GIS-601

M.A. /M.Sc. Geo-informatics

RECENT TRENDS IN GEOINFORMATICS
PART-II

DEPARTMENT OF REMOTE SENSING AND GIS
SCHOOL OF EARTH AND ENVIRONMENT SCIENCE
UTTARAKHAND OPEN UNIVERSITY
HALDWANI (NAINITAL)

RECENT TRENDS IN GEOINFORMATICS PART-II

DEPARTMENT OF REMOTE SENSING AND GIS
SCHOOL OF EARTH AND ENVIRONMENTAL SCIENCE
UTTARAKHAND OPEN UNIVERSITY

Phone No. 05946-261122, 261123
Toll free No. 18001804025
Fax No. 05946-264232,
E. mail info@uou.ac.in
htpp://uou.ac.in

Board of Studies

Programme Coordinator

Dr. Ranju J. Pandey
Department of Geography & NRM,
Department of Remote Sensing and GIS,
School of Earth and Environment Science,
Uttarakhand Open University, Haldwani

RECENT TRENDS IN GEOINFORMATICS PART-II

Wrap-up Material	
Units Written by	
Dr. Suneet Naithani Associate Professor, Doon University, Dehradun.	1,2,6,7,8,9 & 10
Dr. Harish Karnatak Senior Scientist & Group Head, IIRS, Dehradun.	3,4 & 11
Dr. Manish Kumar Assistant Professor, Central University Haryana. & Amit Ray	5

Title : Recent Trends in Geoinformatics Part-II
ISBN No. :

Edition : First (2021), Second (2022)

Published By : Uttarakhand Open University, Haldwani, Nainital-263139
Printed By :

CONTENTS

RECENT TRENDS IN GEOINFORMATICS PART-II

BLOCK 1: LIDAR

BLOCK 2: WEB GIS

BLOCK 3: CROWED SOURCING

BLOCK 1 : LIDAR

UNIT 1 : PRINCIPLES AND PROPERTIES- DIFFERENT LIDAR SYSTEM, SPACE BORNE AND AIRBORNE LIDAR MISSIONS, TYPICAL PARAMETERS OF LIDAR SYSTEM

1.1 OBJECTIVES

After reading this unit you should be able to:

- Know about LIDAR system.
- Gain knowledge about parameters of LIDAR system.

1.2 INTRODUCTION

LiDAR - Light Detection and Ranging - is a remote sensing method used to examine the surface of the Earth.

Lidar (Light Detection and Ranging) is a remote sensing technique that measures ranges (varying distances) to the Earth using light in the form of a pulsed laser. When these light pulses are integrated with additional data collected by the aerial system, exact three-dimensional information about the Earth's shape and surface properties is generated.

A lidar instrument is made up of three parts: a laser, a scanner, and a specialized GPS receiver. The most frequent platforms for collecting lidar data over large areas are planes and helicopters. Topographic and bathymetric lidar are the two forms of lidar. Bathymetric lidar employs water-penetrating green light to determine seafloor and riverbed elevations, while topographic lidar uses a near-infrared laser to scan the land.

Scientists and mapping experts can use lidar devices to analyze both natural and man-made surroundings with accuracy, precision, and flexibility. Lidar is being used by NOAA scientists to create more accurate coastal maps, create digital elevation models for use in geographic information systems, and assist in emergency response operations, among other things.

1.3 PRINCIPLES AND PROPERTIES – DIFFERENT LIDAR SYSTEM, SPACE BORNE AND AIRBORNE LIDAR MISSIONS, TYPICAL PARAMETERS OF LIDAR SYSTEM

Principle of LiDAR

The idea of LiDAR is similar to that of the Electronic Distance Measuring Instrument (EDMI), in which a transmitter fires a laser (pulse or continuous wave) and the reflected energy is recorded (Figure 1). The distance between the transmitter and reflector is calculated using the laser's time of travel (ToT). Natural items or manmade reflectors, such as prisms, could be used as reflectors. In the case of range LiDAR, this distance is one of the major measures, which, when combined with other measurements, yields the reflector's coordinates.

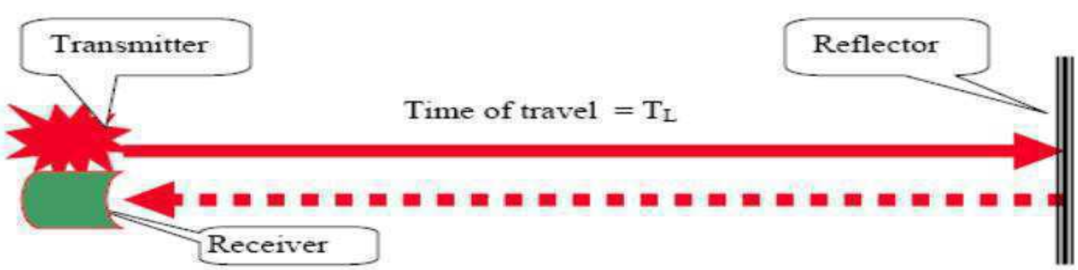


Figure 1: Principle of range measurement using laser
Source: http://home.iitk.ac.in/~blohani/LiDAR_Tutorial/Principle%20of%20LiDAR.htm

The evolution of lidar is depicted schematically in Fig 2, from early ground-based measurements through our first long-duration spaceborne investigations. It displays the initial lidars as ground-based in the 1960s, followed by systems first flown in 1969 on small aircraft, and then lidars flown on large aircraft capable of long-range observations in the late 1970s. Flights onboard high-altitude aircraft began in 1979, with data collected at an altitude of roughly 20 km. The depiction concludes with the first long-duration spaceborne low-Earth-orbit flight, the Geoscience Laser Altimeter System (GLAS), launched onboard ICESat in January 2003, and the first spaceborne lidar employing Shuttle for the 11-day flight of LITE, the Lidar In-space Technology Experiment in 1994.
The initial flights using elastic backscatter for cloud and aerosol studies were trailblazers in the evolution of lidar. This chapter will cover the firsts in airborne and spaceborne missions, as well as flights of lidars utilising alternative techniques such as DIAL, before delving deeper into specific examples of both airborne and spaceborne missions. It will come to a close with a glance ahead. It will not go over airborne or spaceborne laser altimeters, bathymeters, or Doppler lidars in any depth.

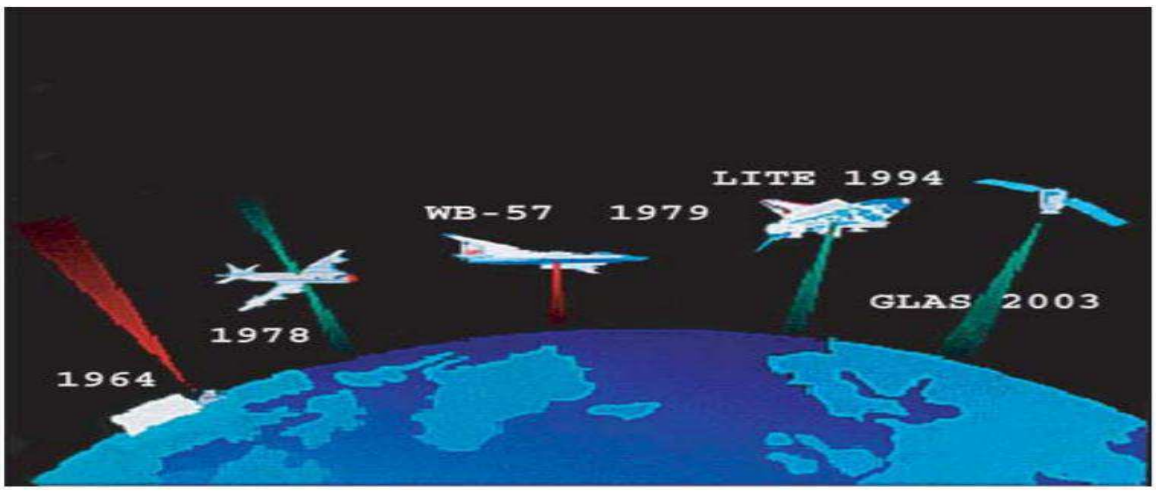


Figure 2: An artist sketch depicting the evolution of lidar "firsts."
Source: https://bit.ly/3ojZuPP

Types of LiDAR systems

With the capability of LiDAR technology, remote sensing, surveying, and monitoring have reached unprecedented heights in recent years. LiDAR (Light Detection and Ranging) is now used in a variety of industries, including forestry, disaster management, weather prediction, construction, archaeology, autonomous cars, and many more.
It's worth noting, though, that there are many different types of LiDAR, and picking the appropriate one is critical to its effectiveness. The many varieties of LiDAR are explained here, first by functionality and then by application.

- *Types of LiDAR by functionality*
  Let's take a look at the many varieties of LiDAR and how they're classed based on how they're set up and how they work.

  a. *Terrestrial LiDAR:* Terrestrial LiDAR, as the name suggests, is a ground-based system. It can be installed on a moving vehicle or at a stationary position. Terrestrial LiDAR data is useful for applications that demand a detailed survey of the ground or a "closer look" at things in either case.

Construction, self-driving vehicles, road surveys, city surveys, and other applications of terrestrial LiDAR are just a few. Terrestrial LiDAR can be divided into two types: mobile and static. Here's how they're different.

b. *Mobile:* A sensor, a global positioning system (GPS), an inertial navigation system (INS), and a couple cameras make up a mobile LiDAR setup. Because the unit is mounted on top of a moving vehicle, such as a car or a train, it is mobile. The LiDAR unit continues to shoot out laser pulses in all directions and read the reflections from this moving car. These priceless point clouds (data points) are then analysed to determine the state of roads and railway tracks, as well as detect undesired road impediments. An advanced rotating LiDAR sensor installed on top of the car identifies the presence of pedestrians and other vehicles on the road in self-driving automobiles.
c. *Static*: In some applications, having the LiDAR unit stationary at one location rather of moving around is preferable. Static LiDAR is used in such applications. The LiDAR unit is mounted on a stationary object, which is commonly a tripod, in this setup. The complete apparatus, including the tripod, can be transferred to a new place if necessary. In essence, this machine is entirely portable, despite the fact that it is not mobile. From a fixed place, a static LiDAR unit continues to deliver laser pulses to the surrounding region. The information is then utilised to deduce the features of the environment. This feature is extremely valuable in a variety of applications, including building construction, mining, and engineering.
d. *Airborne:* When a LiDAR unit is airborne, it indicates the system is installed in an aeroplane or a helicopter that hovers above the earth's surface, emitting laser pulses downward as it moves. When compared to terrestrial LiDAR, airborne LiDAR can scan larger areas in less time. As a result, airborne LiDAR devices are well suited to applications that demand a bird's eye view over a large area. Airborne systems can be divided into topographic and bathymetric LiDAR based on the type of region scanned by the LiDAR unit. Continue reading to learn more about these types.
e. *Topographic*: Topographic LiDAR scans any type of terrain, with laser pulses transmitted down to the earth's surface providing an estimate of the area's varied properties. The altitude of the structures that reflect the laser beams is used to map out the rise and fall of the surface. In a nutshell, topographic LiDAR is used to create a topographic map of a specific area of land. Forestry, urban planning, ecology, infrastructure mapping, geomorphology, and other fields use topographic LiDAR.
f. *Bathymetric:* While topographic LiDAR can detect any type of land from a distance, it struggles when scanning water features. Another sort of airborne LiDAR technology, known as bathymetric LiDAR, is employed to do this work. A bathymetric LiDAR sensor has all of the components of a topographic LiDAR with the addition of the ability to send green laser pulses. These pulses can pass through water and back to the airborne vehicle. The depth of the water bodies can be estimated using the data acquired in this manner. These units can more clearly identify shorelines and elevations when used in conjunction with topographic sensors. LiDAR systems are commonly used in coastal engineering and marine sciences.
g. *Satellite LiDAR:* LiDAR devices can also be installed in spacecraft orbiting the planet. It is now possible to scan larger areas of the globe, as well as the

airspace above it, using satellite LiDAR systems. NASA has used a variety of space-borne LiDAR systems to study cloud positions above the planet, vegetation, the state of the ice at the two poles, and other topics. More improved satellite LiDAR devices that can read particles in the atmosphere are being developed.

- *Types of LiDAR by application*
  Although the underlying operating concept of most LiDAR systems is identical, particular applications necessitate the use of unique LiDAR units. These specifications are used to classify LiDAR systems.
  a. *Differential Absorption LiDAR (DIAL):* This technology is used to monitor ozone levels in the atmosphere by measuring gas concentrations in the atmosphere. DIAL systems can be used on the ground or in the air, and they can also assess pollution levels. Attunable laser is used to generate two wavelengths of laser pulses, one for recording data from the peak of a gas absorption line and the other for recording data from a low-absorption zone. Differential absorption gets its name from this.
  b. *Wind LiDAR:* Measuring the flow of wind with any type of remote sensing technology can be difficult, partly because the wind changes direction quickly. Wind LiDAR was created specifically for this purpose. Advanced LiDAR systems that continually monitor the wind 360 degrees can be particularly useful in determining turbulence, wind speed, and wind direction using many data points.
  c. *Raman LiDAR:* Raman LiDAR is a type of terrestrial LiDAR that is used to identify and analyse water vapour and aerosol levels in the environment. While a traditional LiDAR unit obtains data from the backscattering of reflected laser pulses, Raman LiDAR can detect signals present in the backscattering at various wavelengths. This is known as inelastic scattering. The presence of molecules is primarily responsible for these signals. The LiDAR can accurately determine the presence of aerosols using this technique.
  d. *HSLR LiDAR:* Another method for determining the presence of aerosols in the atmosphere is HSRL, which stands for High Spectral Resolution LiDAR. HSRL, on the other hand, is an airborne system, unlike a Raman LiDAR. To distinguish between pulses reflected by molecules and those sent back by aerosols, the device uses the spectral distribution of reflected LiDAR signals

.

History of Airborne LiDAR

Researchers began developing devices to be carried aboard aircraft after early breakthroughs using lidars to explore the atmosphere in the 1960s and early 1970s. These were appealing because they provided a more "regional" capability, the ability to relocate to the area of concern for the measurements required, or the ability to capture the phenomenon of interest more effectively.

S. Harvey Melfi flew aboard a NASA Langley Research Center (LaRC)T-33 aircraft over Williamsburg, Virginia, at constant altitudes, making measurements with a forward-looking, very modest lidar, while the author made "uplooking" simultaneous measurements with a ground-based zenith-pointed lidar in 1967. The goal of this aerial study was to build lidars that could detect clear air turbulence [2]. The Stanford Research Institute developed and flew the first "downlooking" airborne lidar in 1969 for lower tropospheric aerosol observations during the Barbados Oceanographic and Meteorological Experiment [3]. The first "uplooking" airborne lidar was a two-wavelength elastic backscatter system designed by LaRC that measured aerosol profiles to confirm the Stratospheric Aerosol Measurement-II (SAM II) satellite Earth-orbiting mission launched in October 1978 aboard the Nimbus-7 spacecraft. This "ground truth" experiment took place in November 1978 at Sondrestrom, Greenland. It was followed by similar flights from Poker Flat, Alaska, in July 1979. SAM II and another satellite mission dubbed the Stratospheric Aerosol and Gas Experiment (SAGE), which was launched in February 1979 aboard the Applications Explorer Mission satellite [5], were both validated by the Poker Flat mission. Unlike SAM II, which recorded stratospheric aerosols in the polar regions, SAGE measured not only aerosols but also ozone on a far more global scale. Aerosol backscatter observations from aeroplanes were timed and spaced to match with satellite readings. During satellite overpasses, simultaneous balloon-borne and aircraft aerosol in situ observations were also timed. Validation measurements comprised balloon-borne and rocket-borne measurements at Poker Flat to validate the SAGE ozone readings.

Laser altimeters and lidar bathymeters were also used on aeroplanes in the mid-1970s, however they are not the subject of this book. A NASA C-54 aircraft, for example, flew a N2-Ne laser over the Chesapeake Bay and over Key West, Florida in 1974. It flew at low altitudes in clear water, measuring water depths. The Cloud Lidar System (CLS), developed by NASA Goddard Space Flight Center (GSFC) and flown aboard the WB-57 aircraft in 1979, was the first lidar designed for a high-altitude aircraft. It had to fly autonomously with minimal pilot contact because it was required. The knowledge gathered would be applied to future spaceborne lidar, including simulations required for spaceborne lidar design.

Aerosol and cloud lidars continued to be used aboard aeroplanes throughout the 1980s, 1990s, and into the present, expanding their capabilities as new technologies and uses became available. A wide range of wavelengths and polarisation measuring techniques were used. Airborne lidar applications were expanded thanks to higher repetition rates and more efficient lasers, as well as improved and faster data gathering and storage devices. Airborne aerosol and cloud lidars, for example, circled the globe imaging stratospheric volcanic layers, Saharan dust, stratospheric aerosols, and polar stratospheric clouds (PSCs). Other lidar measurement techniques began to be utilised aboard planes as well. NASA LaRC operated the first airborne UV DIAL system for ozone measurements in a "downlooking mode" as early as 1980.

In 1982, NASA LaRC built and flew an aerial water vapour DIAL system for the first time. It was also set to "down-looking" mode. In 1985, NASA GSFC used a tunable alexandrite laser to build a down-looking aerial DIAL system to measure atmospheric pressure for the first time. Because a reasonably high-resolution ozone measurement was required for the Antarctic ozone hole campaign performed out of Punta Arenas, Chile, in 1987, the LaRC DIAL system was re-configured to make the first "up-looking" DIAL observations of ozone. Data could be taken into the Antarctic vortex by staging out of Punta Arenas. These measurements, along with a variety of other types of ozone photochemistry measurements, were critical in determining how the ozone hole arose. In 1983, an airborne system that used resonance fluorescence to measure Na in the mesosphere flew for the first time. Mesopause density disturbances were explored in this way.

The first DIAL observations were taken aboard a high-altitude aircraft, in this case an ER2, using a technology dubbed LASE (Lidar Atmospheric Sensing Experiment), which was designed to profile water vapour in the troposphere. It served as a test bed for future spaceborne DIAL systems, in this case for DIAL applications, in addition to allowing measurements of water vapour across wide terrestrial distances during ER2 flights.

The Use of Airborne Lidar

A variety of atmospheric applications have benefited from the use of airborne lidars. Studies of the planetary boundary layer, long-range pollution transport, power plant plumes, volcanic aerosols in the stratosphere, mesospheric and stratospheric winds and gravity waves, Saharan dust, PSCs, water vapour and the hydrologic cycle, ozone associated with biomass burning, Arctic and Antarctic ozone associated with the ozone hole and ozone depletion, and the validation of satellite experiment measurements in the stratosphere and troposphere are among. Some of these lidar measurements and investigations were part of a bigger campaign or in-depth field investigation. Airborne lidars enable measurements over large areas with the same lidar in down looking or up looking configurations in less time than a ground-based or ocean-based mobility system can. They enable measurements in regions that are difficult or impossible to access in some situations. This particular feature can be seen in missions to the polar regions to measure PSCs and/or ozone. To ensure that measurements can be taken, airborne equipment can fly above weather systems. They can measure large-scale patterns because they can fly faster than air mass movements. The increased backscattering owing to higher atmospheric density with increasing range somewhat compensates for the inverse range-squared (1/r2) decrease in backscattered signal for down-looking devices. This is also true for spaceborne lidar. An up-looking lidar, like a ground-based lidar, suffers from a 1/r2 signal loss between near and far-field measurements, necessitating a very wide dynamic range. The lack of flying chances, the inherent costs of aircraft operation, and the somewhat increased system complexities due to space, window, and power limits are all downsides of airborne lidars. This is especially true for aeroplanes that fly at high altitudes, such as the ER-2. Aircraft platforms additionally limit the size of telescopic receivers that can be used, as well as increasing AC frequency, vibration, temperature, voltage, and G-force fluctuations. In the case of an airborne lidar, the issue of eye safety is magnified, and it must be properly addressed for all aerial trips. To provide an eye-safe level for viewers on the ground or in lower or higher-flying aircraft, power is reduced, height is increased, and output wavelength is changed.

History of Spaceborne LiDAR

During the early atmospheric uses of lidars, it became evident that a spaceborne lidar orbiting Earth would provide a massive science reward. It's no surprise, then, that NASA and ESRO (later ESA) formed groups to investigate the potential of lidar on satellite platforms in the 1970s and 1980s. The natural platforms for showcasing lidar capabilities were Spacelab and Shuttle, due to the hefty weight and high-power requirements of those early lidars.

Specific plans for building and flying a spaceborne lidar have typically been linked to a specific space project, such as NASA's Spacelab and Shuttle missions, or the European Space Agency (ESA), NASA, and Japan's NASDA's joint Earth Observing System (EOS). Technical feasibility studies for these programmes have been carried out by groups of scientists who specialise in lidar and/or atmospheric research.

The Atmospheric, Magnetospheric, and Plasmas in Space (AMPS) payload for Spacelab/Shuttle, Atmospheric Research Using Spacelab-Born Lasers, the Shuttle Atmospheric Lidar Research Program, the ESA Space Laser Applications and Technology (SPLAT)Workshop, the Lidar Atmospheric Sounder and Altimeter (LASA) Instrument Panel Report, and the LASER Atmospheric Wind Sounder (The German ALEXIS Phase A research, as well as ESA studies of Laser Sounding from Space and the Atmospheric Laser Doppler Instrument, were among the later studies. The Proceedings of International Laser Radar Conferences (ILRCs) and archival publications are both excellent sources of knowledge on the use of lidars from space to study atmospheric composition and structure. These studies and papers have demonstrated the feasibility of using spaceborne lidars to

measure the height of the planetary boundary layer, the vertical distribution of aerosols, clouds, and trace gases like ozone and water vapour, tropospheric winds, and the vertical distributions of atmospheric pressure and temperature, with varying degrees of resolution and sophistication. For many of these characteristics, lidars give a novel measurement capability, greater vertical resolution for others, and daylight capability for others. With today's technology, all of these measurements are possible to some extent.
The EOS was created with the goal of better understanding the underlying, global-scale processes that govern the Earth's environment. Beginning in the late 1990s and continuing into the first decade of the new millennium, a series of polar and low-inclination platforms were to be flown. Two generic EOS lidar facilities were requested: LASA for a US platform and ATLID (Atmospheric Lidar) for an ESA platform. Members of the LAWS research facility instrument team for the Japanese platform were also invited to submit proposals. A number of intriguing and well-thought-out concepts were submitted. LASA's installation, on the other hand, has been postponed by NASA. The LAWS activity was approved for Phase B research, but it was later deselected in 1994. Phase B studies for the ESA facility were also carried out, but no authorisation to proceed was given.
Similarly, the Experimental Lidar in Space Experiment (ELISE) was being developed by NASDA for flight aboard the Mission Demonstration Satellite 2. (MDS-2). The MSD-2 mission was cancelled, however, due to the H-II rocket's launch failure in November 1999. As a result, no long-duration flights were permitted or performed in the late twentieth century. However, during its approved ice altimetry mission, the modification of the EOS research facility instrument GLAS (Geoscience Laser Altimeter System) was investigated with the goal of performing some atmospheric measurements such as cloud top and height of the planetary boundary layer (PBL) determinations. Furthermore, in the first decade of the twenty-first century, ESA opted to pursue spaceborne lidars, as will be discussed later in this unit.

