Geographic Information Technology Training Alliance (GITTA) presents:

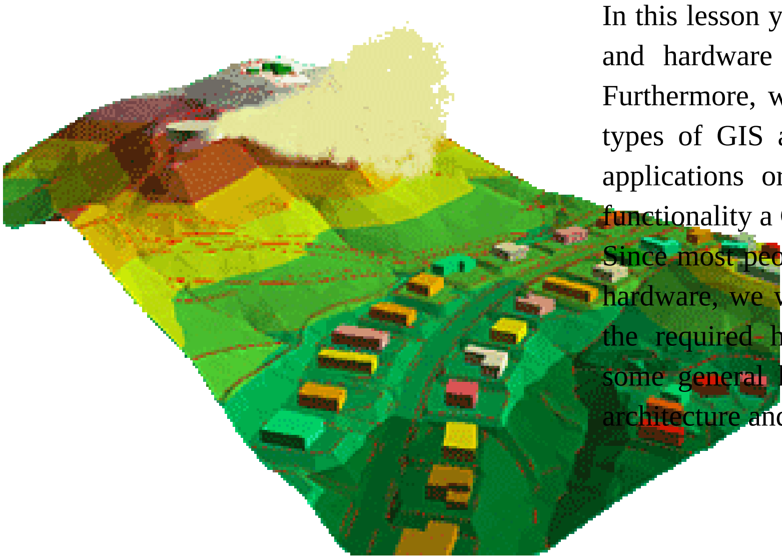
What do we need to work with a GIS?

Responsible persons: Christine Najar, Helmut Flitter, Simon Speich, Stefano Canepa, Valeria Agnolotti

Table Of Content

1. What do we need to work with a GIS?

Introduction



In this lesson you will learn about the software modules and hardware components necessary to run a GIS. Furthermore, we will give an overview of the existing types of GIS application software (referenced as GIS applications or GIS software) and discuss the main functionality a GIS should offer.

Since most people are more familiar with software than hardware, we will start with the software first. Second, the required hardware will be presented and finally some general knowledge about the underlying system architecture and internet technologies will be provided.

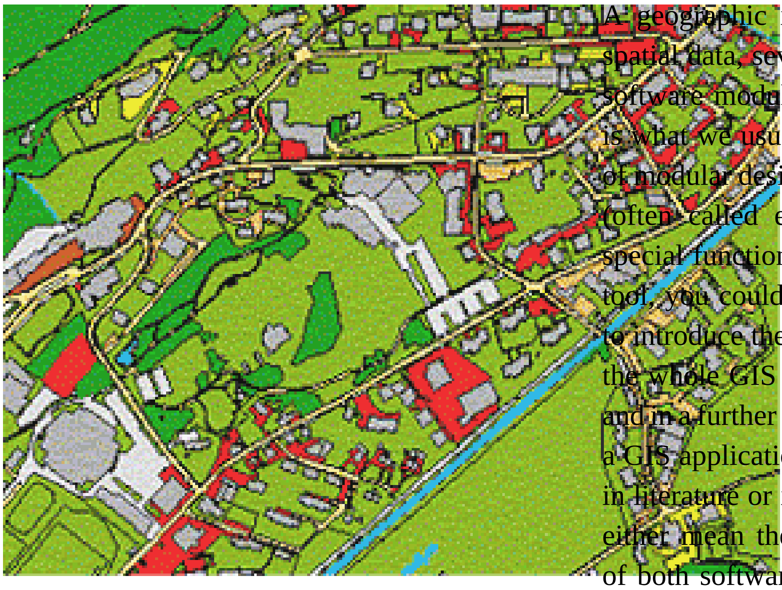
Visualization of 4D data (3D geometry and time)

Learning Objectives

- You know the different software and hardware components that are part of a GIS and what these parts are used for.
- You know the basic functions a GIS application offers.
- You understand the common software architectures (monolithic, client/server, middleware).
- You are familiar with terms like HTTP, TCP/IP, etc.

1.1. GIS software: functionality

Introduction



Example of a map of Davos (www.geoswiss.ch).

