



PDHonline Course L153G (5 PDH)

Geographic Information Systems (GIS)–Hardware and software in GIS

Instructor: Steve Ramroop, Ph.D.

2020

PDH Online | PDH Center

5272 Meadow Estates Drive
Fairfax, VA 22030-6658
Phone: 703-988-0088
www.PDHonline.com

An Approved Continuing Education Provider

Lecture 2 Content

Geographic Information Systems (GIS)

- Hardware and software in GIS (Continued)**

This is a continuation of the hardware and software discussion as stated on the course description.

<ul style="list-style-type: none">❑ Data Storage<ul style="list-style-type: none">- storage media- fixed disks- dismountable devices- volumes- files❑ Software<ul style="list-style-type: none">- programs- operating systems- compilers and languages- additional programs	<ul style="list-style-type: none">❑ Editors and Word Processors❑ Databases<ul style="list-style-type: none">- functions of a database❑ Spreadsheets❑ Statistical Packages
---	--

L153 - Introduction to GIS Lecture 2
Dr. Steve Ramroop

2

We are going to generally look at the different storage devices, what is understood by the term software, how is software generated, and the various operating systems used in a computer. Then we are going to look at the editors and word processors that are particularly useful in GIS. In addition, the general context of a database is presented and then identifies some common functions of a database that is common required in GIS. We are then going to look at spreadsheets and statistical packages which will be useful as common software capable of integrating with the GIS software.

□ Data Storage

Storage Media:

- computers can use several different media for storing information
- necessary for the storage of programs and data
- media differ by:
 - storage capacity
 - speed of access
 - permanency of storage
 - mode of access
 - cost

The storage of data is an important consideration when deciding upon the use of programs and data in the computer. Consideration is needed in terms of what is an acceptable storage space to store the data and the multiple application programs. In terms of GIS, the data sets and the applications programs are typically large, and as such, the storage media must be very large. In addition, the speed at which the data sets and the program are accessed is another consideration. Processing of data should ideally be fast without having the end user waiting long periods of time.

The data stored and program must also be permanent (until they are intentionally removed). The mode of access refers to sequential access or random access to the data and the software files. Mode of access also refers to the security set to access the data and the software files in order to make changes.

Cost is another consideration for storage media. Today's storage media is relatively cheap and is accessible to everyone who is interested in computers.

Fixed disks:

- **most costly memory next to main/internal memory**
- **ranges from 10 Megabytes for typical PC to hundreds of Gigabytes in large “disk farms”**
- **random access but slower than internal memory**
- **permanent (i.e. does not disappear when the power is turned off), though data can be erased and modified**

The fixed disks are the hard drives. They are built in the CPU that houses the computer motherboard and other basic components used to run the computer. It is very important to have large disks in order to execute the top of the line software because they are designed for fast and high efficient computers. There is always a minimum specification defined to run any software, and GIS is by no means any different. Typically, the data sets and the present software are very large applications. The data and the software also influence the requirement for high specs of RAM.

Dismountable devices:

- dismountable devices can be removed for storage or shipping**
- include:**
 - **floppy diskettes**
 - up to 1.44 Megabytes for PC**
 - random access**
 - outdated**
 - **magnetic tapes**
 - tens of megabytes for standard tape**
 - access is sequential, not random**
 - can take minutes to reach a particular set of data on the tape, depending on where the data is stored**

Dismountable devices refer to the devices that are used to physically move data from one computer to another. The use of floppy diskettes is limited to small files that do not exceed 1.44MB. The access to floppy diskettes is random access. Floppy diskettes are rarely seen or used nowadays.

Magnetic tapes are being used by some organizations however its use is slowly being phased out. The drawback with magnetic tapes is that accessing data is sequential.....which means that to locate data / information the entire tape must be searched.

- **optical compact disks (CDs)**
 - around 750-800 Megabytes per CD
 - random access, the delay reaching a given item of data maybe 1 second or more
- **Jump drives or Flash drives**
- **DVD**
- **Bluray Disk**

Volumes:

- **a volume is a single tape, CD, diskette or fixed disk, DVD, Bluray Disk, i.e. a physical unit of storage**

Optical disks are popular in today's IT world. The access is random and as a result, it is faster rather than searching the entire disks when compared to magnetic tapes. CDs today are typically 750-800MB. They can be read only or rewritable CDs.