Parameters of LiDAR system
LiDAR Parameters
1. Detection Range
2. Field-of-View (FoV)
3. Scan Pattern
4. Cross Talk Immunity
5. Detection Rate
6. Multiple Returns
7. Range Precision and Accuracy

- *Detection Range*: The detection range is perhaps the most important of the many LiDAR metrics by which the sensor is evaluated. It refers to the greatest distance at which an object can be detected. This is mostly determined by the laser source's power; the higher the power, the further an object may be detected. Eye safety standards limit the maximum laser power that can be used. Other elements that affect range include LiDAR parameters like laser type and aperture size, as well as reflecting object qualities like size, distance, reflectivity, and diffusion or specular reflection, as well as environmental effects like weather and temperature.

  In an intrusion detection system, for example, a wide detection range is critical. A LiDAR can be used to detect each object that enters a predetermined area, and it can be mounted along a wall or fence to do so. The algorithms also enable for object

classification, allowing for alerts to be sent only if certain criteria are met. However, the detection range of LiDAR can be a limiting factor.
If the wall is exceptionally lengthy and the detection range of the LiDAR is limited, additional LiDAR sensors would be required to cover the entire perimeter for foolproof protection. As a result, a long-range LiDAR is desirable in order to make the system more manageable and economically viable.

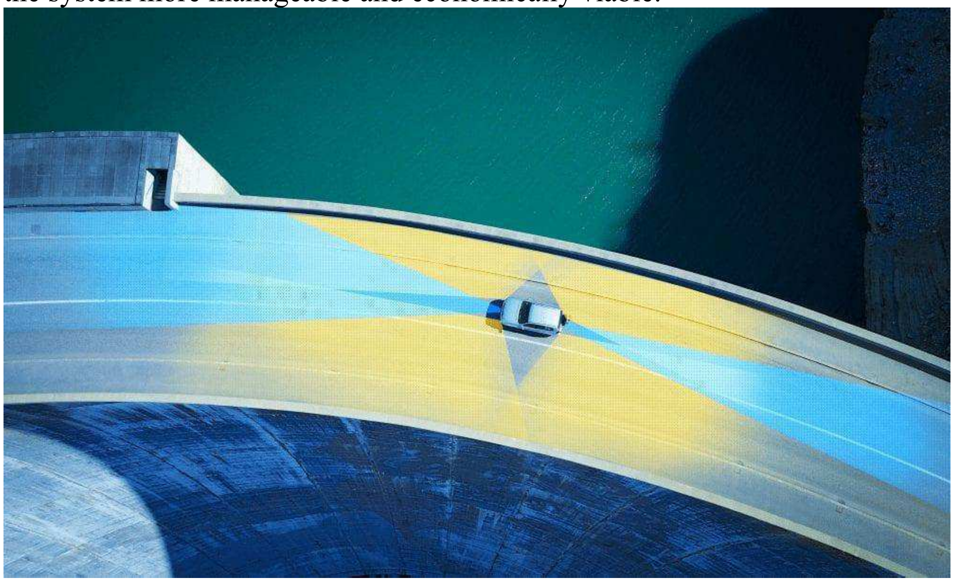


Figure: 3
Source: https://www.blickfeld.com/blog/understanding-lidar-performance/

"The Detection Range of a LiDAR describes the farthest distance where it can detect an object."

The elements that influence the range can be loosely divided into three categories:
- LiDAR characteristics (e.g., laser type, laser source strength, aperture size)
- external factors (e.g., rain, fog, snow, sunlight)
- object attributes (size, distance, reflectivity, diffuse or specular reflection, etc.)

For a MEMS-based LiDAR, Blickfeld's Cube has an unusually long range. This is owing to the mirror's proprietary design, which has a diameter of more than 10 millimetres and thus a big aperture.
This permits a large fraction of reflected photons to be directed onto the photodetector, allowing the Cube to consistently detect even weakly reflecting objects from up to 250 metres away.

- *Field-of-View (FoV)*: The angle covered by the LiDAR sensor, or the angle at which LiDAR signals are transmitted, is known as the field-of-view. Depending on the LiDAR technology, it varies greatly. Spinning LiDARs, for example, generate their field of view by rotating the 16 to 32 stacked laser sources mechanically, allowing them to provide a 360-degree view of the surroundings. Solid-state scanning LiDARs use fewer lasers, in Blickfeld's case only one, that strike one point at a time, making them less complex and hence less expensive while also being more robust. The beam is deflected, or "scanned," to illuminate the area of view point-by-point.

When it comes to customising the field-of-view, the Blickfeld LiDAR has a lot of options. Assuming that the number of laser signals emitted per scan cycle remains constant, decreasing the angle, and therefore the vertical FoV, results in a denser point cloud, whereas raising the vertical FoV results in the LiDAR returns being spread farther apart. The point spacing for the horizontal FoV can be altered while keeping the FoV the same.

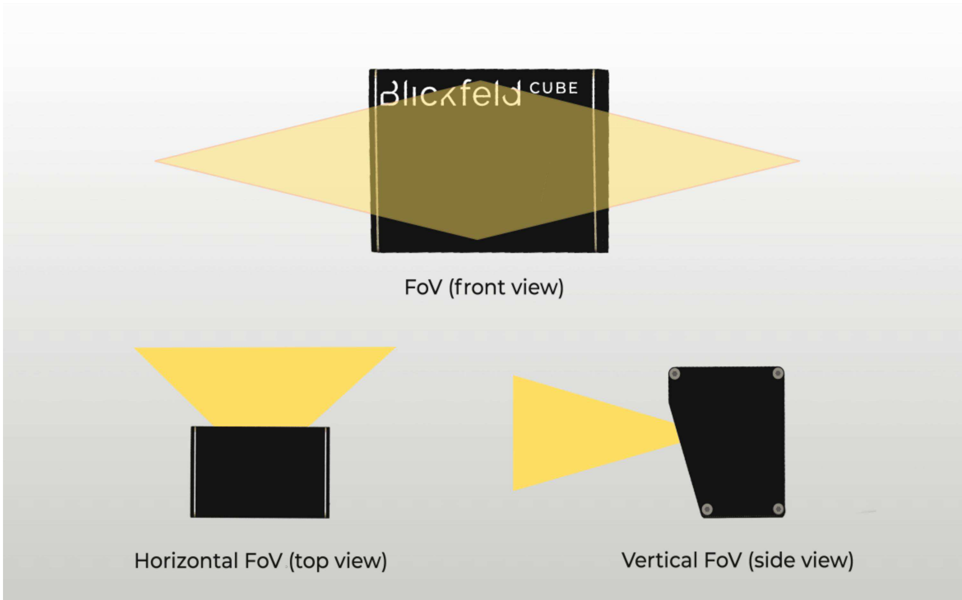


Figure: 4
Source: https://www.blickfeld.com/blog/understanding-lidar-performance/

"The field-of-view is the angle in which LiDAR signals are emitted"

Depending on the LiDAR technology, the FoV changes greatly. FoV requirements vary depending on the application's objectives as well as a variety of other parameters like the sort of objects to be scanned and their surface qualities.
FoV requirements vary depending on the application's needs as well as a number of other criteria such as the sort of objects to be scanned and their surface qualities. A dense forest canopy, for example, requires a larger field of vision to collect returns from the ground. Application situations in cities with tall buildings and small streets, on the other hand, will choose a narrower angle to obtain returns from the street level.

- *Scan Pattern:* The scan pattern is the most significant and intriguing of the many LiDAR factors to consider while scanning LiDARs. Beam deflection devices or scanner units are used in scanning LiDARs to deflect the laser beam in different directions for range measurements, resulting in unique patterns in the point cloud. Different properties, such as the number of scan lines or point density, can be seen in these patterns. The elements of the scan pattern can be crucial depending on the application for which the LiDAR will be employed. For example, in a people counting application, a high-resolution point cloud can be critical depending on the number of people present in a given location. A large number of scan lines is necessary to obtain the requisite resolution.

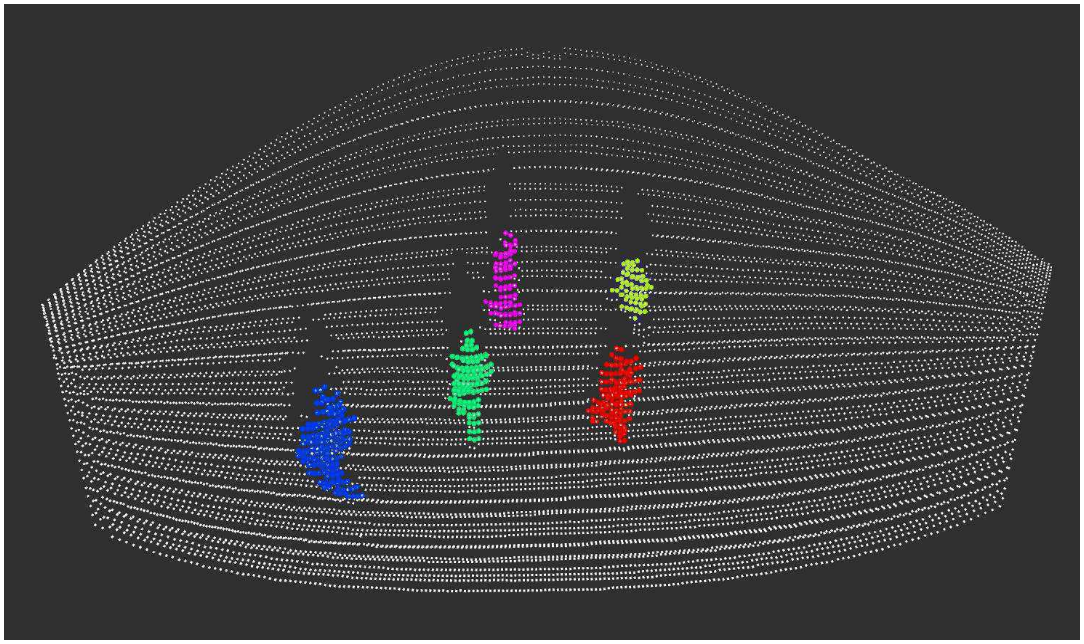


Figure: 5
Source: https://www.blickfeld.com/blog/understanding-lidar-performance/

The scan lines that make up the scan pattern can be easily changed, even while utilising the LiDAR, which makes Blickfeld's sensors unique. The sensor can be customised to meet the application's specific requirements, such as switching from a general view to a high-resolution view by adjusting the scan line density.
Scanning LiDARs deflect the laser beam in various directions to produce range measurements, resulting in scan patterns, which are unique patterns in the point cloud. These patterns exhibit a variety of features that allow for a variety of uses.

- *Cross Talk Immunity:* Cross-talk immunity is another important LiDAR property in many real-time applications where several sensors may present at the same time. For example, within an autonomous vehicle's range of view, the LiDAR sensor could take up laser signals from another vehicle's LiDAR sensor, resulting in false detection. As a result, a road impediment, for example, could be misidentified, resulting in unnecessary and potentially dangerous emergency braking. If the parasitic echo is strong enough, it could divert the LiDAR's attention away from an object, posing an even greater safety risk.

Sunlight is also a significant obstacle because it generates noise and decreases the signal's dependability as well as the detecting range. There are numerous approaches for ensuring that LiDAR sensors are not affected by cross-talk. Two of them have been used by Blickfeld:

Spectral filtering: The Cube has a wavelength of 905 nm, which is close to the infrared part of the electromagnetic spectrum. In front of the detector is a filter that allows only electromagnetic waves of similar wavelengths to pass while blocking others. As a result, the Cube is unresponsive to lasers from other LiDARs operating at different wavelengths, such as 1550 nm. The detector, however, is still susceptible to being fooled by signals from a 905 nm laser. A technique known as spatial filtering is used to combat this.

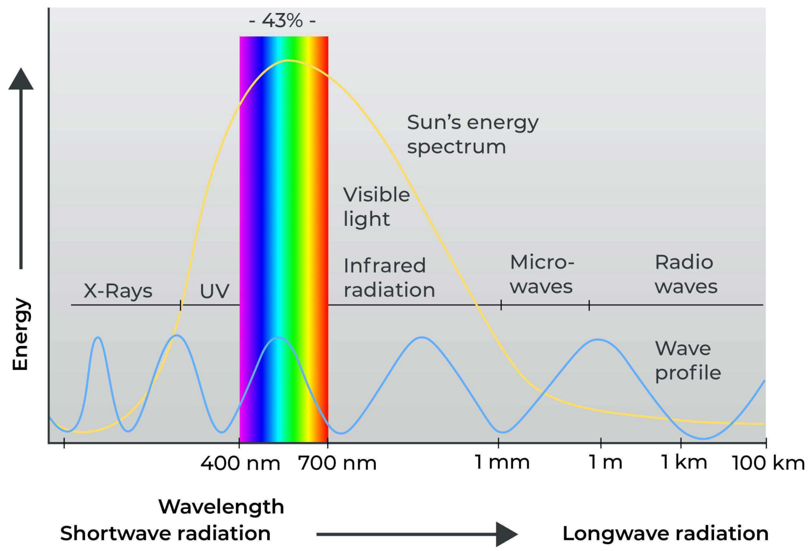


Figure: 6
Source: https://www.blickfeld.com/blog/understanding-lidar-performance/

Spatial filtering: The Blickfeld Cube's coaxial design allows it to receive a laser's reflect ion along the same path as the emission through the beam deflection device. This guarantees that the detector only captures photons that are transmitted in a single direction and is 'blind' to photons that are sent in other directions.
In order for the Cube to detect and create a false echo, another LiDAR sensor on an autonomous vehicle would have to emit a laser pulse at precisely the same angle but in the opposite direction, as well as the proper time window. The probability of this depends on a number of factors, including the distance between the LiDARs, scanning speeds, beam divergence, and relative orientation, all of which make the error exceedingly rare.
Spectral and spatial filtering, for example, can assist reduce sunlight noise and lessen the influence on range performance, which is a significant difficulty for LiDAR sensors. The detection range of the Cube is measured at a backlight level of 100 klux.
Crosstalk happens when numerous sensors give signals to a LiDAR that isn't the sender, resulting in false echo detection.
Crosstalk can be reduced using techniques like spectral and spatial filtering.

- *Detection Rate*: The detection rate (DR) or true-positive rate (TPR) is the percentage of frames in which a real target point is detected. The false-positive rate (FPR) on the other hand, estimates the percentage of frames in which an echo in the point cloud is identified despite the absence of a physical target.False detections are undesirable because they degrade the precision of the point cloud and hence the object recognition's dependability. In our intrusion prevention example, if the LiDAR detection rate is low, there would be a higher number of false alarms in surveillance applications, making the results untrustworthy.Blickfeld Cubes have a detection rate of

90%, with a false positive rate of less than 0.2 percent. This makes it particularly beneficial in applications where accuracy and consistency of outcomes are critical.

The detection rate (DR) or true-positive rate (TPR) is the percentage of frames in which a real target point is detected. False detections are undesirable because they degrade the precision of the point cloud and hence the object recognition's dependability.

- *Multiple Returns*: After sending out the beam, LiDARs typically get many reflections as the beam widens or diverges over time. As a result, while portion of the beam may strike the closest target, some of it may hit a target further away. This means that the beams will return at different times, resulting in several returns being recorded. If the LiDAR can only evaluate one return, it will only show one target, which is determined by the algorithm or the intensity of the reflection. Only the target nearest to the sensor is usually recorded, leaving the target behind unnoticed. LiDARs with multiple return capability can identify data on targets that are partially obscured by objects. With the same number of transmitted laser pulses, the amount and depth of data gathered increases.

  A beam sent by LiDAR usually widens or diverges as it travels further, and it may hit numerous targets, resulting in various returns at different times. LiDARs with multiple return capabilities may detect partially obscured targets, increasing the amount and depth of data collected.

- *Range Precision and Accuracy:* Range precision and accuracy are critical LiDAR metrics to differentiate, as they are frequently confused with one another.
  1. Range Precision: Precision is a metric for how repeatable LiDAR parameters are. Repeated measurements of the same target with high precision will be extremely near to the mean value, whereas low precision will result in a huge dispersion of data around the mean. For applications like speed-camera readings, where the vehicle's speed must be computed using the distance between the LiDAR and the moving target in a short time span, range precision is critical.Range precision is determined by the distance between the sensor and the target, as well as the target's characteristics, such as reflectance and angle of attack. The range precision of Blickfeld's Cube is less than 2 cm. Precision is a metric for how repeatable LiDAR parameters are. Repeated measurements with high precision are close to the mean value.
  2. Range Accuracy: Accuracy refers to how near a measurement is to the true value, or how close the measured target distance is to the true distance. Distance values should be very close to the actual distance and within the required range accuracy for a range accurate LiDAR. LiDAR range accuracy would be beneficial in applications that require absolute distance measurements, such as volumetric measurements. High precision will be crucial in recognising the topography beneath when utilised in drones to generate elevation maps. This information can be used to generate 3D models of the crops and track droughts or different stages of growth in order to optimise water use. Accuracy refers to how close a measurement is to the true value; for example, a range of accurate LiDAR distance values would be very close to the true distance.

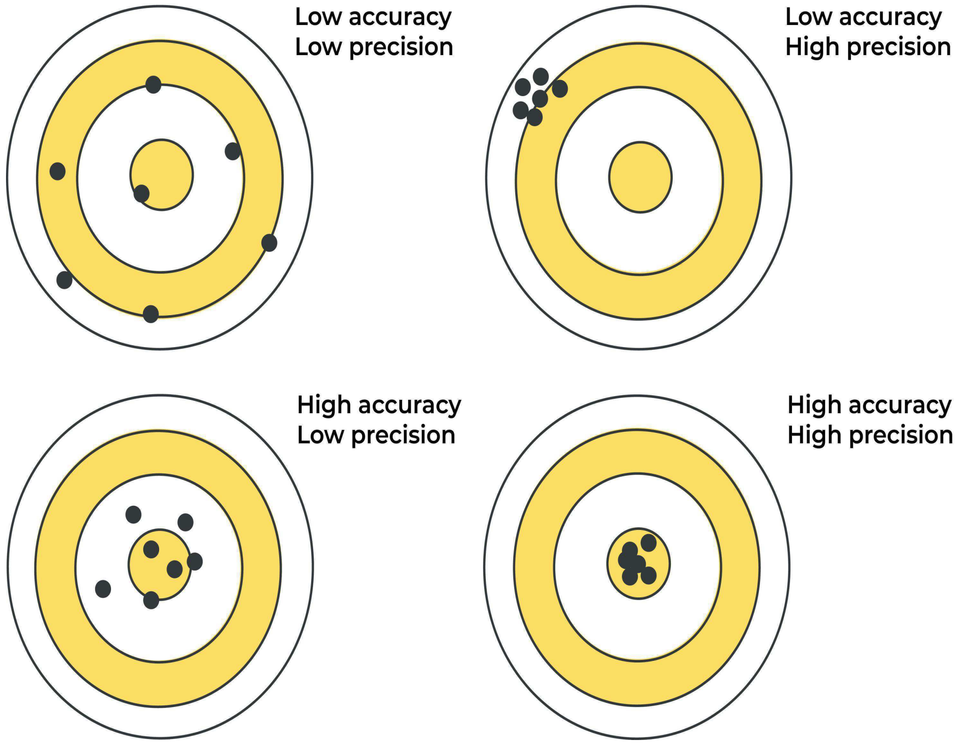


Figure: 7

Source: https://www.blickfeld.com/blog/understanding-lidar-performance/

1.4 SUMMARY

A LiDAR system estimates the heights of objects on the ground using a laser, GPS, and an IMU. Waveforms are used to generate discrete LiDAR data, with each point representing a peak energy point along the returned energy.Discrete LiDAR points have three values: x, y, and z. The height is determined by the z value.Various approaches can be used to estimate tree height and canopy cover using LiDAR data.

Because LiDAR is still being effectively employed in a variety of applications, understanding the many forms of LiDAR gives consumers an advantage when choosing the proper system. LiDAR systems are grouped into several different varieties based on their numerous features, technology, setup, and applications. LiDAR systems can be employed in a variety of ways, from the air to the ground, and from mobile to static. LiDAR has the potential to be employed in a wider range of applications in the future, thanks to advancements in technology such as differential absorption and inelastic scattering. LiDAR has proven itself to be invaluable to the growing field of autonomous mobility and that growth will make these powerful sensors more and more available over the coming years. But autonomous mobility is just one application and LiDAR shows promise to be able to do so much more.

1.5 GLOSSARY

- *Accuracy:* The degree to which an estimated value (for example, measured or computed) is close to a standard or acknowledged (actual) value of a given quantity.
- *GeoTiff:* GeoTIFF is a TIFF 6.0 compliant raster file format extension for storing georeference and geocoding information by linking a raster image to a specified model space or map projection. TIFF 6.0 is a GeoTIFF file. To specify projection types,

coordinate systems, datums, ellipsoids, and other things, GeoTIFF employs numerical codes.

- *ASPRS:* The American Society for Photogrammetry and Remote Sensing (ASPRS) is a photogrammetry and remote sensing learned society based in the United States. The International Society for Photogrammetry and Remote Sensing has a member organisation in the United States.
- *Raster:* When you think about images, you usually think of raster (or bitmap) images. When scanning or photographing an object, these are the types of images that are created. Raster images are made up of pixels, or tiny dots, each with its own colour and tone information, which are combined to produce the image.
- *Metadata:* Any information that is descriptive or supportive of a geospatial dataset, such as formally structured and formatted metadata files (for example, Federal Geographic Data Committee [FGDC] metadata formatted in extensible Markup Language [XML]), reports (collection, processing, quality assurance/quality control [QA/QC]), and other supporting data (for example, survey points, shapefiles).
- *EDMI:* EDMI instruments are highly dependable and practical surveying instruments that can measure distances of up to 100 kilometres. Engineer Supply's EDM equipment delivers dependable, accurate distance measurements that are displayed on an easy-to-read digital screen.
- *Aerosols:* Aerosols are very small particles suspended in the air. When these particles are large enough, we can see them because they scatter and absorb light. Because of the scattering of sunlight, vision is reduced (haze) and sunrises and sunsets are reddened.
- *Oceanography:* Oceanography, often known as oceanology, is the study of the ocean from a scientific standpoint. Ecosystem dynamics, ocean currents, waves, and geophysical fluid dynamics; plate tectonics and the geology of the sea floor; and fluxes of diverse chemical compounds and physical qualities inside the ocean and across its limits are all issues covered by this essential Earth science.
- *Meteorology:* Meteorology is a field of atmospheric science that studies the atmosphere and its phenomena, particularly weather and weather forecasting.
- *GIS (geographical information system):* Computer programmes that acquire, store, alter, analyse, and display spatial data are part of a system of geographically referenced information.
- *Resolution:* The smallest unit that a sensor can detect or that a raster DEM represents is called resolution. The precision with which a measurement can be taken. The linear size of an image pixel or raster cell is frequently referred to as "resolution."
- *Electromagnetic radiation:* Electromagnetic radiation (EMR) is made up of electromagnetic (EM) field waves that travel over space and convey electromagnetic radiant energy. [1] Radio waves, microwaves, infrared, (visible) light, ultraviolet, X-rays, and gamma rays are all examples of electromagnetic radiation. The electromagnetic spectrum includes all of these waves.