A geographic information system is a combination of spatial data, several different hardware components, and software modules. Only one of these software modules is what we usually call a GIS application, which can be of modular design, too. You would only buy the modules (often called extensions) you wanted. If you needed special functionality, let's say a special 3D visualization tool, you could add it easily later. In this unit, we want to introduce the different software modules that make up the whole GIS together with their general functionality, and in a further step, we will discuss the specific functions a GIS application should offer. The use of the word GIS in literature or in this module might be confusing. It can either mean the GIS application only or the whole set of both software and hardware that make up the whole information system (see definitions from **lesson 1**).

1.1.1. Software modules: general functionality

On the basis level, there is the system software, which includes the operating system (OS) and its services such as file operations and internet connectivity. We will talk about this in a little more detail in the next unit **'System architecture'**. On a higher level there are specific software such as the database application and the GIS application itself, which make use of functions provided by the operating system. Some GIS come with a small database already integrated, but in order to process larger quantities of data you need an additional full stand-alone database server. Another module is the user interface (UI), which can be considered as software as well as hardware. Some of its functions are provided by the operating system, some by the application itself. Click on the text buttons (User interface, GIS software, etc.) to display the corresponding information.

Only pictures can be viewed in this version! For Flash, animations, movies etc. see online version. Only screenshots of animations will be displayed. [link]

1.1.2. GIS application: specific functionality

Although most GIS software packages share similar features, programs vary in price and functionality. Some have a number of built-in functions whereas others have additional modules/extensions that can be purchased separately to perform specific types of analysis.
The basic functions GIS software should offer can be grouped into data capture, data management, data analysis and representation:

- Data capture - no data, no GIS.
  Data capture functionality could either be the importing of digital data from a disk, network or database, or digitizing it through the use of peripheral devices.
- Data management - quality is essential.
  One step beyond data capture is data editing and maintenance. Editing involves functions such as adding, duplicating and deleting spatial and attribute data as well as creation and management of metadata. Using tools to detect errors and to perform transformations is another aspect of data management.
- Data analysis - ask questions.
  A GIS is typically used to answer questions related to space. For the purpose of finding an answer, a GIS should not only offer tools to query and retrieve data, but also functions for data integration and modelling.
- Visualization - make it visible.
  Any analysis of data is only useful if the final output is also understood by the intended recipient. A GIS offers tools to present data in various ways, ranging from displaying tabular information, over graph/ chart production, to printing maps. Some GIS offer advanced visualisation beyond two dimensions, such as *digital elevation models* [1] (DEM) or digital movie files for "fly throughs". Output can also be shared by storing it on a disk or a webserver, which allows for dynamic views of the data.

Aside from these basic functions, a GIS should also offer ways for customizing and creating new tools and applications according to your needs. Depending on the software, this can either be done through well known programming or scripting languages such as Visual Basic or Python, or through a programming language specifically provided by the application (eg. ESRI's 'Avenue').
Solve the puzzle below to learn more about GIS software functionality.

Software functions

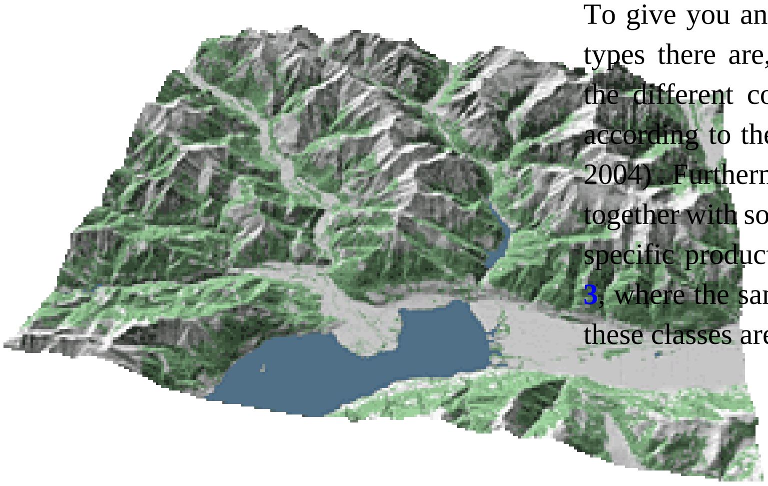
Only pictures can be viewed in this version! For Flash, animations, movies etc. see online version. Only screenshots of animations will be displayed. [link]

Drag and drop the puzzle pieces from the scrolling pane to the empty slots. If you mouse over the pieces, a short explanation will be displayed in the text area below. As soon as all slots are occupied you will be shown the correct answers. To start over click the reset button and the pieces will be randomly shuffled and put back in place.

