a singular and a plural word. Prior to about 1980, data was considered as the plural of "datum", but the term datum has become largely obsolete in technology discussions.

According to McGee, around 1964 the term "data base" was "coined by workers in military information systems to denote collections of data shared by end-users of time sharing computer systems." Also, in the early days of computerized storage, we focused on structured data that could be easily captured and organized. More recently we have also wanted to store less structured and unstructured data like video clips, photographs, maps and entire documents.

A database may be created for storing relevant personal or job specific data, or a shared department or group database may be created. In some cases we can create a more comprehensive database that includes subjects and data relevant to many or all people in an organization or very large diverse group. These types of databases are called personal, departmental and enterprise. Databases vary in physical size, topics, purpose, data and complexity.

: *What is a database?*

Patterns do exist in how data is stored for similar purposes which aids in designing new databases.

A database is more than data it also has an organizing framework or structure. The structure may be rigid and formal or more unstructured. Transactional or operational data is typically organized for a specific purpose and task. The data model usually reflects a real situation like hotel room availability. The database is then used to make a reservation or to find a hotel with vacancies depending upon the scope of the data.

Operational databases provide detailed information and summaries of transactions that occur in an organization. Each transaction is an event that we want to record. Often we want to archive data from operational databases. Sometimes we keep the transactions history in a new database called a decision support data warehouse. In the data warehouse, transaction data can be combined with data about external events and other factual data to create information that improves our understanding of a decision situation.

Data is often stored in a single file, but that limits how one can organize the data or encourages duplication of data in multiple files. Today, the most common model for organizing data is called a relational model. Data is organized in multiple files that are somehow related. Then the related files can be combined using a computer program to sort, retrieve, and display data in many different, yet related, ways without duplicating the data that is stored. From an Information Technology perspective, the term "database" refers to the data itself and associated data structures for organizing the data. A database is a structured set of data stored in a computer-based system.

A database management system (DBMS) refers to software that operates databases, providing storage, user access, security, backup and other data management facilities. Database management systems are often categorized according to the database model that they support, such as a relational data model is supported by a relational database management system (RDBMS).

Database textbooks variously define these key terms. For example, Hoffer et al. define data as "stored representations of objects and events that have meaning and importance in the user's environment" and database as "an organized collection of logically related data". Rob and Coronel define data as "raw facts. The word 'raw' is used to indicate that the facts have not yet been processed to reveal their meaning" and they define database as "a shared, integrated computer structure that houses a collection of: End user data, that is, raw facts of interest to the end user and Metadata, or data about data, through which the data are integrated (p. 5)". Kroenke defines database using technical jargon as "a self-describing collection of integrated records (p. 206)". Date (2003) defines a database as "a collection of persistent data that is used by the application systems of a given enterprise". A database is called by various authors a data bank, data mart, data store

: *What is a database?*

and a data warehouse.

Well-known database management systems (DBMS) include IBM DB2, Microsoft Access, Microsoft SQL Server, MySQL, Oracle, and SAP HANA.

Fact-based decisions require accurate, timely information or what is called "good" information. We need to collect, store, analyze, process and retrieve data from our databases to provide good information to decision makers.

References

Codd, E.F. "A Relational Model of Data for Large Shared Data Banks". Communications of the ACM, Volume 13, Number 6, 1970, pp. 377–387.

Date, C.J. An Introduction to Database Systems, Addison-Wesley (8th edition), 2003.

Gahan, M. An introduction to databases, October 18, 2000 at URL
<http://www.ucl.ac.uk/archaeology/cisp/database/manual/node1.html> .

Hoffer, J. A., M. B. Prescott, and F. R. McFadden. Modern Database Management (8th edition), Upper Saddle River, NJ: Pearson Prentice Hall, 2007.

Kroenke, D. M. Database Concepts, Upper Saddle River, NJ: Pearson Prentice Hall, 2003.

McGee, W.C. "Data Base Technology," IBM Journal of Research and Development, Volume 25, Number 5, 1981, pp. 505-519.

Rob, P. and C. Coronel. Database Systems, Cambridge, MA: Course Technology, 2000.

: *What is a database?*

The Linux Information Project. "Database Definition," 2006 at <http://www.linfo.org/database.html>, last accessed August 5, 2014.

Citation:

Power, D. J. "What is a database?" DSS News, Vol. 11, No. 23, November 21, 2010.

revised 08/05/2014

Author: Daniel Power

Last update: 2017-07-09 01:36