1.6 ANSWERS TO CHECK YOUR PROGRESS

1. *What does lidar stand for?*

   Lidar is a technique for determining ranges (varying distance) that involves using a laser to target an object and measuring the time it takes for the reflected light to return to the receiver. Due to variances in laser return durations and changing laser wavelengths, Lidar may also be used to create computerised 3-D renderings of places on the earth's surface and ocean bottom. It has uses on the ground, in the air, and on mobile devices.

2. *What are different LiDAR parameters?*
LiDAR Parameters
    - Detection Range
    - Field-of-View (FoV)
    - Scan Pattern
    - Cross Talk Immunity
    - Detection Rate
    - Multiple Returns
    - Range Precision and Accuracy

3. *What are the advantages of LiDAR system?*
LiDAR has the following advantages:
    - all data is geo-referenced from the start
    - high precision - ability to cover huge regions quickly
    - faster turnaround, less labour intensive, and cheaper costs than photogrammetric technologies
    - can gather data in steep terrain and shadows - can output DEM and DSM

4. *What are the disadvantages of LiDAR system?*
LiDAR disadvantages include the difficulty to penetrate dense canopy, which causes elevation model mistakes. It is difficult to read and handle very huge datasets. It has no worldwide protocols.

5. *What is the difference between LIDAR and RADAR?*
Despite the fact that the primary objective of LiDAR and RADAR is the same – to detect the existence and volume of distant objects – it is critical to understand the differences between the two technologies. Light Detection and Ranging (LiDAR) is a remote sensing technology that uses light. The premise underlying LiDAR in the instance of yellow scan is simple: shoot a tiny infrared laser beam onto a surface and time how long it takes for the laser to return to its source. It is feasible to obtain a point cloud of the environment using a LiDAR with a 360° viewing angle (for example, using a revolving mirror). Then, using specialised software, a 3D image is created that accurately reproduces the shape of the LiDAR in space.
The RADAR system operates similarly to LiDAR, with the exception that it employs radio waves rather than laser or LED light. It uses a rotating or fixed antenna to transmit radio waves and measures the time of flight of the reflected signal. RADAR has a wavelength range of 30 cm to 3 mm, whereas LiDAR has a wavelength range of micrometres (Yellow scan LiDARs work at 903 and 905 nm). The RADAR's wavelength allows it to identify objects from a considerable distance and even through fog or clouds. However, the antenna's size limits its lateral resolution. At a distance of 100 metres, conventional RADAR has a resolution of several metres. LiDAR is a small solution that allows for great precision in 3D mapping. Yellows can LiDAR systems have a resolution of a few millimetres at a distance of 100 metres. LiDAR is utilised for laser altimetry and contour mapping for this reason. Radar, on the other hand, is utilised for air traffic control, aircraft anti-collision systems, and radar astronomy.

1.7 REFERENCES

- https://en.wikipedia.org/wiki/Lidar
- https://www.geospatialworld.net/blogs/what-is-lidar-technology-and-how-does-it-work/
- https://oceanservice.noaa.gov/facts/lidar.html
- https://bit.ly/3rwxKJV
- https://desktop.arcgis.com/en/arcmap/latest/manage-data/las-dataset/types-of-lidar.htm
- https://www.blickfeld.com/blog/understanding-lidar-performance/
- https://lidarradar.com/info/the-different-types-of-lidar-systems

1.8 TERMINAL QUESTIONS

- Which nations (particularly in the EU) offer a free LIDAR data download and use service?
- Could you give a brief description of your national LIDAR datasets?
- What are the low-cost LiDAR sensors for 3D mapping that are available in India?
- What is the accuracy of open source lidar topography and bathymetry processing software?
- What is the typical LiDAR data accuracy?
- What advantages does LiDAR have over most other data collection techniques?

UNIT 2 : DATA PROCESSING- GEOMETRIC CORRECTION, DATA QUALITY ENHANCEMENT, FILTERING LIDAR, MAPPING APPLICATIONS

2.1 OBJECTIVES

After reading this unit you will be able to know about:

- Data Processing & Information
- Geometric correction
- Filtering LIDAR

2.2 INTRODUCTION

Any computer procedure that turns data into information is known as data processing. It is common to presume that the processing is automated and takes place on a mainframe, minicomputer, microcomputer, or personal computer. Data processing systems are often referred to as information systems to emphasize their practicality because data is most valuable when well-presented and really informative. Both phrases are approximately equivalent, performing similar conversions; data-processing systems often convert raw data into information, and information systems typically take raw data as input to produce information as output.

2.3 DATA PROCESSING- GEOMETRIC CORRECTION, DATA QUALITY ENHANCEMENT, FILTERING LIDAR, MAPPING APPLICATIONS

Geometric correction (also known as geo correction) is the process of transforming the X and Y, Dimensions of a remotely sensed image so that spatial distortions in the original image are eliminated or minimized and the output X and Y dimensions correspond to a predetermined geographic reference system. Latitude and longitude are examples of geographic reference systems, as are the Universal Transverse Mercator projection system and the British Ordnance Survey National Grid. Geometric correction is carried out to:

1. Two or more remotely sensed images are mosaiced together to form a single composite image.
2. Compare two or more remotely sensed photos from different periods of the same region.
3. Find interesting points and characteristics on the mathematically corrected image.
4. Within a GIS, combine a remotely sensed image with data from other sources, such as national park map layers or population census data.
5. To precisely determine distance and area from the geo corrected image.

Geometric adjustment is especially important because satellite flight trajectories do not always correspond with true north or the grid orientation of most geographic reference systems (see figure below).

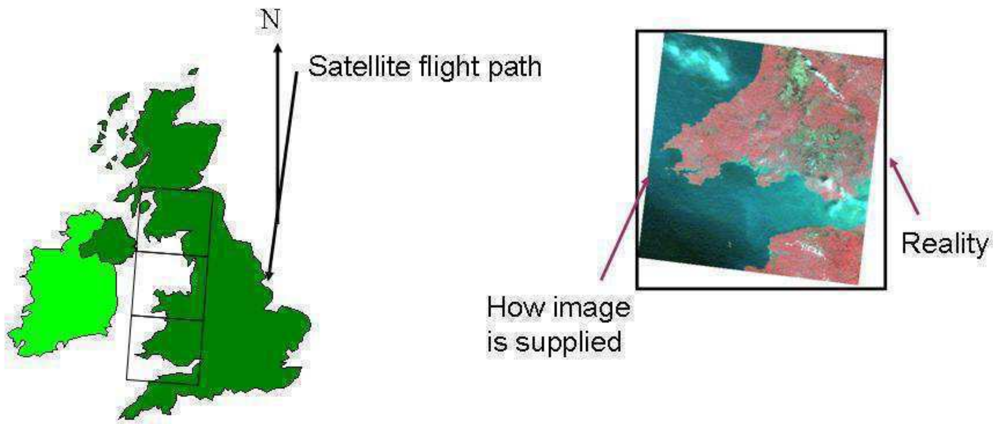


Figure: 2.1
Source: https://www2.geog.soton.ac.uk/users/trevesr/obs/rseo/FlightPath.JPG

Often, satellite imagery is provided in a format that has already been geometrically corrected. However, satellite imagery is frequently offered without any geometric modification. In either scenario, understanding how geometric rectification occurs is beneficial.
Using ground control points is one method of performing geometric rectification (GCPs). These are precise places that can be identified on remotely sensed imagery as well as in the target geographic reference system. The coordinates of ground control points in the target reference system can be determined either by conducting a GPS survey on the ground or by detecting GCP positions on a previously georeferenced second image or digital map layer. GCPs (represented as red dots) on a SPOT satellite image that may also be identified on a digital topographic map of the same area are depicted below.

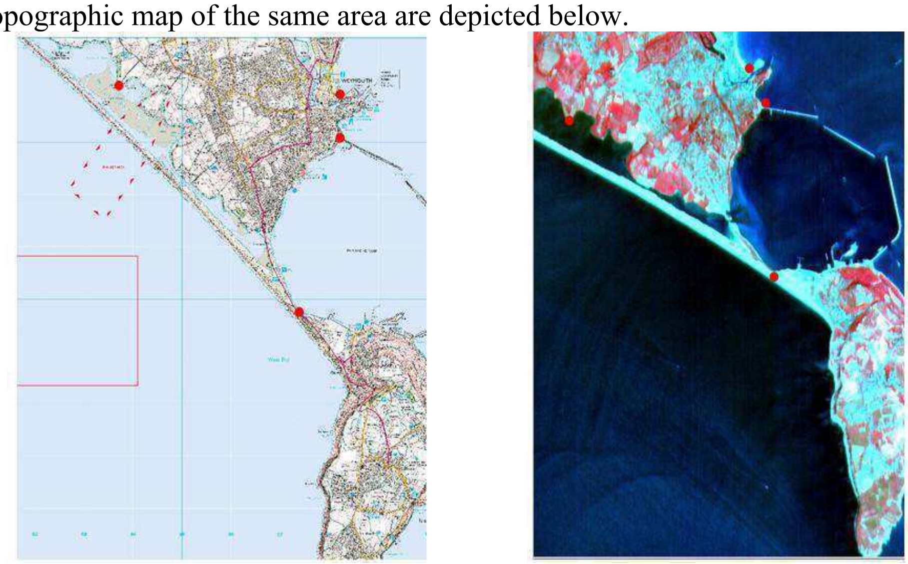


Figure: 2.2
Source: https://www2.geog.soton.ac.uk/users/trevesr/obs/rseo/SelectingGCPs.JPG

To geo correct an image, at least 20 ground control points should be gathered (more points are required for some more advanced correction techniques), and these points should be evenly spread around the image.

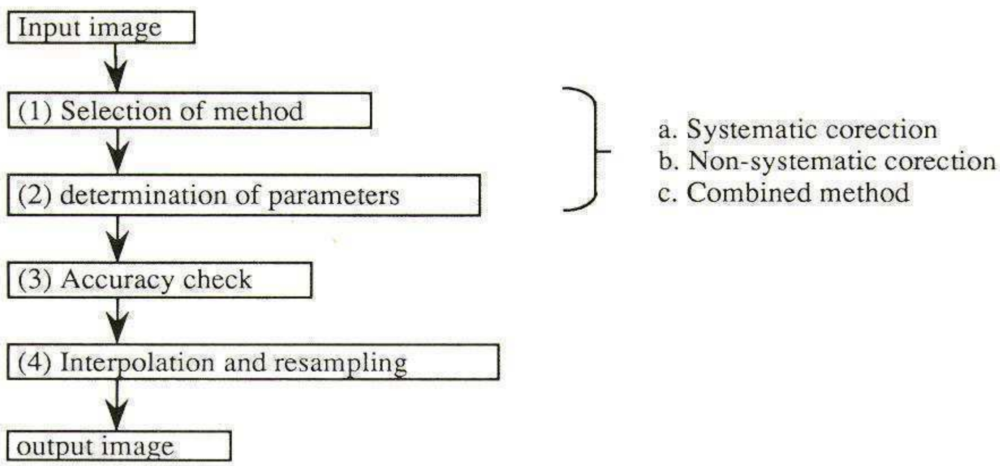


Figure 9.4.1 The flow of geometric correction

The steps to geometric correction are as follows:

(1) Method selection: A suitable method should be chosen after taking into account the geometric distortion's properties as well as the available reference data.

(2) Parameter determination: Calibration data and/or ground control points should be used to determine unknown parameters that define the mathematical equation between the image coordinate system and the geographic coordinate system.

(3) Check for precision: The geometric correction's accuracy should be examined and validated. If the accuracy falls short of the standards, the method or data utilised should be examined and corrected to avoid errors.

(4) Resampling and interpolation: The geo-coded image should be created using the resampling and interpolation technique. Geometric correction can be accomplished in three ways, as detailed below.

a. Systematic correction

When geometric reference data or sensor geometries are provided or measured, geometric distortion can be theoretically or systematically avoided. The geometry of a lens camera, for example, is defined by the collinearity equation with calibrated focal length, parameters of lens distortions, coordinates of fiducial markers, and so on. Tangent correction for an optical mechanical scanner is a sort of system correction. In general, systematic correction is sufficient to eliminate all errors.

b. Non-systematic correction

The least square method will be used to determine polynomials to transform from a geographic coordinate system to an image coordinate system, or vice versa. The accuracy is determined by the order of the polynomials, as well as the number and distribution of ground control points.

c. Combined method

The systematic correction is applied first, followed by the reduction of residual errors using lower order polynomials. The purpose of geometric correction is often to acquire an error within one pixel of its real position.

Data quality Enhancement

LiDAR systems have become the preferred technology for collecting topographic data for a variety of applications, including environmental monitoring, contour mapping, transportation planning, emergency management, military simulation, ortho-rectification, oil and gas exploration, mining, hydrology, shoreline management, 3D city modelling, and forest mapping, thanks to the direct acquisition of a high density and accurate 3D point cloud. A LiDAR system comprises of a laser ranging and scanning unit and a Position and Orientation System (POS), which includes an integrated GNSS and an Inertial Navigation System (INS). The laser ranging unit estimates the distances between the sensor and the mapped surface, while the platform's location and orientation are determined by the onboard GNSS/INS component.

The ground coordinates of the point cloud are calculated using the laser scanner measurements, position and orientation information, and mounting parameters (i.e., parameters specifying the spatial and rotational offsets between the system 2 components). A calibration technique is used to determine the mounting settings. Lidar systems, like any other spatial data gathering, require operations aimed at verifying and evaluating the quality of the resulting data. The term "Quality Assurance – QA" is used in this document to refer to pre-mission operations aimed at assuring that a process will offer the quality required by the user. On the other hand, the phrase "Quality Control – QC" refers to post-mission procedures for assessing the finished product's quality. QA is primarily concerned with the calibration, planning, implementation, and review of data gathering operations, as well as the creation of management controls. A slower speed is recommended in forested areas.

To improve the point density and have more pulses penetrate the ground, a smaller scan angle, a greater overlap percentage, and/or a lower flying height may be required. Another significant QA item is the selection of the proper mission time based on the GNSS satellite constellation distribution. For example, during the survey, at least 4 well-distributed satellites with elevation angles greater than 15° are often required. Furthermore, the aircraft should maintain a predetermined distance from the GNSS base station. For LiDAR system calibration, original observations (GNSS, IMU, and laser measurements) or at least the trajectory and time-tagged point cloud should be available.

To illustrate what is meant by quality control activities, consider the well-established photogrammetric techniques for evaluating the finished product's internal/relative and external/absolute accuracy.

We commonly utilize the a-posteriori variance factor and the variance-covariance matrix resulting from the bundle correction technique to evaluate the internal/relative quality (IQC) of the results from a photogrammetric reconstruction exercise. Checkpoint analysis employing independently measured targets is typically used for the external/absolute quality (EQC) evaluation. Standard photogrammetric IQC measures are not achievable since the LiDAR point cloud is not computed using redundant data that are altered in an adjustment operation. Furthermore, the irregular and sparse character of the LiDAR point cloud complicates the EQC process. The LiDAR surface is compared to independently collected control points in a typical EQC approach (Hodgson and Bresnahan, 2004; Wotruba et al., 2005).

Unless carefully designed targets are used, this approach does not provide precise 3 verification of the horizontal quality of the LiDAR points, in addition to being expensive (Csanyi and Toth, 2007). Because the horizontal precision of LiDAR points is known to be inferior to the vertical accuracy of these points, this incapacity is a big disadvantage.

Another quality control activity is "Strip Adjustment – SA," which detects and mitigates the impact of systematic differences among overlapping LiDAR strips. As a result, strip adjustment goes beyond quality control in that it is intended to not only detect but also reduce or eliminate the influence of biases in system settings and/or measurements on the point cloud by enhancing the compatibility between overlapping LiDAR strips.

A method for decreasing vertical differences between overlapping strips is proposed by Crombaghs et al. (2000). This method does not address planimetric disparities, which may be bigger in magnitude than vertical discrepancies. For finding inconsistencies and enhancing the compatibility between overlapping strips, Kilian et al. (1996) devised an adjustment process comparable to the photogrammetric strip correction. The disadvantage of this method is that it relies on specific spots to connect overlapping LiDAR strips and control surfaces. The identification of distinct points (e.g., building corners) is challenging and unreliable due to the uneven form of the LiDAR points. Where planar features are employed in the strip adjustment technique, Kager (2004) suggests more suitable primitives. To determine the relationship between discrete points in one LiDAR strip and TIN patches in the other, Maas (2002) suggested a least-squares matching approach.

Rather than enhancing the compatibility between nearby strips, this research concentrated on discovering inconsistencies between conjugate surface elements. To increase the compatibility between overlapping LiDAR strips, Pfeifer et al. (2005) used segmented planar patches, while Vosselman (2004) used linear features.

Bretar et al., (2004) provided an alternate way for increasing the quality of LiDAR data by employing photogrammetrically derived surfaces. The fundamental disadvantage of this technology is that it is dependent on obtaining aerial imagery over the same area. Filin and Vosselman (2004) suggested a strip adjustment approach in which natural surfaces are used to estimate the 3D offsets between LiDAR strips, which are locally approximated by planes. In general, strip adjustment processes are limited by the fact that the transformation functions used to increase the compatibility of overlapping strips are generally empirical.Furthermore, the strip adjustment procedure does not employ the predicted transformation parameters to deduce what went wrong in the data collecting system (i.e., what caused the discrepancies?).

A thorough understanding of the sensor model is required for the establishment of effective QA/QC procedures for LiDAR systems and related data. More particular, the user community should understand how the system measurements and parameters are combined to create the point cloud's ground coordinates. We must also consider the effect of systematic and random mistakes in system settings and observations on point cloud coordinates. After gaining a thorough awareness of these concerns, one might choose from the following options:

1. What parameters should be taken into account during the calibration procedure? Such factors should be based on the potential for systematic errors in a LiDAR system, as well as the capability of estimating these values during the calibration operation (i.e., whether these parameters are correlated or not).

2. What are the various system calibration scenarios? Models based on system raw data, time-tagged point cloud coordinates, and the system trajectory, or only the point cloud coordinates for system calibration, for example, should be researched and/or created.

3. What is the best flying configuration for a precise calibration of the system? We must determine the suitable flight configuration (e.g., flight pattern, flight height, overlap %, scan angle, etc.) based on the calibration parameters analysis. The flight configuration should ensure that the calibration parameters are decoupled.

4. What is the best way to keep the system calibration process under control? Control points have traditionally been used to calibrate mapping systems (e.g., photogrammetric systems). This control is inappropriate for LiDAR systems. As a result, additional control information such as LiDAR control targets, control planar patches, and/or control surfaces should be examined.

5. To analyse and quantify systematic inconsistencies and noise levels in the acquired data, effective, practical, and cost-effective quality control processes should be created. These procedures should be simple enough for any LiDAR mission to be carried out by the end user.

To throw some light on these concerns, the remainder of this article will cover the LiDAR mathematical model, potential systematic and random mistakes, and their impact on the point cloud, as well as some discussion of current calibration and quality control techniques, as well as future prospects.

3. Error Budget for LiDAR

The quality of a LiDAR system's derived point cloud is determined by random and systematic mistakes in the system's readings and parameters.

The precision of the system's measurements, which include location and orientation readings from the GNSS/IMU, mirror angles, and ranges, determines the size of the random errors. Biases in the mounting parameters relating the system components, as well as biases in the system measurements (e.g., ranges, mirror angles, and drifts in the position and orientation information) are the main causes of systematic mistakes. The impact of random and systematic mistakes in system measurements and parameters on the reconstructed object space will be examined in the following sub-sections.

Random Error

The goal of researching the impact of random errors is to gain a thorough understanding of the type of noise in the derived point cloud, as well as the precision that can be achieved with a given flight and system configuration. The impact of random errors in system measurements is investigated in this paper using a simulation procedure that begins with a specified surface and trajectory and then proceeds to derive the system measurements (i.e., ranges, mirror angles, position and orientation information for each pulse). The system readings are then subjected to noise, which is then used to reconstruct the surface using the LiDAR equation. The influence of the introduction is shown by the disparities between the noise-contaminated and correct coordinates of the LiDAR sites.

• Noise in the derived point cloud will be identical to that in the position cloud. The noise in the point cloud coordinates as a result is unaffected by the flying height or scan angle.

• Orientation noise (attitude or mirror angles) will affect the horizontal coordinates more than the vertical coordinates for suitable scan angle values (e.g., within the range of 250). The effect is also affected by the flying height and scan angle.

• Range noise primarily affects the vertical component of the computed coordinates for reasonable scan angle values (e.g., within the range 250).

Standard Error

Biases in the mounting parameters relating the system components as a whole could generate systematic mistakes.as well as measurement biases in the system Systematic errors in the laser scanner (e.g., biases in the range and mirror angle measurements) as well as systematic errors in the derived GNSS/INS position and orientation (e.g., differential troposphere and ionosphere, multi-path, INS initialization and misalignment errors, and gyro drifts) are all examples of systematic errors in the system measurements.

4. LiDAR Calibration: Current Practices and Prospects

The ground coordinates of the laser points are determined using the measured ranges and mirror angles, as well as the direct geo-referencing information provided by the onboard GNSS/IMU. LiDAR data gathering is normally done in a strip-wise method. The mounting parameters between the system's components, such as bore-sighting angles and lever-arm offsets, are also required for the determination of the ground coordinates of the laser spots. The mounting parameters, as well as other system parameters, are determined through a calibration procedure, which is normally completed in many steps: Laboratory calibration, platform calibration, and in-flight calibration are the three types of calibration. Individual system components are calibrated at the laboratory calibration, which is performed by the system manufacturer. The eccentricity and misalignment of the laser mirror with the IMU, as well as the eccentricity of the IMU with the sensor reference point, are also determined. The eccentricity between the sensor reference point and the GNSS antenna is established during platform calibration. For the estimation of the LiDAR system characteristics, an in-flight calibration test field made up of control surfaces is used. The differences between the LiDAR-derived and control surfaces are used to fine-tune the mounting parameters and systematic errors in system measurements (mirror angles and ranges).

The following are some of the disadvantages of current in-flight calibration methods:

- They are time-consuming and costly;
- They are usually based on complex and sequential calibration methods;
- They necessitate some effort when surveying the control surfaces;
- Some of the calibration methods require manual and empirical procedures;
- There is no widely accepted methodology because calibration techniques are usually based on a manufacturer-provided software package and the expertise of the user.