[1] Digital Elevation Model. A digital representation of a continuous variable over a two-dimensional surface by a regular array of z values referenced to a common datum. Digital elevation models are typically used to represent terrain relief. This is a common GIS product created at a variety of scales.

1.2. GIS software: categories

To give you an idea of the wide range of GIS software types there are, and to allow for easier comparison of the different commercial products, we classified them according to the German "*GIS-Report*" (Buhmann et al. 2004) . Furthermore, we characterize their possible usage together with some application examples. Descriptions of specific products will be given in more detail in **lesson 3**, where the same classification will be used. As always these classes are somewhat arbitrary.



Digital Elevation Model of Locarno area

1.2.1. Characterization of GIS software categories

Click on the buttons (Internet-GIS, Desktop GIS, etc.) to display the corresponding information or use the next / previous buttons to skip through all categories.

Only pictures can be viewed in this version! For Flash, animations, movies etc. see online version. Only screenshots of animations will be displayed. [link]

1.3. Hardware

A world of components

User interface of a mobile GIS (GPS receiver included)

While hardware was once the main limiting factor in a GIS implementation because of its high cost, large size and low computing power, this is no longer the case. In the same way the hardware's capabilities have almost been exponentially increasing, the costs and size have been decreasing. Today a GIS can be run on a simple desktop computer system without the need for a *mainframe* [2] anymore. Today, the limiting factors are more often the amount of time needed for data collection and management.
The hardware components required to run a GIS can be divided into three general categories according to specific tasks of the system. In this unit, we present what these tasks are and which components are generally needed to solve them. Specific tasks may still need additional components.

1.3.1. Hardware

Data acquisition

Data can be acquired by various tools and methods. New data can be collected in the field by means of surveying and remote sensing devices. Nowadays these acquisition systems are often directly connected to a GIS by means of a mobile device or notebook, enabling the operator to verify the data directly in the field. Data collection is usually done by governments, organisations or private companies. The data is stored on database servers, where it can be accessed and downloaded over the internet or the data can be purchased on storage media.
In the beginning of GIS, and still today, maps, plans or photos often exist, and make digitizing necessary, without the need to collect new data. In the case of a paper map, this can be done by using a digitizer tablet, where an operator traces the outline of objects, such as roads, streams or boundaries with the mouse and the tablet captures the coordinates of its position. However, this method is increasingly replaced by a process known as "heads-up" digitizing, where the data collector traces a map or image right on the screen with the

[2] A term originally referring to the cabinet containing the central processor unit or "main frame" of a room-filling Stone Age batch machine. After the emergence of smaller "minicomputer" designs in the early 1970s, the traditional big iron machines were described as "mainframe computers" and eventually just as mainframes. The term carries the connotation of a machine designed for batch rather than interactive use, though possibly with an interactive time-sharing operating system retrofitted onto it; it is especially used of machines built by IBM, Unisys and the other great dinosaurs surviving from computing's Stone Age (The Free On-line Dictionary of Computing).

mouse. This means that the maps have to be digitized beforehand using a scanner or, in case of photos, with a photogrammetric device. The automatic interpretation of images instead of manually tracing geometrical objects is currently one of the major research tasks in photogrammetry and related disciplines.

Data management and analysis

A Geographic information system runs on different kinds of computer categories, from a simple PC to a specialized database server. The exact needs depend on the amount of data and the complexity of the processing to be done. The faster the computer (resp. the CPU) and the more RAM, the better the working performance. In order to store the often vast quantities of data you need mass storage devices and means of backing up the data. For communication between devices, a network is needed.

