

3.1 Data representation

3.1.2 File organization and access

Computers can store large volumes of data. The difficulty is to be able to get it back. In order to be able to retrieve data, it must be stored in some sort of order.

Take the phone book for example it contains large volumes of data which can be accessed fairly easily because all the records are stored in alphabetical order. Imagine how difficult it would be to find an entry if the records had just been placed in the book at random. The worth of the book is not just that it contains all the data that may be needed. But that it has a structure that makes it easily accessible. Similarly, the structure of the data in a computer file is just as important as the data that it contains. There are a number of ways of arranging the data that will aid access under different circumstances.

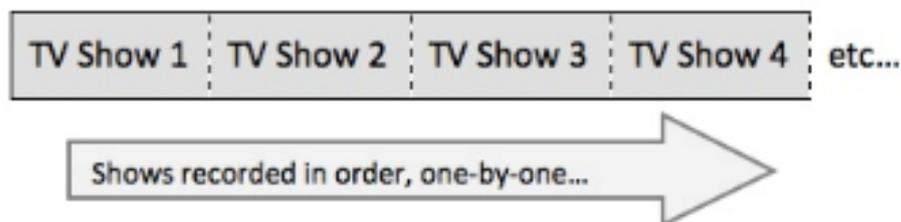
3 methods are:

- Serial
- Sequential
- Random (or Direct)

Serial Organization

Data is stored in the computer in the order in which it arrives (chronological order). This is the simplest form of storage, but the data is much unstructured, so finding a data item again can be very difficult. This sort of data storage is only used when it is unlikely that the data will be needed again, or when the order of the data is not critical. A good example of a serial file is the document you are reading right now. The characters were all typed in, in a random pattern less order. Reading this document would be close to impossible if all the words were in alphabetic order!

Another example of serial organization is shows that come on TV.



When thinking of storage systems, one could presume that all of your data in one folder is located next to each other within the hard drive.

3.1 Data representation



3.1.2 File organization and access

This is false when talking about **random access**. With random access, your information can be pulled from any location on the disk. Meaning, your one folder could have its data scattered about the physical hard drive. The benefit of this type of storage is that you could access data in any order.

Think of it as your CD player, your favorite song ends and you want to hear it again just hit the back button and done. It's fast and nearly instantaneous, unlike sequential.

You could think of **sequential access** like a cassette within a cassette player. When a song finishes and you want to listen to it again, you must rewind the cassette, or if you want to skip a song, you must fast forward the tape. This is used with magnetic tape drives which are used for backup purposes.

Even though in random access media devices it may seem like data could be misplaced or somehow lost in the sea of data. When created, every file is given a unique name by the computer system, otherwise referred to as addressable media, in order to keep track of all the data so it can be accessed later.

Random access and sequential access of data are two separate ways a computer can access data. Random access is the ability to access data in any given location within the hard drive, quickly and efficiently. Most computers use random access today because it saves the user time.

Sequential access requires data being accessed in a sequence. Examples of sequential access would be:

- Data on a disk file
- Magnetic tape disk storage.

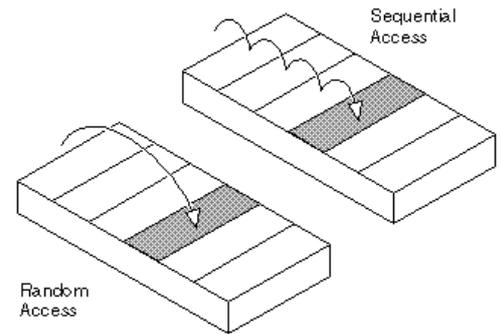
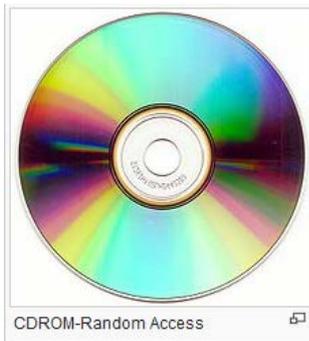
This can be useful to some users, if they are purposely attempting to process a sequence of data elements in order. However, this can also be time consuming for users who are trying to find a certain file on a disk tape, which requires skimming through all of the data in sequence.



3.1 Data representation

3.1.2 File organization and access

An example for comparison between random and sequential access would be the A-Z method. Sequential access would inquire the user to go through letters A-Z to achieve the goal of meeting the letter you desire. Whereas with random access, the user is able to jump directly to point “Z”.



The alternative access method is **direct-access** which allows records to be read and written in any order. Most systems only permit this for files stored on random-access devices such as discs; it is sometimes also permitted on tapes. All records in a direct-access file must be the same length so that the system can compute the location of a record from its record number unlike sequential access where variable length records are allowed. The record length has to be chosen when the file is created and (on most systems) is then fixed for the life of the file.

3.1 Data representation

3.1.2 File organization and access

Sequential organization

Consider a set of students whose records are stored in a computer. The data could have been stored in alphabetic order of their names. It could have been stored in the order that they came in a Computing exam, or by age with the oldest first. However it is done, the data has been arranged so that it is easier to find a particular record. If the data is in alphabetic order of name and the computer is asked for “Zaid” record, it won’t start looking at the beginning of the file, but at the end, and consequently it should find the data faster.

A file of data that is held in sequence is known as a sequential file

Random organization

A file that stores data in no order is very useful because it makes adding new data or taking data away very simple. In any form of sequential file, an individual item of data is dependent on other items of data. ‘Jawad’ cannot be placed after ‘Mahmood’ because that is the wrong ‘order’. However, it is necessary to have some form of order because otherwise the file cannot be read easily. What would be wonderful is if, by looking at the data that is to be retrieved, the computer can work out where that data is stored. In other words, the user asks for Jawad’s record and the computer can go straight to it because the word Jawad tells it where it is being stored.