5. Proposed LiDAR Data Quality Control Strategy
Standard metrics to evaluate the quality of the positional accuracy of LiDAR points are desperately needed by the user community. These metrics should be able to assess the points' relative and absolute quality (i.e., IQC and EQC). The degree of consistency among the LiDAR points in overlapping strips could be used as a potential IQC approach. The idea behind such a process is that in the absence of random and systematic mistakes, conjugate surface elements in overlapping strips should completely match. If consistent inconsistencies are found, the presence of biases in the system parameters and/or measurements can be inferred. Furthermore, after removing systematic discrepancies, the data's noise level can be assessed by evaluating the goodness of fit between conjugate surface elements.

Lidar filtering

- Lidar ground filtering

Three factors influence LiDAR point measurements: bare ground, above-ground objects, and noise.

$$M_{sensor} = E_{ground} + E_{non\text{-}ground} + M_{noise}$$

Where, M_{sensor} is the measurements from the LiDAR sensor, E_{ground} is the elevation of ground. $E_{non\text{-}ground}$ is the height of objects above ground. M_{noise} is the undesired measurements, such as the noise from sensors, airplanes, or birds.
In LiDAR data, the ground points are the measurements from bare-earth terrain that are usually the lowest surface features in a local area. Non-ground points are the measurements from the objects above the bare-earth terrain, such as trees, buildings, bridges, and shrubs. In order to appropriately identify ground points, it is important to understand the physical characteristics of ground points that differentiate them from non-ground points. Ground surfaces can be divided into four categories based on their physical characteristics.

Ground surface steepness: Surface slope is often lower between two adjacent bare ground points than between one bare ground and one non-ground point. As a result, many ground filters classify a point with a slope greater than the maximum ground slope as a non-ground point. The slope threshold value that distinguishes between ground and non-ground points will most likely differ depending on the surface type. A low threshold value, such as 30°, may be appropriate for relatively flat urban areas. Complex surfaces with steeper slopes, such as mountain terrain or high relief forest canopy surfaces, may necessitate a higher threshold to accurately identify ground from non-ground.

Ground Surface Homogeneity: Ground surfaces are relatively continuous and smooth. Trees and buildings are the main non-ground features that should be removed from the measurements. But trees are usually less smooth in texture than bare ground and buildings and may be removed based on morphologic characteristics.
These are the common characteristics that become assumptions of many ground filters. In some cases, ground surfaces may not have all of these characteristics, and filters may mislabel points as non-ground. For example, cliffs have sharp elevation changes and many filters fail to label them as ground.

Difficult Ground Features for LiDAR Ground Filtering
LiDAR ground filtering algorithms make different assumptions about ground characteristics to differentiate between ground and non-ground features. In reality, ground surfaces may have unique characteristics that confound certain ground filters depending on the location and terrain conditions. The following features often confound some ground filtering algorithms:

1. Shrubs, especially those below one meter
2. Short walls along walkways
3. Bridges
4. Buildings with different size and shape
5. Hill cut-off edges
6. Complex mixed covering
7. Areas combined with low and high-relief terrains
8. Lack of reliable accuracy assessment

Shrubs are common error points mis-labelled as ground surfaces. Shrub heights are usually less than one meter and so slope and elevation difference between the shrub and neighboring ground points are similar to those between ground points and neighboring ground points in variable terrain. Shrubs in urban areas are usually found as small patches dotting the ground. Similarly, short walls along steps, bridges, or fences in high relief areas, and other features often present in high-resolution urban LiDAR images are hard to distinguish from ground points. For certain filters (e.g., one-dimensional scanning filters), bridges may also cause difficulty because they are smoothly connected with ground surfaces.
Variable building sizes may cause problems for certain filters. Ground filters based on roving windows sometimes have difficulties removing large or small buildings. This difficulty arises because the filters separate points based on a comparison between the measured value and the estimated value within a certain neighbourhood. If a large building completely contains a roving window, the points in the middle of the building may not be recognized as part of the building since there is no difference between the measured and estimated values.
Jagged hilly edges are difficult to identify for many ground filters. Although many natural surfaces have no sharp change in slope or elevation, certain terrain features do, like cliffs, shores, and riverbanks. Filters based on slope may mislabel ground with slopes larger than the maximum ground slope threshold as non-ground points. Likewise, areas of mixed low and high relief are challenging for ground filters because of the difficulty of selecting appropriate slope and elevation thresholds. Furthermore, most urban environments have a complex matrix of surface features that confound ground filtering algorithms, including buildings, trees, shrubs, bridges, short walls, and even wires.

- Airborne LiDAR Filtering

During previous decades, airborne light detection and ranging (LiDAR) technology, which is an efficient and rapid remote sensing technology for collecting three-dimensional (3D) point clouds over a large area, has been widely employed in various fields, such as digital terrain model (DTM) forest ecosystem investigation, and 3D building modellingIn most applications, a filtering operation for separating point clouds into ground and non-ground points is a preliminary and essential step. Many filtering algorithms have been proposed to automatically filter ground points, and these algorithms can be mainly grouped into three categories, namely, slope-based methods], mathematical morphology-based methods and surface-based methods.

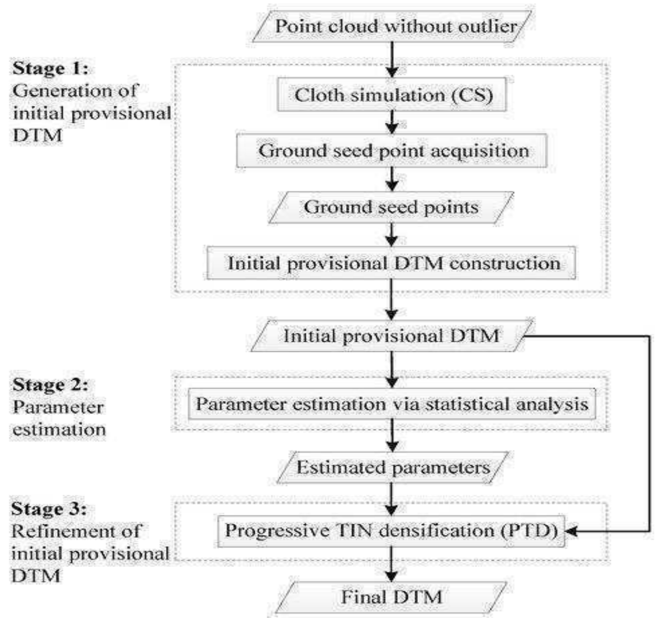
The basic principle of surface-based methods is to gradually approximate the bare earth using a parametric surface, such as triangulated irregular network (TIN) model, weighted linear least-squares interpolation model, active shape model, thin plate spline (TPS), and cloth simulation model. For this type of method, the ground seed points are obtained and then densified iteratively to create a provisional DTM that gradually refines the ground surface based on certain criteria (e.g., elevation). Among the surface-based methods, progressive TIN densification filtering (PTDF) is one of the typical methods that construct the initial TIN-based DTM (i.e., the initial provisional DTM) from ground seed points, which are the lowest points in each grid cell of an entire region dataset. The grid cell size (c) is usually larger than the size of the maximum non-ground object to minimize the influence of non-ground points. Then, remaining ground points are progressively detected from the unfiltered points to update the initial provisional DTM based on elevation and angle criteria. The filter has been widely used in the scientific community and the engineering field because of its promising performance.

The airborne LiDAR system is comprised of three major time-synchronized components:

- a laser scanner unit,
- a GPS,
- and an IMU.

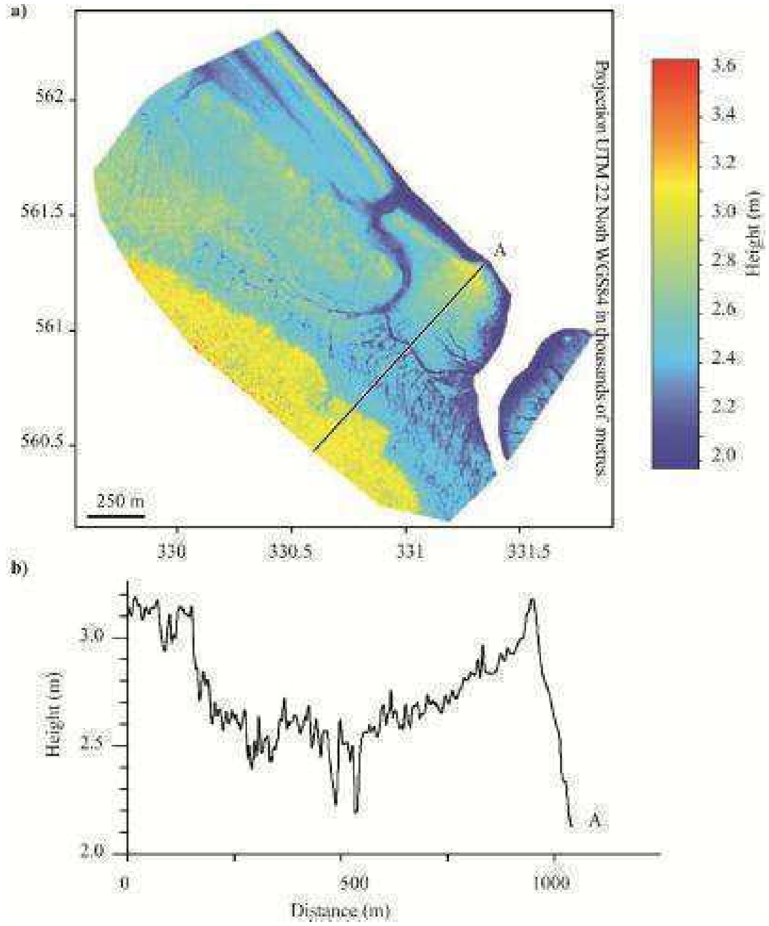
Methods

A series of critical improvements are implemented to increase the accuracy and practicality of CSF and PTDF. First, we use CS to select the ground seed points instead of using the lowest points in a user-defined grid cell. This algorithm provides more ground seed points that are almost evenly distributed in general, thereby generating a high-quality initial provisional DTM. Second, parameter thresholds (i.e., the maximum angle (θ) and the maximal terrain slope (s)) for PTD are derived from the initial provisional DTM based on the statistical analysis theory to improve the automation of the filter. Third, the initial provisional DTM is refined by PTD with adaptive parameter thresholds. The entire workflow is composed of three parts, i.e., generation of initial provisional DTM, estimation Oparameter the sholds and refinementof the initial provisional DTM.



Source: https://www.mdpi.com/2072-4292/11/9/1037/htm#fig_body_display_remotesensing-11-01037-f003

Airborne LIDAR data filtering technology combines the advantages of very high resolution remote sensing with the ability to cover very wide surfaces (several km2), Because it is based on ground measurements of an enormously higher number of locations than a total electronic station or a GPSLiDAR-derived DTMs are more exact than the DTMs generated by the other two approaches. The use of LiDAR in physical investigations of beaches is becoming more common.



Source:https://ars.els-cdn.com/content/image/3-s2.0-B978178548160450008X-f08-12-9781785481604.jpg

Mapping Application

LiDAR is a valuable technology for a variety of industries, ranging from forestry to self-driving cars. Vehicle speed tracking is one of the most prevalent applications for LiDAR. It is used by law enforcement authorities for this purpose as well as for accident investigations.Here are a some of the many LiDAR applications available today:

Agriculture - LiDAR technology can assist agriculture technology (AgTech) firms in identifying regions with optimal sunlight for more effective production. It can also be used to teach machine learning systems how to spot crops that want water or fertiliser.

Archeology - This technology has transformed the field of archaeology, allowing professionals to discover previously unknown structures all around the world. Two archaeologists from the University of Colorado are on a mission to use LiDAR to scan the entire earth.

Astronomy - LiDAR technology was utilised by NASA (National Aeronautics and Space Administration) to investigate Mars. They were able to make a topographic map and identify snow coming from the sky.

Climate change - LiDAR is a technique used by climate scientists to investigate and track changes in the atmosphere. German scientists have created an aerial LiDAR system that can track atmospheric gases and could possibly be used from space. Botanists use it to keep track of changes in forested areas. LiDAR is also used to calculate glacier changes over time.

Land management — When compared to aerial surveys, land management companies can monitor land resources in real-time, allowing for faster and more efficient mapping. They also utilise it in disaster planning, early warning systems, emergency response (for example, fighting forest fires), and location-based investigations.

Land mapping - NOAA employs LiDAR to build accurate coastline maps and digital elevation models for GIS (geographic information systems). They also use it to help in disaster relief operations.

Oil and gas exploration - LiDAR can identify tiny molecules in the atmosphere since it has a shorter wavelength than other technologies. Differential Absorption LiDAR (DIAL) is a new technique that aids in the detection of oil and gas deposits.

Meteorology - Since its inception, LiDAR has aided meteorologists in the study of clouds and their patterns by detecting microscopic particles in the cloud using wavelength. In meteorology, there are several different forms of LiDAR.

Renewable energy-LiDAR can be used to detect essential criteria for capturing solar energy, such as optimal panel positioning, using renewable energy. It's also used to calculate wind speed and direction so that wind farm operators can build and install turbines.

Robotics - LiDAR technology is utilised to provide mapping and navigation capabilities to robots. The technology is being used to train an autonomous system to recognise the distance between the vehicle and other objects in the surroundings for self-driving cars.

Tsunami modelling - LiDAR is used to inform systems that alert individuals to the possibility of a tsunami in their region. It's also used to figure out how high a coastline is and how deep it is underwater. Experts can anticipate which places will be most harmed by a tsunami using LiDAR data placed into the GIS.

Putting LiDAR Data to Work in Computer Vision

LiDAR data must be appropriately labelled to be helpful for computer vision, and more especially, supervised machine learning, which is a massive operation that can be difficult to scale. The difficulty for AI engineers is converting enormous amounts of unstructured data into structured data that can be utilised to train machine learning models. This necessitates hours of data labelling in order to train machines to read and grasp the visual environment.

2.4 SUMMARY

Lidar, also known as LiDAR, LADAR, aerial laser altimetry, or airborne laser swath mapping, is a widely used data layer for a wide range of applications. Topographic lidar data is often collected from aeroplanes and has profited from recent technological advancements. GPS and IMU technology advancements the method can generate a dense set of very relevant data. Over a vast area, precise elevation measurements are required.

The "raw" point data provided by lidar, as well as processed derivatives such as surfaces and curves (DEMs). Not only are elevation measurements included in point data, but so are Classification values, intensity values, and a variety of additional point properties are all available. DEMs are made up of If the points have been classified, they can represent bare ground or other surfaces. As an example, consider the first surface or the last-return surface. To process the data in a GIS with other data or maps, all of the data must have the same reference system. A geometrical rectification, also known as geo-referencing, is a method that assigns a spatial coordinate system to the content of a map (for example, geographical latitude and longitude)

To differentiate ground from non-ground features, current ground filters of lidar generally use four ground characteristics: lowest feature in a specified area, ground slope threshold, ground surface elevation difference threshold, and smoothness. Ground slope and ground elevation difference thresholds are often used in LiDAR ground filters. Some directional filtering methods refer to the label of prior points in a scan line as well as the elevation of the nearest ground points discovered in earlier iterations. Many iterative algorithms require a ground seed as a fundamental ground elevation reference to begin the filtering process, and researchers frequently use the lowest feature characteristic to find these ground seeds. When using such filters, additional attention should be paid to error filtering in order to exclude probable outliers with very low results. If this preprocessing step is skipped, the algorithms may fail to locate any ground points since the slope and elevation are significantly higher than the threshold.

2.5 GLOSSARY

- **Absolute Accuracy** - A measure that accounts for all systematic and random errors in a dataset. Absolute accuracy is stated with respect to a defined datum or reference system.
- **Accuracy** - The proximity of an estimated (measured or computed) value to a standard or accepted (actual) value of a specific quantity. Concerning the source data and the quality of DEM products.
- **Active sensor** - Sensor that generates the energy that it uses to perform the sensing.
- **Beam divergence** - An electromagnetic beam's beam divergence (for example, the laser used in LiDAR) is an angular measure of the increase in beam diameter or radius with distance from the optical aperture or antenna aperture from which the electromagnetic beam originates.
- **Boresight:** Calibration of a LiDAR sensor system equipped with an Inertial Measurement Unit (IMU) and Global Positioning System (GPS) to determine or establish the instrument's accurate position (x, y, z) with respect to the GPS antenna and orientation (roll, pitch, heading) withrespectto straight and level flight.
- **Control point (calibration point)** - A surveyed point is a point that is used to geometrically alter a dataset in order to determine its positional correctness in relation to the real world. Control points are distinct from checkpoints and should never be used as such on the same project.
- **Coordinates** – A group of 3D numbers that define a point in 3D space. Traditionally, a vertical coordinate would be defined as a 3D coordinate, that is, an x/y coordinate with an associated z-value.
- **Digital Elevation Model (DEM)** - A 3D computer-generated representation of the surface of a landscape. DEMs are commonly used in geographic information systems and serve as the foundation for digitally created relief maps.
- **Design Ground Point Density** — Calculated on hard surfaces, this is the planned/required ground point density (rooftops, roads, man-made structures of planar character).
- **A digital surface model (DSM)**-is a three-dimensional (3D) computer-generated representation of the earth's surface that contains all items on it.
- **Easting** - In a coordinate system, the distance east (positive) or west (negative) from a north-south reference line.
- **Inertial Navigation System (INS)** — A self-contained navigation system made up of many subsystems such as an IMU, a navigation computer, a power supply, an interface, and so on. Estimates velocity, position, and orientation based on observed accelerations and rotations. Due to gyro drift, an unaided INS loses accuracy over time.
- **Metadata** it is any information that describes or supports a geospatial dataset, including formally structured and formatted metadata files. For example, Federal Geographic Data Committee (FGDC) metadata in extensible Markup Language (XML).

- **Projected coordinate reference system** - A method used to represent the curved, 3D surfaces of the Earth on a 2D plane. Essentially, the conversion of location data from a sphere approximation to a planar surface (e.g., UTM).
- **Vertical accuracy** - The measure of the positional accuracy of a dataset with respect to a specified vertical datum, at a specified confidence level or percentile. Vertical accuracy is an indicator of quality for geospatial products.

2.6 ANSWER TO CHECK YOUR PROGRESS

- What is LiDAR?
- How are LiDAR data collected?
- Are LiDAR and Airborne Laser Scanning the same thing?
- What is the resolution of LiDAR data?

2.7 REFERENCES

- https://www.asprs.org/wp-content/uploads/2010/12/AKAM_LiDAR_Calibration.pdfhttps://interpine.nz/10-the-importance-of-lidar-data-quality-assurance/
- https://www.sciencedirect.com/topics/earth-and-planetary-sciences/airborne-laser-scanning
- https://www.mdpi.com/2072-4292/11/9/1037/htm
- http://sar.kangwon.ac.kr/etc/rs_note/rsnote/cp9/cp9-4.htm
- https://www2.geog.soton.ac.uk/users/trevesr/obs/rseo/geometric_correction.html
- https://www2.gov.bc.ca/assets/gov/data/geographic/digital-imagery/geobc_lidar_specifications_v40.pdf
- https://www.mdpi.com/2072-4292/2/3/833/htm

2.8 TERMINAL QUESTIONS

- What format do the data come in, and how large are the datasets?
- What programs are available for processing LiDAR data?
- What are the objectives and importance of geometric correction?
- What is geometric correction in remote sensing?
- Why is geometric correction important?
- What is the filtering process of LIDAR data?
- What are LIDAR data mapping and its application?
- In any form of terrain coverage, what are the proper primitives for identifying conjugate surface elements in overlapping strips made up of irregular groups of non-conjugate points? What is the technique for deriving the involved features/primitives in an automated/semi-automated manner?
- What is the process for identifying conjugate primitives in overlapping strips automatically?

BLOCK 2 : WEB GIS

UNIT 3 : WEB GIS TECHNOLOGY

3.1 OBJECTIVES
3.2 INTRODUCTION
3.3 WEB GIS TECHNOLOGY
3.4 SUMMARY
3.5 GLOSSARY
3.6 ANSWER TO CHECK YOUR PROGRESS
3.7 REFERENCES
3.8 TERMINAL QUESTIONS

3.1 OBJECTIVES

After reading this unit the student will understand:

- Basics of Web and Internet technology;
- Difference between desktop and web GIS;
- Definition and components of Web GIS;
- Different types of Web GIS applications and services; and
- Web 2.0 and GIS

3.2 INTRODUCTION

With the recent advances in broadband and wireless communication technologies as well as the dramatic increase in internet technology it is promising to extend further the reach and range of GIS user working in offices and laboratories in the field or at home would lead to the development of internet GIS or web enabled GIS. The Internet technology as a digital communication medium enhances the capability of GIS data and software application by making them more accessible and reachable to wider range of users, planners and decision makers. The web enabled GIS would facilitate decision making at the strategic, tactical, and operational levels; support for performance of administrative operations; and serve as a gateway for decision-makers and general users to access the system conveniently and effectively. Internet is emerging as a perfect means of GIS data accessing, analyzing and transmission. The World Wide Web, FTP (file transfer protocol) and HTTP programs make it convenient to access and transfer data files across the Internet.

Advanced GIS work is influenced significantly by the high computing capabilities and advanced visualization system using latest state-of-art of information and communication technologies. The Integration of GIS and Internet technology has revolutionize the use and range of geo-spatial data and its applications in planning and implementation of strategies for a wide range of activities including disaster management. The combination of GIS and Internet offers great possibilities, such as the interactive access to geospatial data, real time data integrations and transmission, enhancement of the functions of geographic information, and the access to platform independent GIS analysis tools.

Internet has been changing the ways of data access, sharing and dissemination. It is further changing the means of analysis and visualization of Geographic Information Systems (GIS). Internet GIS is rapidly evolving as Internet and Web technologies change. Internet GIS is a network-centric GIS tool that uses Internet as a major means to access and transmit distributed data and analysis tool modules, and to conduct analysis and visualization.

Internet GIS is the framework of network-GIS that utilizes the Internet to access remote geographic information and geoprocessing tools. Instead of Internet GIS, other similar terms often used include: on-line GIS, Distributed Geographic Information (DGI), and Web GIS. However, some minor differences exist with the use of these terms. Internet GIS and on-line GIS both refer to the use of the Internet to access remote geographic data and geoprocessing tools; Web GIS specifically refers the use of World Wide Web (WWW), one of the many applications on the Internet; Distributed GIS or DGI has the broadest meaning which includes both wired Internet GIS and wireless mobile GIS.

3.3 WEB GIS TECHNOLOGY

BRIEF HISTORY OF INTERNET AND WEB GIS

The short 'history' of Internet GIS is inseparable from the development of Internet itself. In 1993 Internet growth exploded with the appearance of Mosaic, the first World Wide Web browser supporting multimedia, which was regarded as a significant milestone for the Information Revolution. In the same year, Steve Putz of the Xerox PARC center published the first prototype web-GIS. As an early experiment in interactive web services, the PARC Map Viewer (http://mapweb.parc.Xerox.com/map, no longer maintained) served more than 150 million maps. Users without any specialized software or data could view digital maps through a standard web browser with Internet connection. As the pioneer of web-GIS, PARC Map Viewer prompted many GIS users to take advantage of the web and inspired major software vendors to develop their Internet GIS software products. By 1996 major vendors, including Autodesk, ESRI, Intergraph and MapInfo, all introduced their first Internet GIS products.