Data output and presentation

There are several ways of outputting data. Probably the most common way is to display your data in the form of a map on a screen. Other means are sending the data to a printer or plotter. However, if you want to continue handling the geographic information with other programs and tools, or if you want to keep up with the development of new methods and applications, the spatial data needs to be exported in digital form. Output can then either be done on storing media or over a network. Often the data is published on a web- or database server, where other users can access and query the same data. Today, this form of output is getting more and more important, which demands for the establishment of technical as well as communication standards for interfaces and procedures.

Solve the puzzle below to get an idea of the different hardware devices and their function in a GIS.

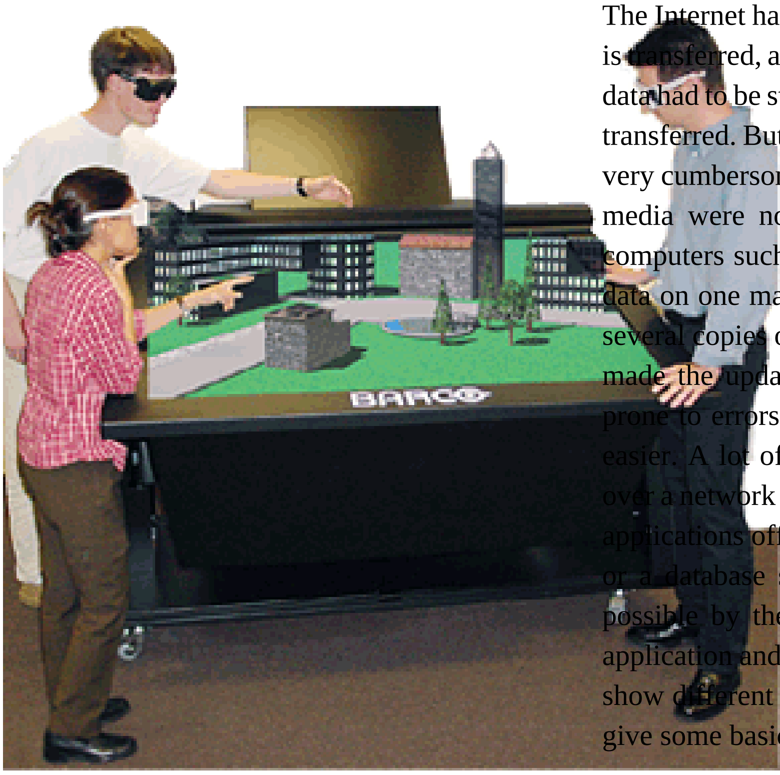
Hardware puzzle

Select a puzzle piece (a hardware picture) from the scrolling pane and place it on a target symbol in one of the 3 categories (input, data management, output). If you mouse over a piece it's name will be shown in the text box below. With the arrow buttons you can scroll or jump to the beginning or end. After having placed all the pieces on the targets, the result will display automatically. Press the reset button to start over.

Only pictures can be viewed in this version! For Flash, animations, movies etc. see online version. Only screenshots of animations will be displayed. [link]

1.4. System architecture

Introduction



3D visualization table ETHZ

The Internet has opened new ways in which (spatial) data is transferred, accessed and used. In the early days of GIS, data had to be stored on physical media before it was to be transferred. But, the amount of the data made this transfer very cumbersome if not impossible, and often the storage media were not compatible between different type of computers such as mainframes, PCs and Macs. Because data on one machine wasn't accessible by another, often several copies of the same spatial data file existed, which made the update process a very difficult task that was prone to errors. Nowadays sharing of data is quite a bit easier. A lot of GIS data and functionality is accessible over a network by means of a web browser and many GIS applications offer ways to connect directly to the internet or a database server. These communications are made possible by the modular architectures of both the GIS application and the operating system. In this unit, we will show different types of system architectures and we will give some basic information about internet technologies.

1.4.1. System architecture

Software architectures

A persisting technology often experiences some major changes during its ongoing development, building upon the standards of a previous generation and (probably) continuing to do so in the future. Over time Geo Information System architecture has moved away from the early monolithic architecture to a client-server, and then to a more distributed architecture.