Volume is a term used to categorized similar data on tape, CDs, diskettes or fixed disks. This is important for large projects in which large amounts of data sets are used and there is need for the proper recognition of similar and related data sets. Typically the files will have some connotation through its file names and are managed by the creator of the data sets.

Files:

- a file is a logical collection of data which can be in the form of a table, document, program, map
- many files can be stored as a single volume
- files are given names, (the rules for naming files vary among types of systems)
- the computer operating system keeps track of files stored in a volume by viewing the directory
- files are identified in the directory by name, size, date of creation and often type of contents
- files are often organized into subdirectories so that the user can group files under specific topics

A file is the simplest logical collection of data and/or information which is stored on a computer. A collection of similar or related files are typically stored in a directory. A directory helps to categorize the data and information. GIS data sets are typically stored in separate directories and they are usually large files that are related to each other. Further subdivisions under a directory are called subdirectories under which there can be further subdivisions until the desired file is stored at the desired location. Files have a name, size, date of creation, and file type.

Software

Programs:

- **a program is a sequence of related instructions, performed one step at a time by the CPU to accomplish a task**

- **there are three types of programs:**
 - **operating systems**
 - **language interpreters and compilers**
 - **application programs**

The software is the program which is responsible for the operations to the computer. The instructions are defined in the software which is encoded using a known computer language and a compiler. The compiler converts the instructions into machine readable operations.

Operating systems is the important software which is used to control the boot up of the computer and also control the operation of the other programs installed on the computer.

Application programs are the software that is installed on the computer to do specific tasks. Examples of software are MS Office, Internet Explorer, and such like.

Operating Systems:

- **an operating system (OS) is the software which controls the operation of the computer from the moment it is turned on or “booted”**
- **the OS controls all input and output to and from the peripherals as well as the operation of other programs**
- **allows the user to work with and manage files without knowing specifically how the data is stored and retrieved**
- **in multi-user systems, operating systems manage user access to the processor and the peripherals and schedule jobs**

There are various operating systems (OS), for example: Windows 95, NT, Linux, Windows 2000, Windows XP, and such like. The OS controls the access to the computer and its programs, and in addition connects to printers, and other peripherals. OS also assist in the networking of multiple computers.

– common operating systems include:

- **IBM PCs and clones use MS-DOS (often called DOS), although there is movement to Windows, NT, OS/2, etc.**
- **UNIX (and others such as AIX, XENIX) is the dominant operating system for workstations. Linux is a shareware which is also growing popular**
- **mainframes commonly use proprietary operating systems developed by their manufacturers - VMS on DEC's VAX series, PRIMOS on Prime, CMS on IBM mainframe, etc.**
- **although functions performed by operating systems are similar, it can be very difficult to move files or software from one to the other**
- **many software packages run under only one operating system, or have substantially different versions for different operating systems**

This slide gives more information about common operating systems. It is important that different operating systems may require the conversion of files between the various OS if such files are to be used on other operating systems. For example, to use a file under the UNIX OS which is obtained from an IBM PC will require the DOS to UNIX conversion. However, the files used between the OS of the IBM PCs are usually read directly without any conversion.

Compilers and programming languages:

- Since computers operate on electricity and binary operations, all instructions executed by computers must be provided to the CPU in machine code.**
- However, humans do not have to interact with computers at this level.**

Compilers are the computer program which converts the language specific code into machine readable format. Human beings interact with the program at the programming level. Any changes in the program, needs to be done at such a level and not directly at the machine readable code level.

- **programs can be written in very specialized languages, called *assemblers*, which allow programmers to take advantage of the specific capabilities of particular machines by addressing the basic operations directly**
- **these languages are system specific and cannot be transported from one type of computer to another**
- **most programs are created using standard high level languages as JAVA, C++, Pascal, FORTRAN, BASIC which are common across most computer systems, from micro to mainframe**

There are various computer programming languages that range from the simplistic approach, such as BASIC, to that of the high level of programming such as JAVA. Software developers typically make use of these languages to develop software capable of performing operations that satisfy our various user-needs. The compilers are usually designed to run using the various computing platforms such as IBM PCs, UNIX, VAX, and such like.