The Internet is shaping the ways of traditional Geographic Information Systems (GIS) function. It is changing the access and transmission of GIS data, applications and visualizations. Internet has facilitated three major changes in GIS:

a. Access to data;
b. Transmission of data; and
c. Access to GIS analysis functionalities.

Internet provides a means for data access. It greatly improves the accessibility of data in general, and GIS data in particular. GIS data providers, including government agencies and private organizations, are discovering the convenience of publishing GIS data on the World Wide Web. GIS data warehouse and Clearinghouse are two common forms of Internet data access systems. Internet is also a perfect means of GIS data transmission. The World Wide Web and FTP (file transfer protocol) programs make it convenient to transfer data files across the Internet. They provide GIS users easy access to acquire GIS data from different sources of data providers. GIS users can download data almost instantly from the Web over the Internet without the hassles of calling or writing around to find who has what data, making a formal request, and then waiting for days or weeks for the data disks to be arrived. Web browser is providing Internet users access to GIS applications software residing in Web servers. End users could work with GIS data interactively on the Web browser without owning GIS software on his/her local machine.

We will be using Internet GIS and Web GIS terminology synonymously in this chapter for better understanding of the readers.

WHAT IS INTERNET/WEB GIS?

Internet is a global network of computers connected through communication devices to one another. It is a means for GIS users to exchange GIS data, conduct GIS analysis and present GIS output (Karnatak et al, 2007). Therefore, Internet GIS is a special GIS tool that uses Internet as a means to access and transmit remote data, conduct analysis and make GIS more accessible and usable for its user (s).

Internet or Web GIS should have all or almost all functionalities the traditional GIS software have. In addition, Internet/ Web GIS should have additional functions that take advantage of the Internet and its associated protocols such as World Wide Web (WWW) and File Transfer Protocol (FTP). These additional functions include exchanging remote data and application programs, conduct GIS analysis on the Internet without owning GIS software on the local machine, and presenting interactive maps and data on the Internet. The key features of the Internet GIS are object-oriented, distributed and interoperable application. Each GIS data and functionality is an object. These objects are resided on different servers on the Internet and are assembled and integrated when needed. These data and analysis functionalities can be exchanged or interoperable by any other systems on the Internet.

- **Web GIS is an integrated client/server network system**

The concept of client/server involves splitting an application into tasks between the server and client. A client/server application has three components: a client, a server, and a network each of them is supported by specific software and hardware. The client sends a request to the server, which processes the request and returns the result to the client, the client then manipulates the data and/or results and presents to the user. Internet/Web GIS applies the client/server concept in performing GIS analysis tasks. It breaks down the task into server side and client side. The client can request for data, analysis tools or modules from the server. The server either performs the job itself and sends the results back to the client through the network, or sends the data and analysis tools to the client for use on the client side.

- **Web GIS is an Interactive System**

World Wide Web provides a natural interactivity to the Internet through its hypertext linkage. The users can browse the web page by page through the hypertext linkage. However, each web page is a static image that published in a web server. Similarly, many map pages those were published on the web are also static map images. Typically, static maps are portrayed at many sites to provide some geographic context to a site's information such as where the location of a citizen service such as shopping centers, medical services, public transport etc. Most of them are static image file like JPEG or gif (Graphics Interchange Format) file. The user can perform basic queries, zoom to the location and pan the map for locating the geographical object (s).

- **Web GIS is a Distributed System**

One advantage of the Internet is that it can access distributed database and perform distributed processing. Information and applications can be resided on different computers across the network. Internet/Web GIS takes advantage of this distributed system so that the GIS data and analysis tools can be resided in different computers on the network. GIS data and analysis tools are individual components or modules. Users can access those data and application programs on-demand from anywhere on the network. The user does not have to install the data and application programs in his/her local computer. Whenever it sends requests to the server, the server would deliver the data and analysis tool modules for just-in-time performance.

- **Web GIS is a Dynamic System**

Because Internet GIS is a distributed system, database and application programs reside on computers that publish them. Those data and application programs are updated by those who manage them. Once the data and application programs are updated, they are available to every user on the Internet. In other words, Internet/Web GIS is dynamically linked with the sources. The GIS is updated as soon as the sources are updated. This dynamic linkage with the sources always keeps the data and software current. The Internet/Web GIS can also link with real time information, such as satellite images, traffic movements and accident information by real time connection with the relevant information sources.

- **Web GIS is Cross-Platform**

Web GIS can be accessible cross platforms. It doesn't matter what operating system the user is running. The Web GIS is not limited to any one kind of machine or operate system. As long as one has access to the Internet, he/she can access and use Web GIS, giving that the Web GIS providers provide platform neutral or cross-platform GIS tools.

- **Web GIS is a Graphical Hypertext Information System**

With the help of the hypermedia system in the web, Internet GIS can link different map pages through a hypermedia hot link. For example, one can browser a national map to the state map, to the city map by clicking on the hot link that links the national map to the state maps and further to the city maps. Web/Internet GIS, on the other hand, enables users to manipulate GIS data and maps on the Internet interactively. Users can perform basic GIS functions such as zoom, pan, query, and labeling using Web browser. It can perform spatial queries like where is the closest restaurant and hotel is. It should be able to perform more advanced analyses such as spatial analysis and network analysis.

TRADITIONAL GIS VS. WEB GIS

Traditional GIS system consists of a single software package, plus data on a single machine. This stand-alone model, which lacks software compatibility and networking capability, can no longer meet the requirements of complicated situations, which are often multi-disciplinary, multi-platform, multi-software and multi-user. Internet GIS separates the user interface, data storage and processing, which are normally on a single machine in traditional GIS; thus it is able to provide interoperability, reusability, and flexibility, which are impossible to realize in a stand-alone GIS. Some advantages of Internet over traditional mainframe and desktop GIS include:

- World-wide access of GIS data and information;
- Much more accessibility-easier and faster access expands potential GIS users;
- Data, information or service distribution can be anywhere, anytime to desktop computers, laptops PDAs or mobile phones;
- No proprietary GIS products are necessary; people who couldn't afford expensive GIS software can access free data and services or pay a minimum cost;
- Reduce duplication of data collection and simplify the updating of data and information;
- Able to combine data and information from many different sources including local data;
- Improve the efficiency of GIS database management and reduce the cost of GIS database maintenance.

Some major differences between traditional and distributed GIS are listed in Table below. Although Internet GIS has many advantages over the traditional model, it is still at an early stage of development. At present it is not possible to provide the same functionality as traditional GIS since analysis functions are still very limited compared with desktop GIS.

Table 3.1. Traditional vs. Web GIS

Traditional GIS	Web GIS
Data needs to be gathered separately in large volume and at high cost.	Data can be published and shared; the overall cost of data collection and maintenance can be largely reduced.
Duplication of effort due to lack of communication and sharing system.	Avoids duplication because of a sharing system.
Proprietary or product specific GIS software are required for all users.	Proprietary software is not necessary for all users
GIS operations can only be available locally.	GIS services can be leased, subscribed or purchased through Internet.
Maintaining and updating the information separately costs money and effort.	Increase the potential volume of available information, distribute the workload, and simplify the updating of information.
The cost and availability of GIS data and software limit the potential pool of GIS users.	GIS users are greatly expanded due to the flexibility and lower cost.

WEB GIS TYPES AND SERVICES

Web GIS offers different kind of services to its users. These services are interoperable and accessible through simple web browser or any other web client software application. Some of the most popular type services available through web GIS are:

- **GIS Data Download**: Web GIS applications provide facility to browse the GIS data in web browser environment with a physical data download facility. The data download is made available through HTTP or FTP protocols. The GIS data is also made available as data service in web GIS application. The data services are allowing to access GIS data in a GIS software using standard web URL. Such data services are permitting specific GIS analysis to its users without physical download of the data.

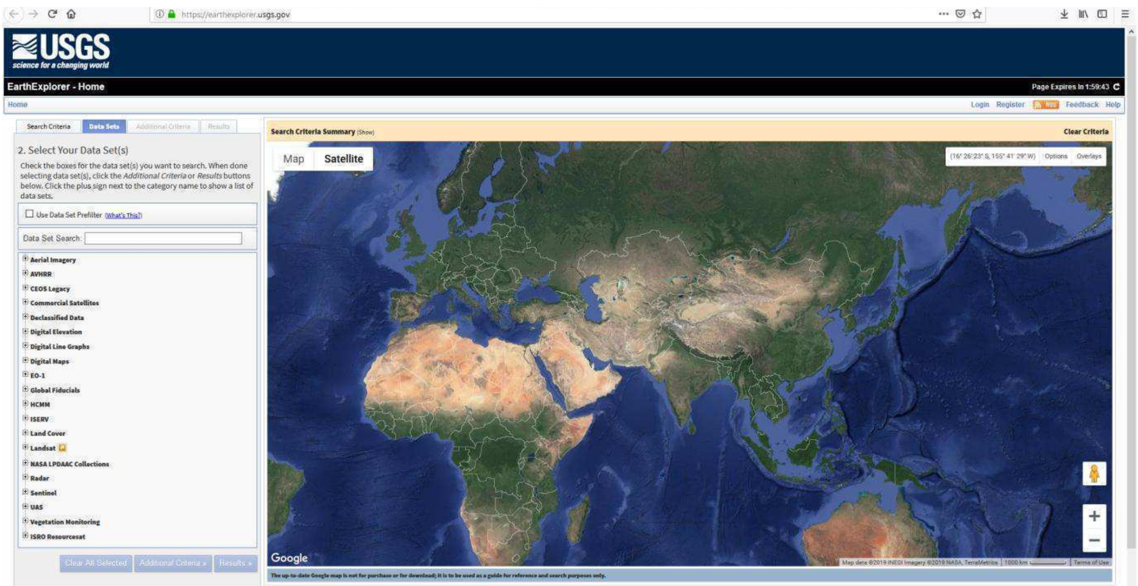


Figure 3.1- GIS Data download in Web GIS application (https://earthexplorer.usgs.gov/)

- **Static Map Display**: In a web GIS application, the GIS data is presented to its users as a static image such as JPEG, GIF, and PNG etc. Generally, such static Images are generated on-the-fly based on user's request. This capability of Web GIS makes it more accessible and user-friendly non-GIS professionals. The typical client-server communication for static map publishing in web GIS application shown in Figure 3. 3. The request is sent to the web server (also known as HTTP server) for the map. The request is processed GIS server and appropriate static map is generated for visualization and further responded back to the user as HTTP response. The role of GIS server is to create web compatible map using GIS data and applying cartography on it. Uses simple Web publishing technology, a two-tier client/server model. Maps are graphic images embedded inside the HTML page. The server can only provide ready-made files. Users have no interaction with the map except clicking.

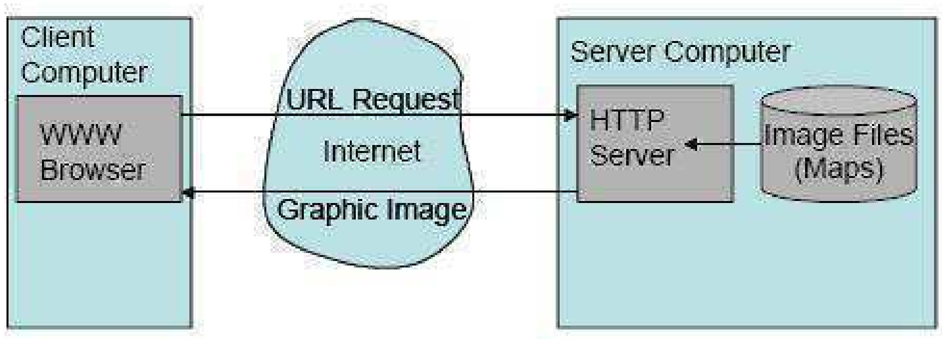


Figure 3.2- Static Map Publishing

One of the example from a web GIS application for Forest Fire Risk and advisory system is shown in Figure 3.3 for Uttarakhand state.

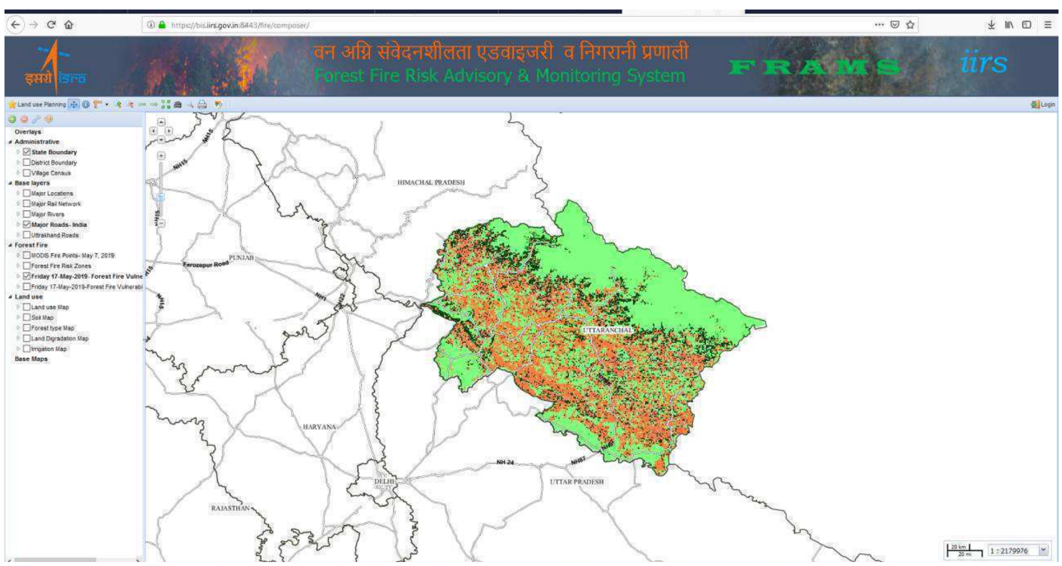


Figure 3. 3- Static Map Visualization in Web GIS application (https://bis.iirs.gov.in)

- **Metadata Search using Catalogue**: Metadata is referred as "data about data". It means it gives characteristics and definition of GIS data. Through metadata the users can get an appropriate GIS data for their project or study. The web GIS platform provides a clearing house or gateway to its users. This metadata gateway provides various modes of GIS data searching in the online data repositories.

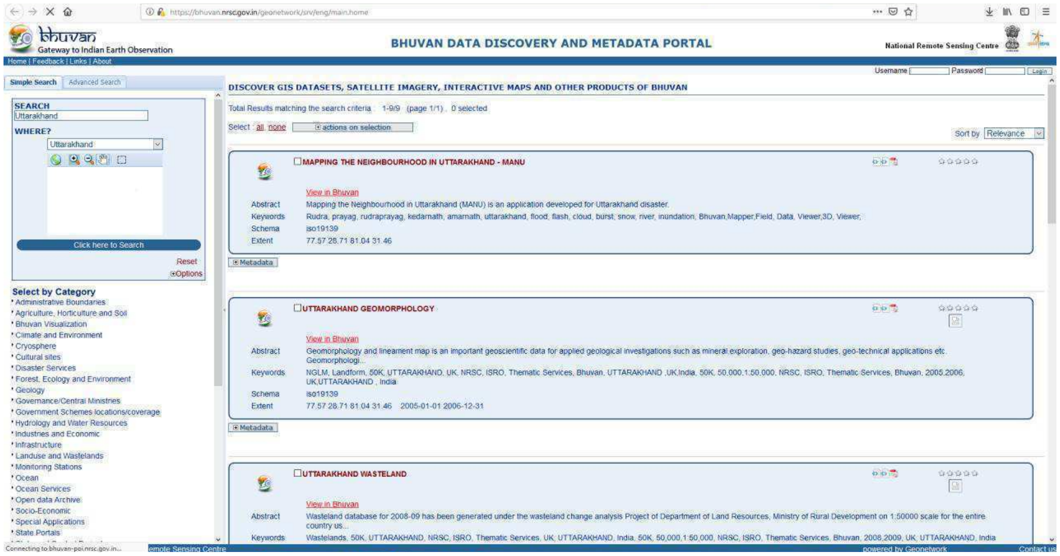


Figure 3.4- Metadata Search in Bhuvan Geoportal

- **Dynamic Geo-visualization**: In Web GIS applications, the maps are created 'on the fly' according to the scale, location or themes specified by the audience. The dynamic map browsing or visualization is one of most attractive component of Web or Internet based GIS. The users are able to access the original GIS data interactively which is very help to make effective use of GIS data in planning and decision making. The recent advancement in web based technologies are permitting dynamic and interactive geo-visualization in 2D and 3D environment. The 3D visualization of GIS data provides the terrain information which is capable of presenting virtual reality of the geography.

Figure 3.5- 3D Geo-visualization in Internet GIS (CartoSAT Data)

- **GIS Query and analysis:** The web GIS applications are allowing to perform basic and advanced location-based queries. These applications also provide very interactive platform to perform various spatial analysis such and vector and raster operations.

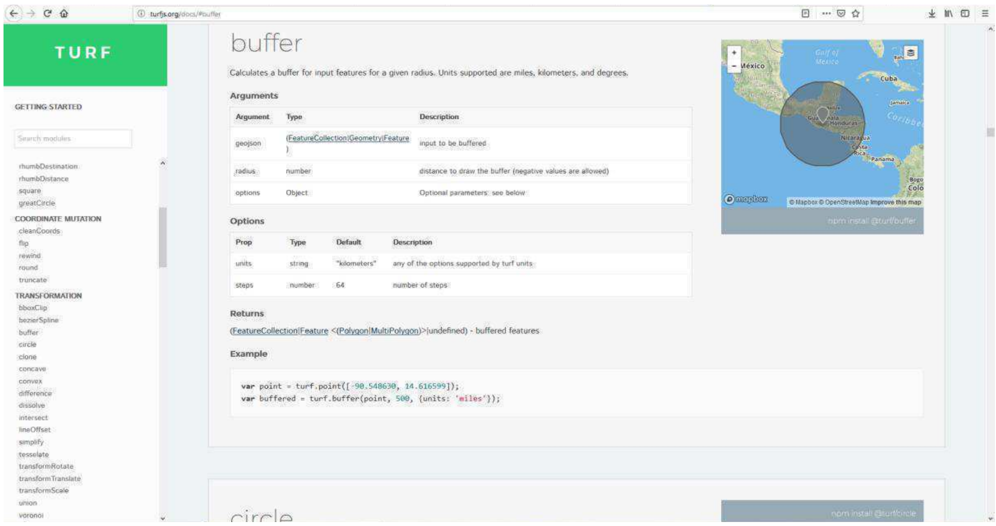


Figure 3.6- Various GIS query and analysis in Web GIS application (Ex. http://turfjs.org)

- **GIS Data Processing:** Online GIS data processing is emerging as one of the exciting application of Web GIS technology. The advancements in computation and internet bandwidth is allowing GIS user (s) to perform complex scientific geo-data analysis on-the-fly. The user (s) are able to perform GIS analysis without having any GIS data or software installed in their local computer. In fact in some cases now such data analysis are possible through mobile platforms.

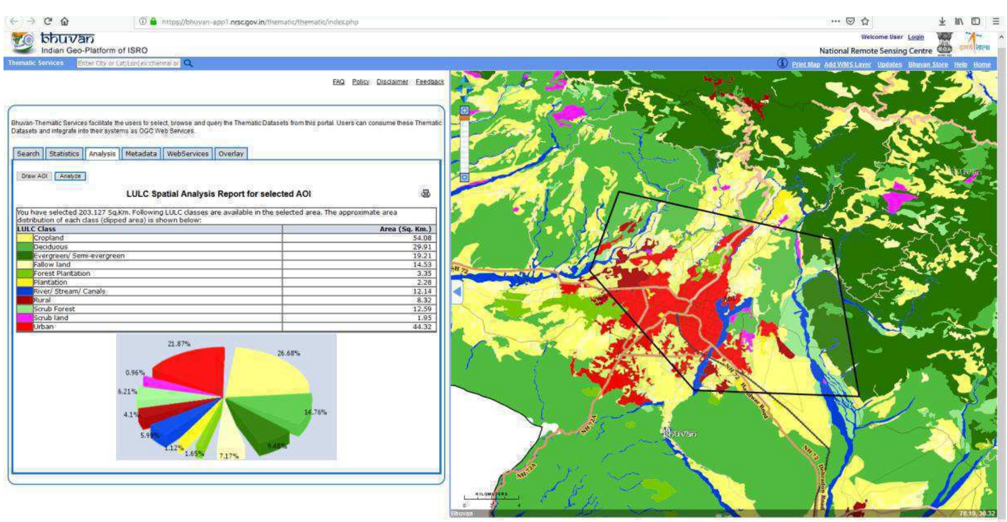


Figure 3.7- On-the-fly analysis of Land use map of Uttarakhand in Bhuvan Geo-portal

Growth and evolution of Web GIS technology is shown in Figure 8. The advancements in Information and Communication Technologies (ICT) has enabled many innovative tools and functionalities in Web GIS applications. The grown in this technology during last decade can be considered as pioneering development which is leading to develop an intelligent GIS applications. Such applications are also known geo-web applications which can solve many real world location specific problems.

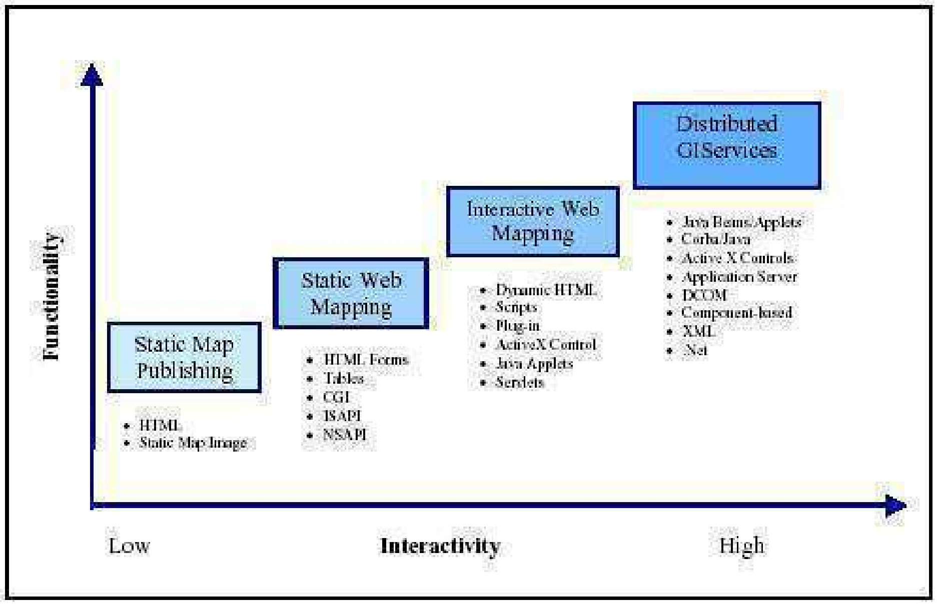


Figure 3.8. Evolution of Distributed GIS

WEB GIS TECHNOLOGY: AN OVERVIEW

Internet, a client/server system is a perfect means of GIS data accessing, analyzing and transmission. The World Wide Web, FTP (file transfer protocol) and HTTP programs make it convenient to access and transfer data files across the Internet. The Internet provides GIS users easy access to acquire GIS data from central server system to diverse data source in distributed environment. The World Wide Web is a fast becoming standard platform for Geographic Information System (GIS) and related technologies. It is a means for GIS users to exchange GIS data, conduct GIS analysis and present GIS output in the form of maps, report and web services.