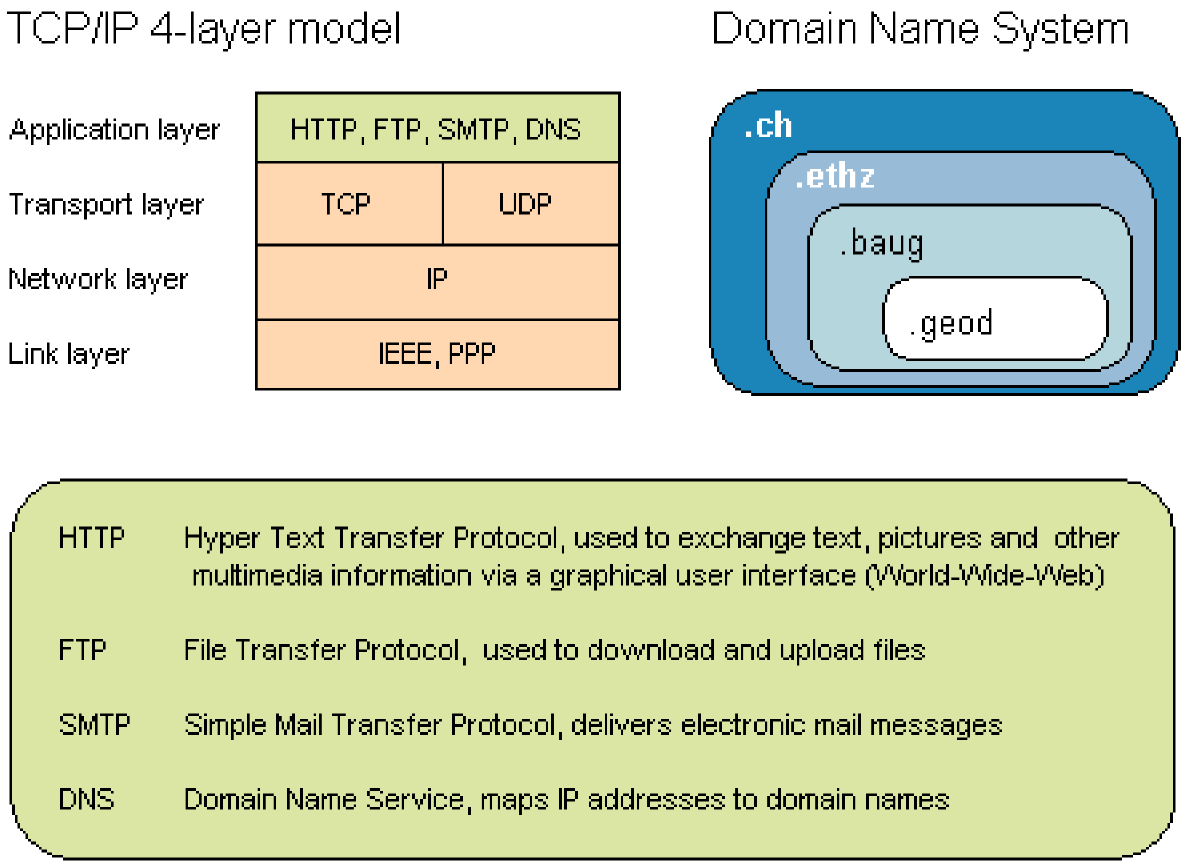
At the beginning of application development, monolithic applications were a bit like rocks (Greek monolithos, consisting of a single stone: monos, single, alone + lithos, stone). They were made up from one large, single program containing all the functionality. As the scope of the programming tasks increased, the need for reusability did, too. Therefore, the software's tasks were separated into well-encapsulated entities, which could be reused whenever the same problem came up, or if one of the entities was replaced by a new software generation. Unfortunately, the level of reuse was mainly at the level of source code within an application, where it was not accessible from other applications. So the next step was to create software modules with specific tasks such as the application or database server, whose functionality can be accessed by a client-computer. Over time database applications not only grew in complexity but they also took over a lot of processing rules and tasks (stored procedures), which were not really modular. This contradicted the idea of reusability, so finally the middleware architecture evolved, where a special middle layer handles all the processing between client and server, and over which communication between all kinds of applications / modules is possible due to its open and modular architecture.