- **The original human readable program is referred to as source code**
- **A compiler is a program designed to convert a program written in a high level language to the machine instructions of a specific computing system or “platform”**

The original program written by the software developer using specific computer languages is referred to the source code. This slide gives a simple definition of a compiler. In GIS there are compilers that are included with the standard software which further facilitate the customizations and extension of the standard software capability.

Application Programs:

- application programs are programs used for all purpose other than performing operating system chores or writing other programs**
- Examples includes GIS, word processors, spreadsheets, statistics packages, graphic programs, payroll system, etc.**

Another term used to refer to software which is capable of processing specific input from the computer user is called application programs. The application programs that are common to us are MS Office, Internet Explorer, and such like.

Editors and Word Processors:

- **are packages designed to modify or edit the contents of files**
- **are most often used to edit written text or programs**
- **editors and word processors are WYSIWYG (“what you see is what you get”)**
- **the screen shows a picture of the contents of the file at all times**
- **well known examples are:**
 - **Commercial : Word Perfect, Wordstar, Microsoft Word, Vi**
 - **Shareware: Latex**

The editors and word processors are the text processing application programs that usually come with the basic operating system installation. The common Notepad, VI editors are common with the OS installation. There are however, the more popular commercial word processors such as MS Word that are application programs that provide more processing capability and a better graphical user interface. There are also free word processors such as Latex which is command driven. Usually word processors do not give you what you see unless you create a postscript file. That is if you store your document in one computer and then take the same file to another computer then the format changes. With postscript files, what you see is what you get and it is transferable between computer platforms. The word processors that come with the operating system installation are used together with other application programs for example, in writing computer programs to extend the processing capability.

- **linkage to a printer is essential so that the user can obtain a “hardcopy” of a file’s content**
- **many mainframes offer their users several editors, unfortunately there is little standardization of editors**
- **an editor is the most important system to learn after the operating system**

After typing in the document, users would like to get a printout. This is facilitated if the computer is connected to a printer. The OS establishes the communication between the application program and the peripheral such as the printer.

❑ Databases

- **Data storage files which are managed by a Database Management System (DBMS).**
- **DBMS are packages designed to create, edit, manipulate and analyze data.**
- **To be suitable for a database, the data must consist of *records* which provide information on individual cases, people, places, features, etc.**
- **Each record may contain several *fields* each of which contains one item of information**
- **The number and interpretation of the fields must be constant for each class of records**
 - e.g. each record in the class of “streets” may contain fields for name, length, surface type, etc.
- **Field contents can be of many types - numeric or text, fixed or variable length**

Databases are one of the important aspects in every GIS application. They typically store data sets using various data models and structures. There are various databases. In order to execute the application programs data stored in databases are required. A typical database has rows and columns where each row is called a record and each column describes a specific characteristic of the data. Database is a separate course series which is not covered in detail in this Introduction to GIS series.

Functions of a database:

- **creating and editing records, using customized screens**
- **printing reports using customized report forms, including subtotals and totals**
- **selecting records based on user-specified rules**
- **updating records based on new information**
- **linking records**

This slide shows some of the common functions of a database. The points are self explanatory. These functions (among many others) are all available.

Types of database:

- 1. Network**
- 2. Hierarchical**
- 3. Relational**
- 4. Others (Object Oriented Database and Temporal Database)**

- although the first three are used, the relational model has been most successful within GIS**
- well known relational database management systems (RDBMSs) include dBase, Oracle, Info, etc.**
- many databases use the same language, SQL (Standard Query Language), for formulating queries**

These are some examples of database types. We will discuss each one of them in greater detail later on in the course series. To search databases the common Structured Query Language called (SQL) is used. There are other high level query languages.

□ Spreadsheets

- are programs which allow the user to work with numerical and character data in tabular form**
- column and row totals, percentages, etc. are automatically updated as data items are changed**

□ Statistical Packages

- offer a range of types of statistical analyses**
- data is primarily numerical**

Other application programs are Spreadsheets (e.g. Excel), and Statistical packages (e.g. SPSS).

– Statistical Software may include :

- **database functions, such as editing, printing reports**
- **capabilities for graphic output, particularly graphs but many produce maps**

The statistical programs are very useful to further process and analyze numerical data sets. The results from such processing can be stored in another database which can be used in other application programs.

... The End ...