Internet has facilitated five major changes in GIS: (1) Online access the data for query and visualization; (2) Sharing and distribution of data; (3) GIS Data analysis and processing; (4) Online transaction processing of data; and (5) linking of real/near-real time data. The Internet GIS applications provide all or almost all functionalities of traditional GIS software in addition to power of internet and related technologies. The users of Internet GIS application can use advanced GIS tools for analyzing their data without having any specific data or GIS software installed in their machine. The advent of web 2.0 in internet technology has opened a new dimension for geo-spatial data applications by introducing Geoweb 2.0 for GIS data and applications. The Geoweb 2.0 allows more interactive and latest GI system as a mashup architecture which is very effective for development of decision support tools for any decision problem where integration of latest information, public participation is important.

The most important advantage of this technology is "dynamic" nature of GIS. For example, once any client (s) or database administrator updates the data or information at server end, it will available for all the clients on web at the same time. The Internet GIS can also link with real time information, such as satellite images, traffic movements and accident information etc. by real time connection with the relevant information sources like sensor web. And also these GIS applications are cross-platform which means it can be accessed using any operating system or platform. The Internet GIS applications can categorize into three major categories i.e.

- Server-side applications;
- Client-side applications and
- Mix of server and client side application

Server-side applications completely rely on GIS server (usually reside on a remote server) to perform all GIS activities including data analysis and processing, client-side applications perform GIS activities at client (user) end by using local data and application and the mix of server and client side application allows to use local and remote data application together for producing GIS products.

WEB GIS SYSTEM WORKFLOW:

A typical GIS application usually includes three essential elements i.e. presentation (refers to user interface), logic (refers to processing) and data (refers to database or database management system). The relationship between these three elements is that one element sends the request to other element and the other elements responses the request (Karnatak, et al, 2008). This making and fulfilling of request is called client/server-computing model (Shan and Earle, 1998). The element that makes request is called *client* and the element that fulfills the request is called *server*. In the typical GIS environment the user sends the request through GUI of GIS program to have a query operation; this request will be sent to the logic element, which will make further request to the data element and search for significant data and will send back the result to the user via GUI.

The client/server model is based on distributed computing where one client program, installed in client's machine, communicates with another program installed in server machine for the purpose of exchanging information. In the client/server environment, once the client and server software processes recognize a common set of protocols, they need not depend on specific hardware or operating system requirements. The client and the server can exist in the same host computer, or be distributed across a wide variety of computers at widely distant locations. The client/server computing is developed from a distributed database model where everything except data management is handled in the client computer. In this configuration the client needs more computing resources, both hardware and software. In the technological development most of the application processes shifted towards server end, and distributed system. A distributed system is a 'thin client' architecture where interface management is distributed between client and server and application logic and data management part is entirely handled by powerful server (s). The term 'thin client' refers to the least amount of processes that take place at the client end in comparison to the huge amount of processes that take place in the server end.

The client/server is a software defined model of computing, not a hardware defined. The client/server environment is generally heterogeneous in nature where the hardware platform and operating system of the client and server are not usually the same. In such cases, the communications mechanism may be further extended through a well-defined set of standard i.e. Application Program Interfaces (APIs) and Remote Procedure Calls (RPC). The client/server computing provides the capability to use the most cost-effective user interface, data storage, connectivity, and application services.

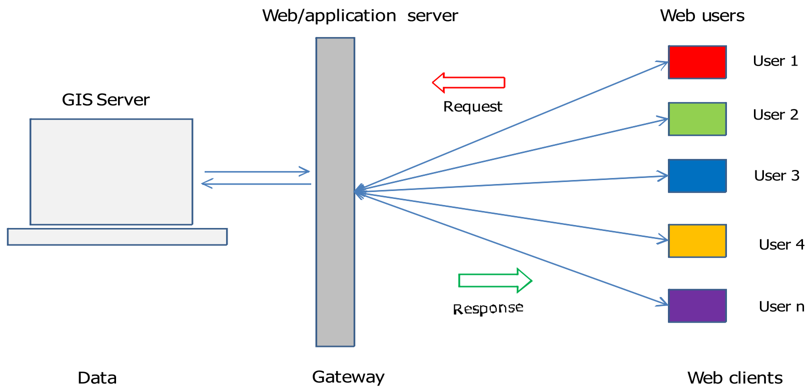


Figure 3.9- A typical client/server communication

The client/server system can be further implemented in a tier system. In the two-tier structure the presentation component would be located at the client machine and the logic component and data component could be located at the server machine. The two-tier architecture is good for small applications and gives good performance when the number of users connected with server is less. In the three-tier system, the presentation element is located at the client machine and logic element at the server machine and data element is stored in the database server. An n-tier system expands the logic or data element into multiple computers or multiple components. In an n-tier system each element of client/server architecture can be further extended to different machine as per requirement. For example the data server can be further extended into spatial database server to non-spatial database server and logic element into application server to map server in typical distributed GIS environment. The multiple data server can be linked together by data category server or data access program like ODBC, JDBC etc. The selection of suitable system configuration in client/server architecture depends upon the size of application, data and number of anticipated users connected with server at a time. The tier architecture can also configured logically or physically, where each component i.e. presentation, logic and data can reside at same machine but in different partition or at different machine.

From a hardware platform perspective, three-tier hardware architecture involves three classes of computers: the client (usually a PC or PDA); the middle tier (usually a workstation server); and the back-end server (usually a mainframe computer). Selection of suitable hardware is also dependent upon the size of application, data and number of anticipated users connected with server at a time.

WEB GIS SERVER COMPONENTS:

In a typical development environment of web based GIS application following server components are essential:

DATABASE SERVER

The database server may have a file based system or Relational Database Management System (RDBMS) based or a combination of files and RDBMS. In a typical web GIS application the spatial data is organized in RDBMS environment which allows better performance, data security, data consistency and many more advantages of RDBMS for GIS data sets. Following are well known and famous RDBMS based database server software solutions available for geo-spatial data sets:

Table 3.2- Important GIS database servers with RDBMS support

S. No.	Database server software	Strength	Official Web address	Status
1	PostgreSQL + POSTGIS	Performance and Advanced analysis.	http://postgis.refractions.net http://www.postgresql.org	Open source (Freeware)
2	ArcSDE+ Selected RDBMS	Technical support.	http://www.esri.com	Commercial
3	Oracle Spatial	Support for JAVA and store common spatial data types in a native Oracle environment.	http://www.oracle.com/technology/products/spatial/index.html	Commercial
4	My SQL	Compatibility with PHP and other open source s/w.	http://www.mysql.com	Open source (Freeware)
5	TerraLib	Time series analysis and supported by many RDBMS.	http://www.terralib.org/	Open source (Freeware)
6	SpatiaLite	Spatial extensions for the open source SQLite database.	http://www.gaia-gis.it/spatialite	Open source (Freeware)
7	IBM DB2	Strong error handling.	http://www-01.ibm.com/software/data/db2	Commercial

GIS OR MAP SERVER

Map server or GIS server is a software package or program, which is responsible for rendering the GIS data into web browser. Since the standard web servers and browsers supports only standard image and data formats like .jpeg, gif .txt, .html, .xml etc. To represent or publish geo-spatial data in web compatible format there is a need of intermediate software components called as a map server or GIS server. Today many of the Map server products available either as commercial product or as an open source/freeware products for map publishing in Internet environment. The available Map servers' products are either based on CGI (Common gateway Interface) or Servlet based connectors. The Important GIS/map server products available at present are shown in table 3.3.

Table 3. 3- Important GIS/Map server products and their strengths

S. No.	Map/GIS server software	Strength	Official Web address	Status
1	Geoserver	Performance, security, vector support and OGC web services.	http://www.geoserver.org	Open source (Freeware)
2	UMN Mapserver	Performance, raster and vector support.	http://www.osgeo.org	Open source (Freeware)
3	Arc GIS Server	Technical support and GIS processing.	http://www.esri.com	Commercial
4	SkylineGlobe	3D visualization.	http://www.skylineglobe.com	Commercial
5	Mapguide	Support and quick customization.	http://mapguide.osgeo.org	Open source (Freeware) and Commercial
6	Degree	OGC web services	http://www.deegree.org	Open source (Freeware)
7	ERDAS APOLLO Server	Raster support.	http://www.erdas.com	Commercial

8	Intergraph Geo web server	Engineering based application.	http://www.intergraph.com	Commercial

APPLICATION SERVER

An application server in a GIS environment especially in web GIS applications is software which provides customized software applications with services like query system, GIS analysis and processing, report generation, data security and authorizations etc. Many times the application servers are a part of GIS/map server. In general the web GIS application servers are customized by using API (Application programming Interface like Open Layer), GIS Objects, and spatial libraries (e.g. GDAL, OGR, Geotooletc). Many application servers like JBOSS (java application server) are bundled with GIS servers like ERDAS Appolo server and available as a single package. The development of application server using open API like Open Layer API, Google API, Yahoo API are becoming very popular in user community due to its easy development and interactive support from user community. The development of application server can be done using Java SDK, .net framework, PHP, Javascript etc.

Table 3.4- Important application development environment for web GIS applications

S. No.	**Database server software**	**Strength**	**Official Web address**	**Status**
1	Open Layers	AJAX library for accessing geographic data layers of all kinds	http://openlayers.org	Open source (Freeware)
2	GeoBase	Geocoding, navigation and route optimization.	http://www.geobase.org	Open source (Freeware)
3	Geomajas	Aggregation and transformation of GIS data sources	http://www.geomajas.org	Open source (Freeware)
4	GeoTools	GIS data creation, editing and processing using JAVA framework	http://www.geotools.org	Open source (Freeware)
5	GDAL/OGR	Compatibility with any development environment.	http://www.gdal.org	Open source (Freeware)
6	GEOEXT	Rich GUI using open layer API.	http://www.geoext.org	Open source (Freeware)

WEB SERVER

A web server is a computer program which uses the client/server model and the World Wide Web's Hypertext Transfer Protocol (HTTP), serves the files that form web pages to web users (whose computers contain HTTP clients that forward their requests). The primary function of a web server is to deliver web pages on the request to clients. This means delivery of HTML documents and any additional content that may be included by a document, such as images, style sheets and scripts. A user agent, commonly a web browser or web crawler, initiates communication by making a request for a specific resource using HTTP and the server responds with the content of that resource or an error message if unable to do so.

Table 3.5- Popular web server products

Product	Vendor	Web Sites Hosted	Percent
Apache	Apache	397,867,089	64.91%
IIS	Microsoft	88,210,995	14.39%
Nginx	Igor Sysoev	60,627,200	9.89%
GWS	Google	19,394,196	3.16%
Resin	Caucho Technology	4,700,000	0.77%
lighttpd	lighttpd	N/A	N/A
Sun Java System Web Server	Oracle	N/A	N/A

*<u>**Source**</u>: Wikipedia- https://en.wikipedia.org

WEB 2.0 AND GIS

Web 2.0 is referred as an interactive web application with public participation which allows participatory information sharing, interoperability, user-centered design, and collaboration on the World Wide Web. The integration of web 2.0 with GIS data application is known as GeoWeb 2.0. The Web 2.0 technologies and standards enable web as a platform by allowing user participations in web application. In the realization of Web 2.0 new knowledge and services are created by combining information and services from different sources which is known as ‘mashups’. In this environment users are not seen simply as customers or readers but as contributors and co-developers or sometime owner of the web pages. Web 2.0 can be realized by combining several web computing technologies such as AJAX, Open API, REST, XML, XHTML/CSS, RSS/GeoRSS and other related technologies. In the realization of Web 2.0 new knowledge and services are created by combining information and services from different sources available on web. This requires standardized protocols, published interfaces and adequate business models. This trend and advancement of web technology leads toward realization of Spatial Mashup for GIS data and services. The spatial mashup can be build using open map APIs like Open Layer or Google APIs where AJAX based architecture is adopted for interactive and fast accessing of geo-spatial data and services. Many AJAX based applications are developed in Internet domain by using these open APIs for interactive web mapping services. Geo-web portals can be published by using spatial mashup technology for geo-visualization, online mapping and editing, query and analysis, web processing and other related operations for GIS. Some of the case studies are presented with their unique functions: map viewer, channel editor, harvest tool, and many kinds of extensions for OGC WMS and catalogue services. The hybrid mashup is towards more intelligent web services handling geo-based resources. In this architecture content provider provides data, API provider provides API to access and modify the data and client web browser displays mashup pages to the user, and may perform mashup logic, being implemented using AJAX.

3.4 SUMMARY

Web and internet GIS are emerging as one of the important source of geospatial data and information. The advancement in internet and related technologies are directly influencing the web GIS and related areas. The online data repositories and geo-portals are getting popular among user communities. The various components of web GIS technologies are available in free and open source environment which gives it great opportunity to develop innovating applications. The integration of Web 2.0 and GIS is another interesting development in this domain where interactive GIS applications can be developed for user defined thematic applications.

3.5 GLOSSARY

1. Traditional GIS- System consists of a single software package, plus data on single machine.
2. Server- A computer that stores information that a number of computers can share.
3. Database- A large amount of data that is stored in a computer and can easily be used.
4. WWW- World Wide Web

3.6 ANSWER TO CHECK YOUR PROGRESS

1. What is Internet and Web GIS?
2. What are the different components of Web GIS?
3. What is GIS/Map Server?
4. What is Web 2.0?

3.7 REFERENCES

- Armstrong, M. P., Densham, P. J. (1990). Database Organization Strategies for Spatial Decision Support Systems. International Journal of Geographical Information Systems, Vol. 4(1), pp. 3-20.
- Aydin, G., 2007a. Service Oriented Architecture for Geographic Information Systems Supporting Real Time Data Grids, Ph.D. Thesis, Indiana University, IN, USA, pp. 12-47 Josuttis, 2007.
- Bonczek, R., C. Holsapple, and A. Whinston, (1978) Mathematical Programming within the Context of a Generalized Data Base Management System, R.A.I.R.O. RechercheOperationalle/Operations Research, Vol. 12, No. 2, pp. 117-139.
- Brian E. Mennecke (1995) Understanding the Role of Geographic Information technologies in Business: Applications and Research Directions, Journal of Geographic Information and Decision Analysis, vol.1, no.1, pp. 44-68.
- Densham, P. J., Goodchild, M. F. (1989), Spatial Decision Support System: A Research Agenda, Proceedings of GIS/LIS '89, Florida, pp. 707-716.
- Fedra, K. (1997). Spatial Decision Support in Resources Management and Environmental Applications, Proceedings of the Eleventh Annual Symposium on Geographic Information Systems held in Vancouver, British Columbia, Canada, 302-305.
- Fürst, J., Girstmair, G., and Nachtnebel, H.P. (1993). Application of GIS in Decision Support Systems for Groundwater Management. In HydroGIS93: Application of Geographic Information Systems in Hydrology and Water Resources, edited by K. Kovar and H.P. Nachtnebel, IAHS Publ. no. 211: 13-21.
- GangothriRajaram, Harish Chandra Karnatak, SwaminathanVenkatraman, Manjula KR and Kannan Krithivasa, 2018 "A novel Computational Knowledge-Base Framework for visualization and quantification of Geospatial Metadata in Spatial Data Infrastructures", Geoinformtica, Springer Science Publication, DOI: 10.1007/s10707-018-0317-6.
- Harish Chandra Karnatak, R Shukla, VK Sharma, YVS Murthy, V Bhanumurthy, 2012 "Spatial mashup technology and real time data integration in geo-web

application using open source GIS–a case study for disaster management", Geocarto International 27 (6), 499-514.

- Harish Chandra Karnatak, S Saran, K Bhatia, PS Roy, 2007 "Multicriteria spatial decision analysis in web GIS environment", Geoinformatica 11 (4), 407-429.
- Harish Chandra Karnatak, Sameer Saran and P. S. Roy, 2005 "Spatial services, a click away" article in Geospatial Today Volume 3 Issue 5 pp-42-46 January-February 2005.
- Harish Karnatak, PLN Raju, YVN Krishnamurthy and A. Senthil Kumar, (2015), "E-learning Based Capacity Building in Geoinformatics", ISG Volume 21, Issue 1-2, pp 4-13, ISSN: 0972-642X.
- Hendler, J. (1999). Is There an Intelligent Agent in Your Future? Nature, Vol. 398, No. 6723, March 11, 1999.
- Hopkins, L., (1984). Evaluation of methods for exploring ill- defined problems, Environment and Planning B 11:339-48.
- K. Sahina and M. U. Gumusay, 2008, "Service Oriented Architecture (SOA) based web services for Geographic Information Systems", International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B2. Beijing 2008.
- Konsynski, B. R., (1980). On the Structure of a Generalized Model Management System,'' Proceedings of the Fourteenth Hawaii International Conference on the System Sciences,Vol. 1, pp. 630-638.
- Kotzinos, D. and Chrysoulakis, N., 2003, Design of GIS Web Services for Environmental Monitoring: Using Satellite Imaging to Calculate Vegetation Indices, Geographical Information Systems and Remote Sensing: Environmental Applications, Volos, Greece.
- Lee K. 2010. Intelligent Geo-Web Services Based On Hybrid-Mashup Using Open Source Geo-Spatial Software, WebMGS 2010 1st International Workshop on Pervasive Web Mapping, Geoprocessing and Services XXXVIII-4/W13, ISSN Number: 1682-1777.
- Nogueras-Iso, J., F. J. Zarazaga-Soria, R. Bejar, P. J. Alvarez and P. R. Muro-Medrano, 2005. OGC Catalog Service: a Key element for the development of Spatial Data Infrastructure, Computers & Geosciences, 31, pp. 199-209.

- Open Geospatial Consortium (OGC) official website- http://www.opengeospatial.org (last accessed 2nd April, 2019).
- PeasrsonInformit official website- http://www.informit.com/articles/printerfriendly/336265 (last accessed 2nd April 2014).
- Sayar, A., 2008, GIS Service Oriented Architecture, Community Grids Laboratory, Indiana, USA.
- Sayar, A., Aktas, M., Aydin, G., Pierce, M. and Fox G., 2005, Developing a Web Service-Compatible Map Server for Geophysical Applications Technical report, Indiana University, Indiana, USA.
- Shan Yen-Ping and Earle Ralph H. (1998). Enterprise computing with objects: from client/server environments to the Internet, publisher-Wesley Longman Publishing Co., Inc. Boston, MA, USA, ISBN: 0-201-32566-7.
- Singh, H., Garg, R.D., Karnatak, H.C. , 2019 "Online image classification and analysis using OGC web processing service", Earth Science Informatics, Springer Berlin Heidelberg, https://doi.org/10.1007/s12145-019-00378-z Print ISSN 1865-0473 Online ISSN 1865-0481, pp-1-11
- Sprague Jr. R. H. and H. J. Watson (1996). Decision support for management, Upper Saddle River, N.J.: Prentice Hall publication.
- Turban, E., (1995). Decision Support and Expert Systems, 4th Edition. Englewood Cliffs, NJ: Prentice Hall publication.
- Walsh, M.R. (1993). Toward Spatial Decision Support Systems in Water Resources, ASCE Journal of Water Resources Planning and Management, 119, 158-169.
- Watkins, D.W., McKinney, D.C., Maidment, D.R. and Lin, M.D. (1996). Use of Geographic Information Systems in Ground-Water Flow Modeling, ASCE Journal of Water Resources Planning and Management, 122, 88-96.
- World Wide Web Consortium (W3C) official website- http://www.w3.org/ (Last accessed on 2nd April 2019).
- Zhong-Ren Peng, Ming-Hsiang Tsou, (2003). Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks, ISBN: 0-471-35923-8m March 2003.

Acknowledgements

This chapter is a compilation specifically intended for enhancing knowledge on Web GIS technology and its various components. The data content is compiled from various web sources and all the resources and their authors are gratefully acknowledged. The contents presented in this chapter are pure for academic purpose.

3.8 TERMINAL QUESTIONS

1. What are the different types of services available in Web GIS platform?
2. What is the difference between Map server and application server?

UNIT 4 : WEB GIS ARCHITECTURES

4.1 OBJECTIVES

After reading chapter student will understand:

- The System architecture of client-server computation;
- Computation architecture of Web GIS;
- Data and information flow in Web GIS;
- Concept of Service Oriented Architecture (SOA);
- SOA and Web GIS;
- OGC Web service specifications for geospatial data and services; and
- Concept of Mobile GIS and its architecture.

4.2 INTRODUCTION

A typical GIS application usually includes three essential components i.e. presentation (refers to user interface), logic (refers to processing) and data (refers to database or database management system). In a typical workflow of GIS, the communication between these three components is essential. The relationship between these three components is that one component sends the request to other component and the other component responses the request. This making and fulfilling of request is called client-server computing model and used in online software applications and services. The component that makes request is called *client* and the component that fulfills the request is called *server*. In the typical GIS environment the user (s) sends the request through user interface of a GIS software to get the data and information from GIS databases. The user's request is sent to the logic component, which make further request to the database component and search for appropriate data. The database component process the query and respond back the result to the GIS user again through user interface of the GIS software.

The client-server model is based on distributed computing where one software program is installed in user's machine which communicates with another software program installed in server machine for the purpose of information exchange. The communication between client and server is established by following certain standards and procedure which are known as *Protocol*. In the client-server environment, once the client and server software processes recognize a common set of protocols, they need not depend on specific hardware or operating system requirements. The client and the server can exist in the same host computer, or be distributed across a wide variety of computers at extensively distant locations. The client-server computing is developed from a distributed database model where everything except data management is handled in the client computer. In this configuration the client needs more computing resources, both hardware and software. In recent technological development most of the application processes shifted towards server end, and distributed system. A distributed system is a '*thin client'* architecture where interface management is distributed between client and server and application logic and data management part is entirely handled by powerful server (s). The term 'thin client' refers to the least amount of processes that take place at the client end in comparison to the huge amount of processes that take place in the server end.

The client-server is a software defined model of computing, not a hardware defined. The client-server environment is generally heterogeneous in nature where the hardware platform and operating system of the client and server are not usually the same. In such cases, the communications mechanism may be further extended through a well-defined set of standard i.e. Application Program Interfaces (APIs) and Remote Procedure Calls (RPC). The client-server computing provides the capability to use the most cost-effective user interface, data storage, connectivity, and application services.