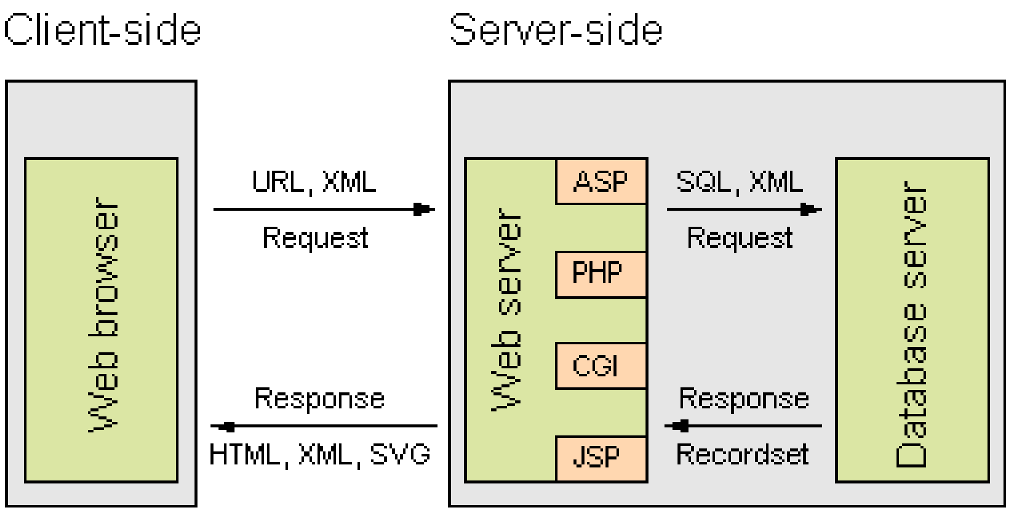
Only pictures can be viewed in this version! For Flash, animations, movies etc. see online version. Only screenshots of animations will be displayed. [link]

GIS and internet

Data exchange in the internet is based on different protocols, which are generally handled by the operating system. One important protocol is the TCP/IP (Transmission Control Protocol/Internet Protocol) network protocol, actually a suite of layered protocols. This means that each protocol layer builds upon the layer below, adding new functionality. On the lowest level, there are protocols that are implemented right into the network adapter which are responsible for communication with the actual network hardware (e.g ethernet card). Above that are protocols that handle the connection and routing, and on top are the application protocols designed for tasks such as transferring files or sending and receiving E-mail.

Internet addresses can be symbolic or numeric. The symbolic form is for humans, because it is easier to read, for example: http://www.geod.baug.ethz.ch. Its corresponding machine readable form would be 129.132.26.4 and is used by the IP protocol. The mapping between the two is done by the Domain Name System (DNS).





Another important protocol is HTTP (Hyper Text Transfer Protocol). It is most frequently used by web browsers (Firefox, Internet Explorer, Netscape/Mozilla, Opera, Safari and others) to access documents, images, video and sound. It defines how pages are formatted and transmitted, and what actions web servers and browsers should take in response to various commands. For example, when you enter an URL (Uniform Resource Locator) in the form of „Protocol://Domain Name:Port/Directory/Filename", this actually sends an HTTP command to the Web server directing it to send the requested web page back. But before this happens the web server might parse the page first. This is done by a script language (*ASP* [3], *CGI* [4], *PHP* [5], etc), which might then query a database. Only the processed page is then sent back. The browser receives this page, parses it too, and creates the output according to the instructions (*HTML* [6], *XML* [7], *JavaScript* [8], *SVG* [9]).

1.4.2. Test your knowledge

By answering the following questions, you can test your knowledge about this unit. The questions will be presented to you in random order. As soon as you complete the questions on one page, you will gain access to the next page (the next button turns dark blue). After answering all questions your score will be shown. You can press the button "Show Answer" to display some explanation (actually you can do so any time, but this would be cheating :-). Press the "Reset" button to start over.

---

[3] Active Server Pages

[4] Common Gateway Interface

[5] Hypertext Preprocessor

[6] HyperText Markup Language: A Hypertext document format used on the World-Wide Web. Built on top of SGML (Standard Generalized Markup Language).