Web GIS uses client-server computation architecture in its implementation. Typical, the client is a GIS user and server is a host of GIS data and services. To understand it better, let us discuss the client-server computation architecture.

4.3 WEB GIS ARCHITECTURES

Client Server Architecture

Client-Server computation architecture is a shared architecture system where the computation loads of client and server are shared or divided. The client-server architecture is a centralized resource system where server holds all the resources. The servers can be centrally located or distributed across different physical location. The server receives many performances at its edge for sharing resources to its clients when requested. Client and server may be on the same or in a different network. The server is totally stable and scalable to execute the client's request and respond it accordingly. Client computers implement a bond to allow a computer user to request services of the server and to represent the results the server returns. Servers wait for requests to appear from clients and then return them. A server usually gives a standardized simple interface to clients to avoid a hardware/software confusion. Clients are located at workplaces or on personal machines, at the same time servers will be located somewhere powerful in the network. This architecture is useful mostly when clients and the server each have separate tasks that they routinely perform. Many clients can obtain the server's information concurrently, and also a client computer can execute other tasks, for instance, sending e-mails.The client-server system can be further implemented in a tier system such as 1-tier, 2-tier, 3-tier and n-tier.

1-tier architecture

In this category of client-server setting, the user interface, business logic and data logic are present in the same system. This kind of service is reasonable but it is hard to manage due to data variance that allots replication of work. Typically a desktop software comes under 1-tier architecture. One-tier architecture consists of layers.For example, Presentation, Business, Data Access layers within a single software package. The data is usually stored in the local system or a shared drive. Software applications which handle all the three tiers such as audio and video player software, MS Office come under one-tier application. Typical information flow in 1-tier architecture is shown in figure 4.1.

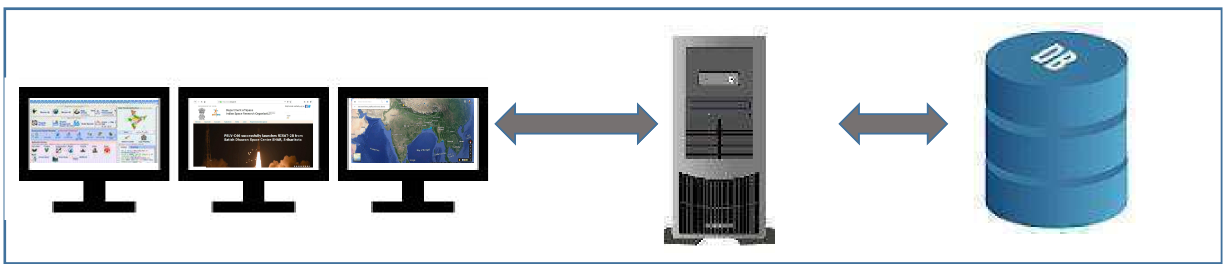


Figure 4.1- 1-tier architecture of client-server model

2-tier architecture

In this type of client-server environment, the user interface is stored at client machine and the database is stored on the server. Database logic and business logic are filed at either client or server but it needs to be maintained. The two-tier architecture is good for small applications and gives good performance when the number of users connected with server are less.

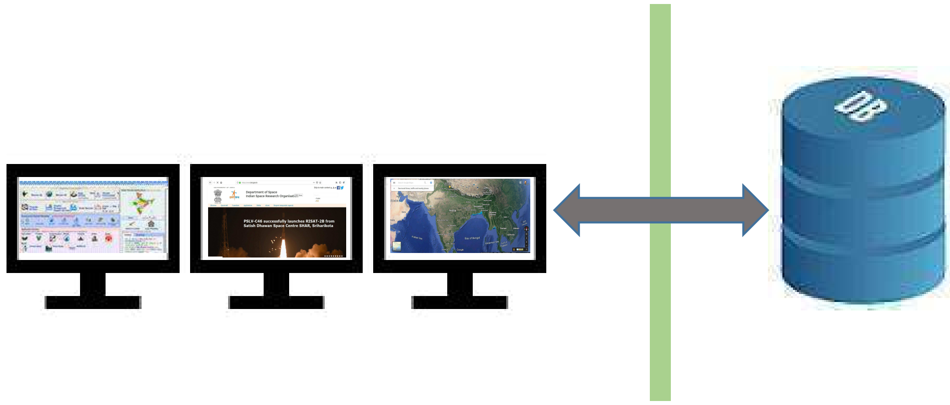


Figure 4. 2- 2-tier architecture of client-server model

If Business Logic and Data Logic are collected at a client side, it is named as fat or thick client architecture. If Business Logic and Data Logic are handled on the server, it is called thin client architecture. In two-tier architecture, client and server have to come in direct integration. If a client is giving an input to the server there shouldn't be any intermediate component to avoid loss of data and information. This is done for rapid results and to avoid confusion between different clients. For instance, online ticket reservations software use this two-tier architecture. Typical information flow in 2-tier architecture is shown in figure 4.2.

3-tier architecture

The Three-tier architecture is split into 3 components viz. the presentation layer (client tier), application layer (business tier) and database layer (data tier). The client system manages Presentation layer; the Application server takes care of the Application layer, and the Server system supervises Database layer.

In the three-tier system, the presentation element is located at the client machine and logic element at the server machine and data element is stored in the database server. In 3-tier architecture, the middleware is used that means client request goes to the server through that middle layer and the response of server is received by middleware first and then to the client. This architecture protects 2-tier architecture and gives the best performance. This system comes expensive but it is simple to use. The middleware stores all the business logic and data passage logic. The idea of middleware is to database staging, queuing, application execution, scheduling etc. A Middleware improves flexibility and gives the best performance.

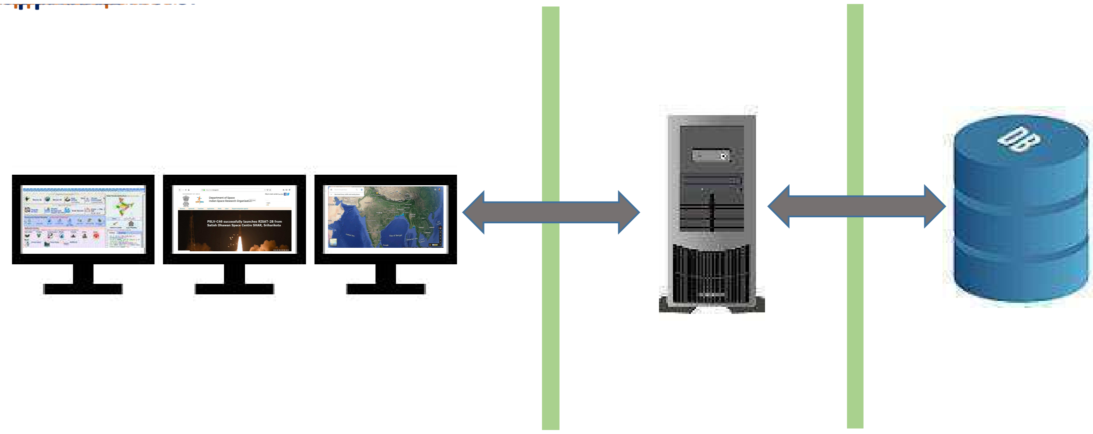


Figure 4. 3 3-tier architecture of client-server model

In most of the online services, there has been growing demands for the quick responses and quality services. Therefore, the complex client architecture is crucial for the business activities. Companies usually explore possibilities to keep service and quality meet to maintain its marketplace with the help of client-server architecture. The architecture increases productivity through the practice of cost-efficient user interfaces, improved data storage, expanded connectivity and secure services.From a hardware platform perspective, three-tier hardware architecture involves three classes of computers: the client (usually a PC or mobile devices); the middle tier (usually a workstation or server); and the back-end server (usually a mainframe computer). Selection of suitable hardware is also dependent upon the size of application, data and number of anticipated users connected with server at a time.

N-tier architecture

An n-tier system expands the logic or data element into multiple computers or multiple components. In an n-tier system each element of client-server architecture can be further extended to different machine as per requirement. For example the data server can be further extended into spatial database server to non-spatial database server and logic element into application server to map server in typical distributed GIS environment. The multiple data server can be linked together by data category server or data access program like ODBC, JDBC etc. The selection of suitable system configuration in client-server architecture depends upon the size of application, data and number of anticipated users connected with server at a time. The n-tier architecture can also configured logically or physically, where each component i.e. presentation, logic and data can reside at same machine but in different partition or at different machine.

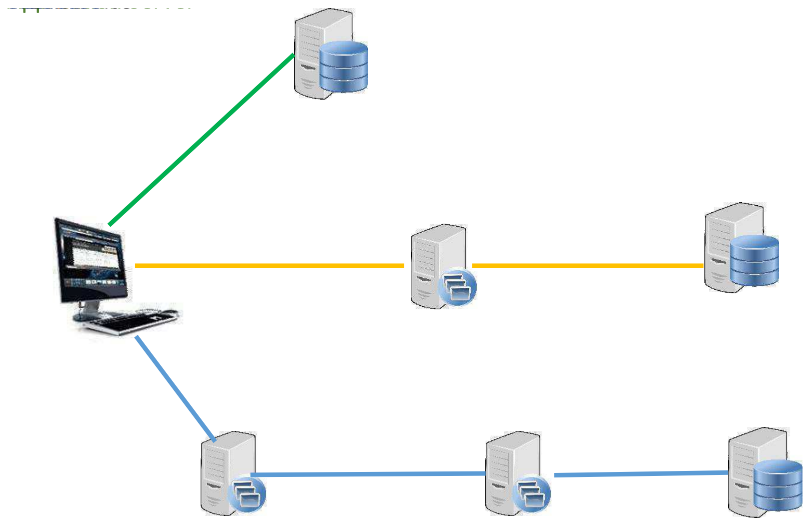


Figure 4.4- n-tier architecture of client-server model

1. Web GIS System Architecture

The Web GIS system architecture can be directly relate with client-server computation architecture. However, there is need for additional technology to provide a web mapping and geo-processing application (s), either at just the server-side or at both the server-side and the client-side. The web server itself is not able to do any geo-processing then it has to pass on requests from the client to map server or GIS server for geo-processing. There are several standard interfaces such as the Common Gateway Interface (CGI) and server Application Programming Interfaces (APIs) to enable the communication between Web server and map server. Figure 4. 5 illustrates a generalized view of a web-deployed mapping application.

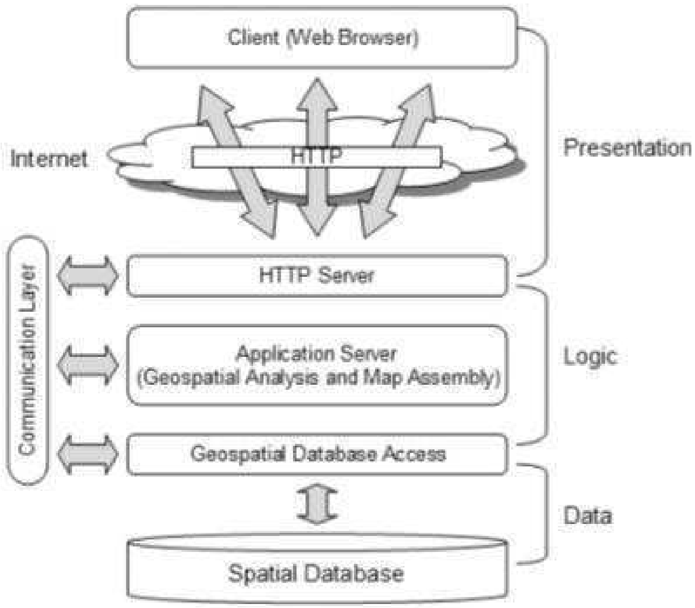


Figure 4. 5. Architectural Logic Diagram of Internet GIS

The interaction based on request and respond scenario of client-server. A server-side application communicates with the map server to provide the request information. The updated map is returned to the browser embedded in the HTML page or applet. Generally, the web system architecture can be categorized by means of the thickness of users:

Server-Side Internet-GIS (Thin Client)

Server side approach focuses on the server side application access & execution on a client-server architecture environment. In this approach the client has no specific applications installed in his computer to communicate with the server, instead the browser with the specified URL itself accesses the application on the server for analysis. The application resides on the server and the client is free from downloading or installing any applications or support tools from the server. The merit of this approach is that a slow Internet connection can be sufficient for the architecture. The only demerit is that the server needs to be fast and configured to attend the requests and queries from various clients at a time.This strategy is comparable to traditional terminal-to-mainframe models for running GIS on local networks. Little processing power is required of the client (a "dumb" terminal in the traditional model), only the ability to submit requests and display responses.

Advantages to Server Side Strategies

- If a high-performance server is used, users can access large and complex datasets that would be difficult to transfer across the Internet and process locally on the client.

- If a high-performance server is used, complex GIS analytical routines can be run quickly even by clients who lack access to sophisticated hardware.
- More control can be exerted over what the user is permitted to do with the data, perhaps also insuring that the data is used correctly.

Disadvantages to Server Side Strategies

- Every request no matter how small must be returned to the server and processed. Responses must then be returned to the client across the Internet.
- Performance will be affected by the bandwidth and network traffic on the Internet between the server and client particularly when responses involve transferring large files.
- Applications do not take advantage of the processing power of the user's own "client" computer, which is used merely to submit a request and display the response.

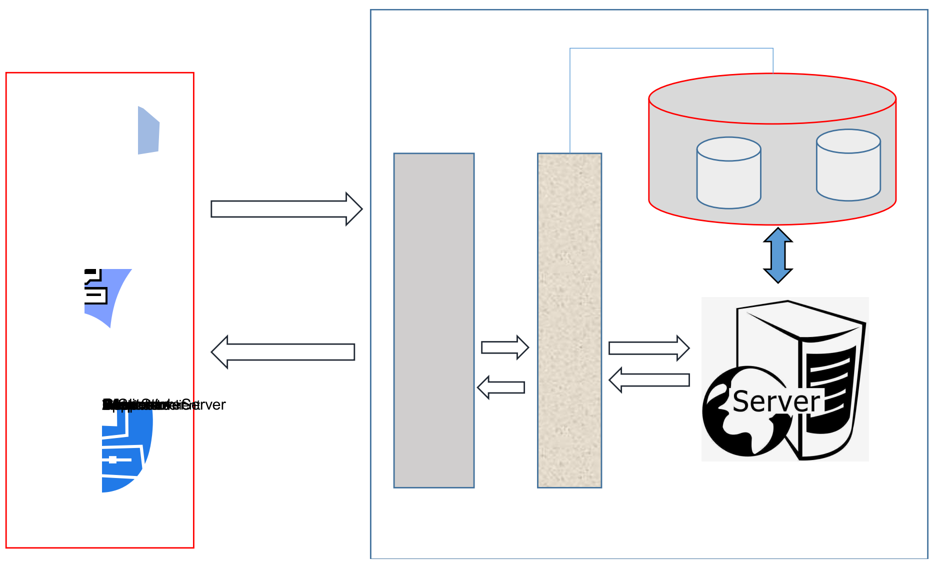


Figure 4.6- Thin Client Architecture

Client Side Approach (Thick Client)

In thick client based Internet architecture the client has to download an application to communicate with the server. The application after being downloaded and installed connects to the server to get access to the required data for analysis. The merits of this architecture is an enhanced application interface for the client user to communicate, since if such an application is present on the server side then the client needs a fast dedicated connection to the server. In the case of thick client:

- Client-side applications attempt to shift some of the work of processing requests to the user's computer, sometimes referred to as a thick client.
- Instead of forcing the server to do most of the work, some of the GIS capabilities are downloaded to the client, or reside there, and data is processed locally.

Advantages of Client Side Strategies

- Applications take advantage of the processing power of the user's own computer.
- The user can be given greater control of the data analysis process.
- Once the server has delivered its response, the user can work with the data without having to send and receive messages across the Internet.

Disadvantages of Client Side Strategies

- The response from the server may involve transferring large amounts of data as well as applets, causing delays.
- Large and complex datasets may be hard to process on the client if it is not very powerful. Complex GIS analytical routines may run more slowly on the client if it is not very powerful.
- Users may not have the training needed to employ the data and analysis functions properly.

Web services and Service-Oriented Architecture (SOA)

The trends in technological advancements in Information technology has made it possible to move towards serviceoriented architectures (SOA) and distributed computing to achieve high interoperability in data and information. In SOA, the services are intended to be independent building blocks that collectively represent a software application environment. The services have many unique characteristics that allow them to participate as part of a SOA.One of the unique quality is complete autonomy from other services which is important for interoperability. This means that each service is responsible for its own domain, which typically translates into limiting its scope to a specific business function (or a group of related functions) (http://www.informit.com).The most widely accepted and successful type of service in Information technology is the XML based Web service. XML stands for Extensible Markup Language which is a markup language similar to HTML and used for data definition while HTML is used for data presentation. The XML based service has two fundamental requirements:

- It communicates via Internet protocols (most commonly HTTP)

- It sends and receives data formatted as XML documents

The web services architecture has three roles: *a provider*, *a requestor*, and *a broker*. The provider creates the web service and makes it available to clients or users who want to use it as XML. A requestor is a client application that consumes the web service at other end. The broker, such as a service registry, provides a way for the provider and the requestor of a web service to interact as a catalogue. The provider, requestor, and broker interact with each other through the three major operations viz. *publish*, *find*, and *bind*. A provider informs the broker about the existence of the web service by using the broker's publish interface to make the service accessible to clients. The information published describes the service and specifies the location of the service as where the service is located. The requestor consults the broker to locate a published web service through uniform resource locator (URL) or web address. With the information it gained from the broker about the web service, the requestor is able to bind, or invoke, the web service. The architecture of SOA is shown Figure 4.7.

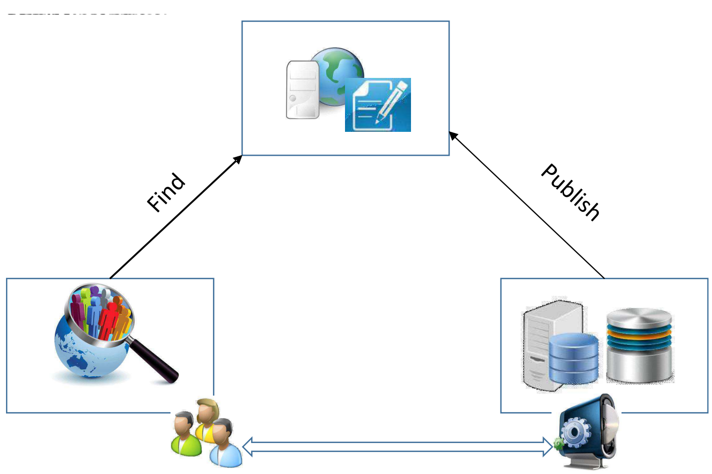


Figure 4.7- Service Oriented architecture

The three essential components shown in Figure 1 can be described as:

- *Service provider*- It publishes services to a registry and makes it available on the Internet for the access of the service consumers.
- *Service requester (client)* - It performs service discovery operations on the service registry in order to find the required service and also get the access to the services.
- *Service registry* – It helps service providers and service requesters to find each other by acting as a registry of the services.

The Web services use SOAP (Simple Object Access Protocol), WSDL (Web Services description Language), UDDI (Universal Description, Discovery and Integration), and WSIL

(Web Services Inspection Language) for communication between two applications. SOAP works with standard Web protocols including XML, HTTP, and TCP/IP, as well as WSDL. Web Services are published on the UDDI registry for discovery by the client. A WSDL document defines characteristics/ and properties of the target service through which the client can knows what the service does.

There are many advantages using SOA. One of the most important advantage is a SOA implementation provides interoperable environment, which is based on reusability and standardized components. SOA creates a framework for software application development in distributed computing environment where the data is available in decentralized rather than locally. This means that inconsistency in local copies of the data and the data available in original repositories are avoided. This approach increased the quality of the outputs in the applications where data from various sources are used which is originally available in different formats. The SOA approach to system development can produce systems that can be flexibly adapted to changing requirements and technologies, and offers easier maintainable and more consistent systems of data and functionality.

Standards for Web Services

Extensible Markup Language (XML) is designed for data definitions and is the base for web services. Web services are XML based software systems designed to support interoperable communications between two systems over a network. The interaction between web services are based on four major internet protocol viz. HTTP, WSDL, SOAP and UDDI and the Web services are recognized by URL. The brief description about these protocols and standards are given here:

- *HyperText Transfer Protocol* (HTTP) is the low-level internet protocol used as transport layer and is responsible for transporting messages between networks based software applications.
- In web services, *Simple Object Access Protocol (SOAP)* is a communication protocol used for exchanging data and message between two systems via HTTP.
- *Web Services Description Language* (WSDL) is used to describe type and other characteristics of web service using public interface.
- *Universal Description, Discovery, and Integration* (UDDI) specifications are used as a service registry. Through UDDI the service providers advertise the existence of their services and by requesters to search and discover already registered services.

The most important advantage in web service implementation is the implementation logic is not known to the users which allows the service to be used on different platforms. The communication between web services are through XML based SOAP protocol and regardless of with which programming language it was developed. These properties allow Web Services based frameworks to be loosely coupled and component oriented. The main advantage of Web services is that the service can be used remotely without the user's actual knowledge and intervention and by multiple users at the same time, eliminating the need for constant updates to locally installed software.

Web services for GIS

In geospatial domain web services are very important to achieve interoperability in data and information available with different data providers. Today, the geospatial services available in internet through various geo-portals are increasing rapidly. There is a need of methodology to locate desired services that provide access, data discovery and analysis capabilities for geospatial data. The interoperability of services across organizations and providers is important for seamless integration and sharing of spatial data from a variety of sources. Different organizations and commercial vendors develop their own data standards and storage structures for geo-spatial data. If GIS services are not interoperable, these data sets cannot interact or overlaid to each other even though they are in the same organization or they belong to same commercial vendor. To solve the interoperability problems in GIS, the Open Geospatial Consortium (OGC) has introduced data and service standards by publishing specifications for the GIS services. OGC is a not-for-profit, international, voluntary, consensus standards organization founded in 1994. The major objectives of OGC are to lead in the development, promotion and harmonization of open geospatial standards. OGC have around 500 members from industry, government, research and university across the world. The GIS service specifications developed by the OGC are based on Service Oriented Architectures (SOA). Such systems unify distributed services through a message-oriented architecture. Web Service standards are a common implementation of SOA ideals (Sayar et al, 2005).