[7] eXtendable Markup Language

[8] Java Script: (Formerly LiveScript) Netscape's simple, cross-platform, World-Wide Web scripting language, only very vaguely related to Java. JavaScript is intimately tied to the World-Wide Web, and currently runs in only three environments - as a server-side scripting language, as an embedded language in server-parsed HTML, and as an embedded language run in browsers (The Free On-line Dictionary of Computing).

[9] Scalable Vector Graphics

Consider the questions only as general statements when answering them, of course you can always find an exception to the question.

Only pictures can be viewed in this version! For Flash, animations, movies etc. see online version. Only screenshots of animations will be displayed. [link]

1.5. Summary

- The software functionality a GIS should offer can be grouped into functions for data capture, data management, data analysis, and representation.
- Many different GIS exist and they cover a wide range of fields. They can be categorized as "General purpose GIS", "Internet GIS", "Desktop GIS", "CAD GIS", "Business Mapping", "GIS complenents", "Mobile GIS" and "Open source GIS" .
- The hardware components required to run a GIS can be divided into four general categories according to specific tasks of the system: data acquisition, data management, data analysis, and data output and presentation.
- A lot of GIS data and functionality is accessible over a network and many GIS applications offer ways to connect directly to the internet or a database server. These communications are made possible by the modular architectures of both the GIS application and the operating system.

1.6. Glossary

ASP:
: Active Server Pages

CGI:
: Common Gateway Interface

DEM:
: Digital Elevation Model. A digital representation of a continuous variable over a two-dimensional surface by a regular array of z values referenced to a common datum. Digital elevation models are typically used to represent terrain relief. This is a common GIS product created at a variety of scales.

HTML:
: HyperText Markup Language: A Hypertext document format used on the World-Wide Web. Built on top of SGML (Standard Generalized Markup Language).

JS:
: Java Script: (Formerly LiveScript) Netscape's simple, cross-platform, World-Wide Web scripting language, only very vaguely related to Java. JavaScript is intimately tied to the World-Wide Web, and currently runs in only three environments - as a server-side scripting language, as an embedded language in server-parsed HTML, and as an embedded language run in browsers (The Free On-line Dictionary of Computing).

mainframe:
: A term originally referring to the cabinet containing the central processor unit or "main frame" of a room-filling Stone Age batch machine. After the emergence of smaller "minicomputer" designs in the early 1970s, the traditional big iron machines were described as "mainframe computers" and eventually just as mainframes. The term carries the connotation of a machine designed for batch rather than interactive use, though possibly with an interactive time-sharing operating system retrofitted onto it; it is especially used of machines built by IBM, Unisys and the other great dinosaurs surviving from computing's Stone Age (The Free On-line Dictionary of Computing).

PHP:
: Hypertext Preprocessor

SVG:
: Scalable Vector Graphics

XML:
: eXtendable Markup Language

1.7. Bibliography

- **Anonymous**. *The Free On-line Dictionary of Computing* [online]. Available from: http://www.instantweb.com/foldoc/ [Accessed 2016-05-02].
- **Bill, Ralf**, 1999. *Grundlagen der Geo-Informationssysteme*. Karlsruhe: Wichmann Verlag.
- **Buhmann, E., Wiesel, J.** (2004). *GIS Report* [online]. B. Harzer Verlag, Karlsruhe. Available from: http://www.harzer.de/gr04web.pdf [Accessed 2016-05-02].
- **David J. Buckley**. *The GIS Primer* [online]. Available from: http://www.innovativegis.com/basis/primer/primer.html [Accessed 2016-05-02].
- **Dr. Michael P. Peterson**. *Geographic Information Systems* [online]. Available from: http://maps.unomaha.edu/Peterson/gis/ [Accessed 2016-05-02].
- **Geospatial Media and Communications Pvt. Ltd**. *Geospatial Media and Communications* [online]. Available from: http://geospatialmedia.net/ [Accessed 2016-05-02].
- **U.S. Geological Survey**. *Geographic Information Systems - Posters* [online]. Available from: http://erg.usgs.gov/isb/pubs/gis_poster/ [Accessed 2005-01-21].