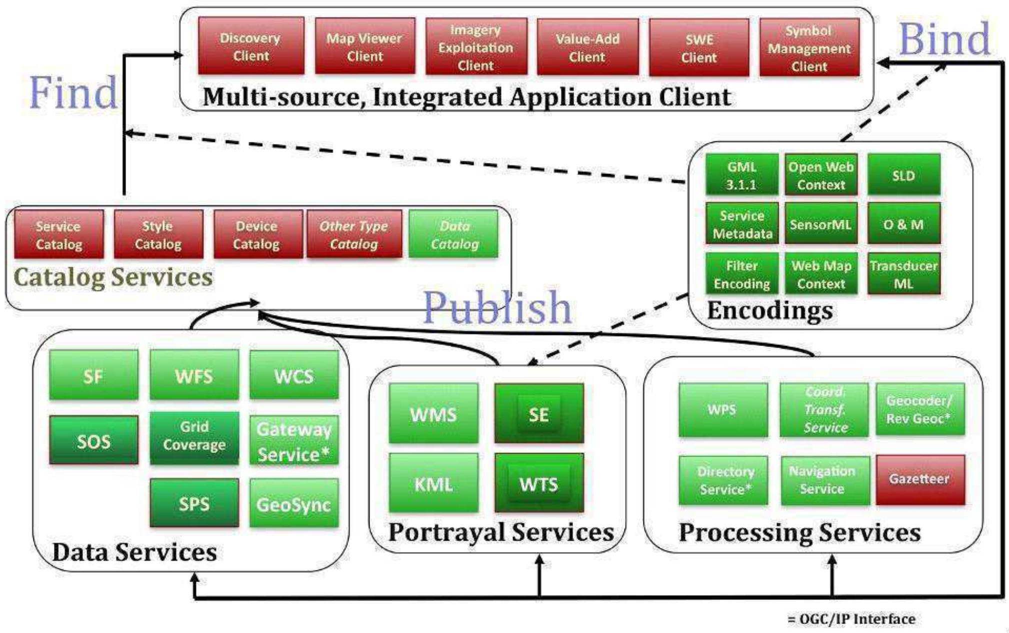


Figure 4.8- Web service framework of OGC (Source http://live.osgeo.org)

The OGC GIS services can be grouped into six major categories (http://www.opengeospatial.org):

Catalogue Services

Registry or Catalog Service allows users and applications to classify, register, describe, search, maintain, and access information about Web Services. Some of the examples of catalogue services are:

- CS Core
- CS-W ebRIM
- CS-W 19115/19119
- CS-W ebRIM for EO

Processing Services

Processing Services provide operations for processing or transforming data in a manner determined by user-specific parameters. The processing services provide generic processing functions such as projection and coordinate transformation, raster and vector data processing for user defined data sets etc. Some of the examples of processing services are:

- ***OpenLS Core Services:***

  Tracking Service Interface Standard version 1.0.1. The OpenGIS Tracking Service Interface Standard supports a very simple functionality allowing a collection of movable objects to be tracked as they move and change orientation.

- ***Sensor Planning Service (SPS):***

  SPS defines interfaces for queries that provide information about the capabilities of a sensor and how to task the sensor.

- ***Web Processing Service (WPS):***

  WPS provides rules for standardizing how inputs and outputs (requests and responses) for geospatial processing services, such as polygon overlay. The standard also defines how a client can request the execution of a process, and how the output from the process is handled. It defines an interface that facilitates the publishing of geospatial processes and clients' discovery of and binding to those processes. The data required by the WPS can be delivered across a network or they can be available at the server.

- ***Coordinate Transformation Service (CTS):***

  CTS provides a standard way for software to specify and access coordinate transformation services for use on specified spatial data. This standard addresses a key requirement for overlaying views of geo-data ("maps") from diverse sources: the ability to perform coordinate transformation in such a way that all spatial data are defined relative to the same spatial reference system.

- ***Web Coverage Processing Service (WCPS)***

  This specification is used for online processing of gridded data i.e. raster and statistical data.

Encoding

These specification defines Symbology Encoding, an XML language for styling information that can be applied to digital Feature and Coverage data. Some of the examples of encoding standards are:

- ***Geography Markup Language (GML):***

GML is an XML grammar for expressing geographical features. GML serves as a modeling language for geographic systems as well as an open interchange format for geographic transactions on the Internet. As with most XML based grammars, there are two parts to the grammar – the schema that describes the document and the instance document that contains the actual data. A GML document is described using a GML Schema. This allows users and developers to describe generic geographic data sets that contain points, lines and polygons

- ***CityGML:***

It is an open data model and XML-based format for the storage and exchange of virtual 3D city models. The aim of the development of CityGML is to reach a common definition of the basic entities, attributes, and relations of a 3D city model.

- ***GML Simple Features***
  - ***Filter Encoding (FE):***

    This standard describes an XML and KVP encoding of a system neutral syntax for expressing projections, selection and sorting clauses collectively called a query expression. These components are modular and intended to be used together or individually by other standards which reference this International Standard.
  - ***GML in JPEG 2000:***

    GML in JPEG 2000 for geographic imagery encoding standard defines the means by which the GML Standard is used within JPEG 2000 images for geographic imagery. The standard also provides packaging mechanisms for including GML within JPEG 2000 data files and specific GML application schemas to support the encoding of images within JPEG 2000 data files
- ***KML:***

KML is an XML language focused on geographic visualization, including annotation of maps and images. Geographic visualization includes not only the presentation of graphical data on the globe, but also the control of the user's navigation in the sense of where to go and where to look.

- ***Observations & Measurements (O&M):***

This standard specifies an XML implementation for the OGC and ISO Observations and Measurements (O&M) conceptual model, including a schema for Sampling Features. This encoding is an essential dependency for the OGC Sensor Observation Service (SOS) Interface Standard.

- ***Sensor Model Language (SensorML):***

The primary focus of the Sensor Model Language (SensorML) is to provide a robust and semantically-tied means of defining processes and processing components associated with the measurement and post-measurement transformation of observations.

- ***Symbology Encoding (SE):***

It defines an XML language for styling information that can be applied to digital geographic feature and coverage data. SE is independent of any OGC Web Services descriptions and could therefore be used to describe styling information in non-networked systems such as desktop geographic information systems.

- ***Styled Layer Descriptor (SLD):***

SLD addresses the need for users and software to be able to control the visual portrayal of the geospatial data. The ability to define styling rules requires a styling language that the client and server can both understand. This encoding Standard defines an encoding that extends the WMS standard to allow user-defined symbolization and coloring of geographic features and coverage data.

- ***SWE Common:***

This standard defines low level data models for exchanging sensor related data between nodes of the Sensor Web Enablement (SWE) framework. These models allow applications and/or servers to structure, encode and transmit sensor datasets in a self-describing and semantically enabled way.

- ***Transducer Markup Language (TML):***

TML is a standard developed to describe any transducer (sensor or transmitter) in terms of a common model, including characterizing not only the data but also metadata defined in XML.

Data Services

Data Services are tightly coupled with specific data sets and offer access to full or a portions of that data. The data services provide access to the data without any loss where a simple URL returns the actual data in the form of XML without any dependency of format and structure of original data at user end.

- ***Simple Features Specification (SFS):***

SFS specifies a common storage model of mostly two-dimensional geographical data (point, line, polygon, multi-point, multi-line, etc.).

- ***Web Coverage Service (WCS):***

This standard specifies how a Web Coverage Service (WCS) offers multi-dimensional coverage data for access over the Internet. This specification allows to publish raster data (satellite and maps) as web service in Internet. Services implementing this language provide access to original or derived sets of geospatial coverage information, in forms that are useful for client-side rendering, input into scientific models, and other client applications.

- ***WCS Transactional:***

WCS-T (T standing for transactional) establishes how to upload complete coverages to a server or modify existing coverages on a server.

- ***Sensor Observation Service (SOS):***

SOS provides an API for managing deployed sensors and retrieving sensor data and specifically "observation" data. Whether from in-situ sensors (e.g., water monitoring) or dynamic sensors (e.g., satellite imaging), measurements made from sensor systems contribute most of the geospatial data by volume used in geospatial systems today

- ***Table Join Service (TJS):***

This OGC standard defines a simple way to describe and exchange tabular data that contains information about geographic objects.

- ***Web Feature Service (WFS):***

This specifications defines an interface for specifying requests for retrieving geographic features across the Web using platform-independent calls. The WFS standard defines interfaces and operations for data access and manipulation on a set of geographic features, including:

  - Get or Query features based on spatial and non-spatial constraints
  - Create a new feature instance
  - Get a description of the properties of features
  - Delete a feature instance
  - Update a feature instance
  - Lock a feature instance

Portrayal Service

These services are used for simple data visualization like map rendering and cartographic representation of the maps. Some of the examples of portrayal services are:

- ***Web Map Service (WMS):***

Web Map Service Interface Standard (WMS) provides a simple HTTP interface for requesting geo-registered map images from one or more distributed geospatial databases. A WMS request defines the geographic layer(s) and area of interest to be processed. The response to the request is one or more geo-registered map images (returned as JPEG, PNG, etc.) that can be displayed in a browser application. The interface also supports the ability to specify whether the returned images should be transparent so that layers from multiple servers can be combined.

- ***Web Map Tiling Service:***

WMTS is an extension of WMS which is used to serve pre-rendered geo-referenced map tiles from server to client. This standard is very useful for enhancing the performance of web applications specifically for geo-visualization.

Others Services

- GeoXACML
- Geospatial Objects
- OWS Common

Architecture of Mobile GIS:

Mobile GIS is an extension of Web GIS where GIS services are made available in mobile client (s). The mobile GIS architecture is a combination of client at mobile and host at server level. The data and information is published at server end for mobile user similar to web based GIS application. The only major configuration required at server end is to tune the data size for low storage and limited computation devices like mobile and PDA. The mobile GIS applications are best suited for smartphone devices where the multimedia components with GPS facility are available for the users. The smartphone devices are an electronic device in between computer notebook (laptop or portable computer) and the mobile phones. The smart mobile devices are having major functionalities of laptops and personal computers and an additional feature of mobility and wireless communication (GSM, CDMA etc.) of mobile phone. The major limitation in these devices is its small screen size and limited computation capabilities in terms of storage, processing and memory. The GIS operations in smart mobile phones are revolution of ICT which has many exciting new applications of GIS in addition to traditional GIS applications. The mobile users are moving in a geographic space and they know their positions and they have access to the wide available geographical data and information from variety of servers. The mobility is “key” for mobile GIS applications. The typical architecture of Mobile GIS is shown in Figure 4.9.

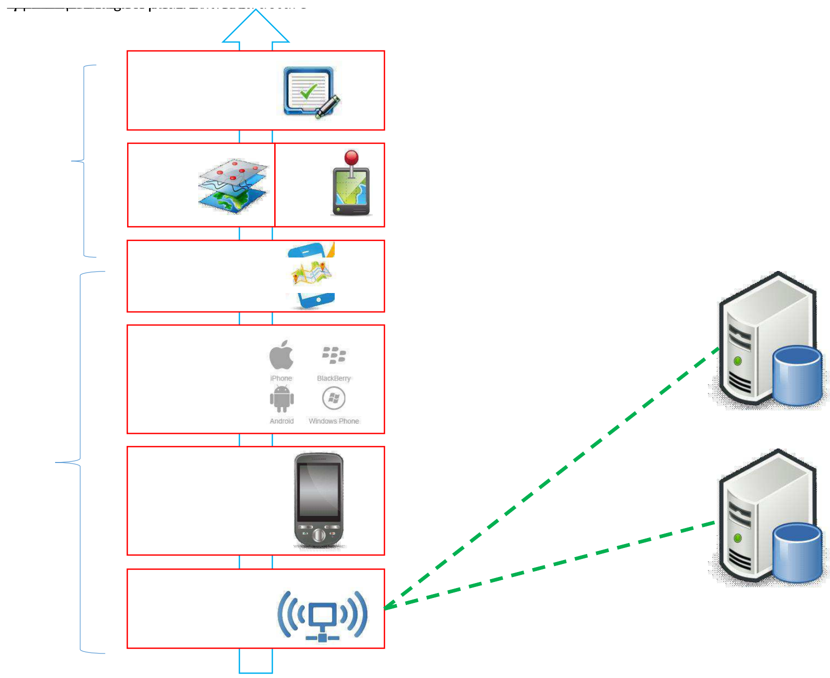


Figure 4.9- typical architecture of Mobile GIS

The mobile GIS is a miniature form of traditional GIS which have five major components namely hardware, software, data, applications, and users. The mobile GIS has the similar GIS components, smart mobile phone as hardware, GIS software for mobile, limited data, and theme specific applications.

The typical architecture of mobile GIS is shown in Figure 1 which have six major layers at client end and one composite layer at server end. The hardware of mobile device is having mobile operating system like Android, IoS, and Windows etc., on which a Mobile GIS software will be installed. The mobile GIS software will use storage and other computation requirement from mobile device and communicate with web and mobile servers through wireless communication like GPRS and CDMA. The theme specific GIS applications like mobile GIS for mapping, mobile GIS for field data collection, mobile GIS for spatial analysis and mobile GIS for Location Based Services (LBS) etc., can be build. The development of mobile GIS applications and software are the top layers of the system and it is variable according to the place of interest and the application to the user. The mobile platform is the combination of the hardware architecture of the mobile device and its operating system, both of them are the base of the mobile GIS that is installed and used on this platform.

4.4 SUMMARY

The Client-server computation architecture is designed and developed for data and information sharing between the users. The client-server computation is having three essential componentsi.e. presentation, logic and data. The software application developed using this computation architecture works in the principle of *request* and *response*. The component that makes request is called *client* and the component that fulfills the request is called *server*. The Web GIS works on client-server computation architecture where different components of GIS are made available either at central or distributed systems. The client side or server side web GIS applications are developed based on the requirements of the users. It is designed to satisfy the requirement of multiple users simultaneously. To make the web GIS an interoperable software application, the concept of Service Oriented Architecture (SOA) is used. The Open Geospatial Consortium (OGC) has developed variety of web service standards for GIS data and information. These GIS web services provides a means of GIS data access and sharing. In the recent past the computation requirements in client-server systems is moving towards mobile based solutions. The mobile GIS is another emerging area where GIS data and information services are made available in mobile platforms such as cell phones, PDA, tablets etc.

4.5 GLOSSARY

1. WMS- Web Map Service
2. WFS- Web Future Service
3. TJS- Table Join Service
4. WCS- Web Coverage Service
5. SPS- Simple Features Specifications
6. TML- Transducer Markup Language

4.6 ANSWER TO CHECK YOUR PROGRESS

1. Define catalogue services?
2. Define web processing services?
3. Define symbology encoding?
4. Define sensor model language?
5. What do you understand by Mobile GIS?

4.7 REFERENCES

- Aydin, G., 2007a. Service Oriented Architecture for Geographic Information Systems Supporting Real Time Data Grids, Ph.D. Thesis, Indiana University, IN, USA, pp. 12-47
- Nicolai M. Josuttis 2007., SOA in Practice the Art of Distributed System Design, O'Reilly ISBN-10: 0-596-52955-4, ISBN-13: 978-0-596-52955-0
- Fedra, K. (1997). Spatial Decision Support in Resources Management and Environmental Applications, Proceedings of the Eleventh Annual Symposium on Geographic Information Systems held in Vancouver, British Columbia, Canada, 302-305.
- H Singh, HC Karnatak and RD Garg, 2018. An automated and optimized approach for online spatial biodiversity model: a case study of OGC web processing service, Geocarto International 34 (2), 194-214.
- Hariom Singh, RD Garg, Harish C Karnatak and Arijit Roy, 2018. Spatial landscape model to characterize biological diversity using R statistical computing environment Journal of environmental management, vol. 206, pp-1211-1223.
- Harish C Karnatak, Hariom Singh and RD Garg, 2018. Online spatial data analysis and algorithm development for geo-scientific applications using remote sensing data", Proceedings of the National Academy of Sciences, India Section A: Physical Sciences, vol. 87, Issue 4, pp- 701-712.
- Harish Chandra Karnatak, R Shukla, VK Sharma, YVS Murthy, V Bhanumurthy, 2012 "Spatial mashup technology and real time data integration in geo-web application using open source GIS–a case study for disaster management", Geocarto International 27 (6), 499-514.
- Harish Chandra Karnatak, S Saran, K Bhatia, PS Roy, 2007 "Multicriteria spatial decision analysis in web GIS environment", Geoinformatica 11 (4), 407-429.
- Hendler, J. (1999). Is There an Intelligent Agent in Your Future? Nature, Vol. 398, No. 6723, March 11, 1999.
- K. Sahina and M. U. Gumusay, 2008, "Service Oriented Architecture (SOA) based web services for Geographic Information Systems", International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B2. Beijing 2008.
- Karnatak H.C, S. Sameer, K. Bhatia and Roy P.S., 2007, Multicriteria Decision Analysis in Web GIS Environment, Geoinformatica, (2007) 11, pp: 407-429: Springer Science Publication DOI 10.1007/s10707-006-0014-8.
- Karnatak Harish, Karamjit Bhatia and Sameer Saran, (2008) "Multi-criteria decision analysis using Spatial Compromise Programming", Proceedings of the 2nd National Conference- INDIACom-2008, on Computing for Nation Development, ISBN No-ISSN 0973-7529, ISBN 978-81-904526-2-5, pp-77-62. February 08 – 09, 2008, New Delhi India.

- Lee K. 2010. Intelligent Geo-Web Services Based On Hybrid-Mashup Using Open Source Geo-Spatial Software, Web MGS 2010 1st International Workshop on Pervasive Web Mapping, Geoprocessing and Services XXXVIII-4/W13, ISSN Number: 1682-1777.
- Open Geospatial Consortium (OGC) official website- http://www.opengeospatial.org (last accessed 2nd April, 20014).
- Peasrson Informit official website- http://www.informit.com/articles/printerfriendly/336265 (last accessed 2nd April 2014).
- Sayar, A., 2008, GIS Service Oriented Architecture, Community Grids Laboratory, Indiana, USA.
- Sayar, A., Aktas, M., Aydin, G., Pierce, M. and Fox G., 2005, Developing a Web Service-Compatible Map Server for Geophysical Applications Technical report, Indiana University, Indiana, USA.
- Shan Yen-Ping and Earle Ralph H. (1998). Enterprise computing with objects: from client/server environments to the Internet, publisher-Wesley Longman Publishing Co., Inc. Boston, MA, USA, ISBN: 0-201-32566-7.
- Wikipedia, the free encyclopedia- http://en.wikipedia.org (Accessed 19 March 2012).
- World Wide Web Consortium (W3C) official website- http://www.w3.org/ (Last accessed on 2nd April 2014).
- Zhong-Ren Peng, Ming-Hsiang Tsou, (2003). Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks, ISBN: 0-471-35923-8m March 2003.
- https://www.softwaretestingmaterial.com
- https://elearning.iirs.gov.in
- https://bhuvan.nrsc.gov.in
- https://www.iirs.gov.in

4.8 TERMINAL QUESTIONS

Q1. Write an essay on web GIS architectures and its type.

Q2. Define server side internet GIS? Explain advantages and disadvantages of server side strategies.

Q3. Define client side approach? Explain advantages and disadvantages of client side strategies.

UNIT 5 : WEB GIS DEVELOPMENT

5.1 OBJECTIVES
5.2 INTRODUCTION
5.3 WEB GIS DEVELOPMENT
5.4 SUMMARY
5.5 GLOSSARY
5.6 ANSWER TO CHECK YOUR PROGRESS
5.7 REFERENCES
5.8 TERMINAL QUESTIONS

5.1 OBJECTIVES

The evolution of the internet allows us to get aware about anything we wish, we are at our will that what type of information we want about something and it is all possible through the process of internet. The Internet allows people from all walks of life to access geospatial data and provides a platform for geo-related data processing that is not limited by location or such things. Geographic information system is a framework which is designed to capture, store, manipulate, analyze, and present spatial or geographic data.GIS is computer-based tool which examines and manages spatial relationships, patterns, and trends in geo spatial world. It uses data that is attached to a unique location. GIS software has enabled user to view spatial data in its proper format as a result, the interpretation of spatial data has become easy and simple to understand. But it is not obvious that, everyone has access to GIS, nor would they be able to spend the time necessary to use it efficiently. The development of Web GIS as an important feature of GIS which led it to become cheap and easy way of disseminating geospatial data and processing tools. Because Web GIS is web-based, it enables users to create a wide range of applications that can reach a considerably bigger audience than traditional disconnected GIS software placed on a user's computer. Instead of being limited to creating and sharing static maps (PDF, PNG, etc.), users may now build dynamic, shareable (access-controlled) online maps and mapping apps that can be accessed via any web browser, mobile device, or desktop GIS software. That why many organizations were interested to distribute maps and processing tools without time and location restriction to users. The Internet technology has successfully made its path to various organizations such as government and private as well as in households too. The ability to get information through internet made spatial data providers to explore the internet resources for disseminating spatial information. To deliver a successful online GIS implementation, it is necessary to think of it as a process rather than a step. The implementation should also respect the availability of technology and the application requirement. In fact, Web GIS is more than a new and advanced technology; it's a new way of thinking about how to create a modern GIS. It's significantly altering our understanding of GIS and the role of GIS experts.

5.2 INTRODUCTION

Web GIS (web geographic information system) is an advance form of geographic system which is available online which means it can be availed through online portals. Web GIS is exactly what it sounds like: Geographic Information Systems on the World Wide Web. It is the most advanced utilized tool of geospatial techniques. A web GIS is a computer software and hardware configuration that allows the sharing of maps, spatial data, and geographic processing operations throughout one' sown network and beyond using common web communications protocols such as HTTP. It is a type of distributed information system, which comprises at least a server as well as a Clint, where the server will be the GIS server and client is in the form of web browser, desktop application, or simply an application of mobile in general we can say, a web GIS system that uses web technologies for management of information among different components of a system. It involves creating, extending, utilizing, web GIS or web mapping solutions to solve some more specific types of problems, or to build complete applications, to consume or produce data and geospatial processing services. As the expansion of internet occurs and availability of web GIS or web mapping became possible, web GIS programming becoming a commonly required skill set in many organizations.

5.3 WEB GIS DEVELOPMENT

Web GIS programming is a type of software development that provides a means of handling internet, browser-based software application development tasks which require unique solutions to web GIS or web mapping problems. Web GIS makes it feasible to perform analytics on geographical or geographic data in ways that were previously impossible. To answer a preset set of questions, spatial data has to be processed, edited, updated, and extracted previously. Now, the data is translated into web maps or services, which are then mashed up with multiple layers to create a web GIS, which can answer any sort of question. The data is ready to utilize and dynamically respond to questions, and the solution is comprehensive. Data does not need to be processed for each individual inquiry any longer. In simplest term Web GIS is a two-tier architecture that consists of a server and one or more clients. Web GIS is far more adaptable, responsive, customizable, and capable of facilitating workflow. The Geographic Information System (GIS) has the ability to do more than just map data. GIS provides a wide range of analytical capabilities which is beyond mapping. Web GIS is enhancing the power of GIS to a wider group and assisting in the decision-making process.

"web GIS is a type of distributed information system. Comprising at least a server and a client, where the server is GIS server and the client is web browser, desktop application,or mobile application.in its simplest form,web GIS can be defined as any gis that uses web technology to communicate between a server and a client." [ESRI,2016].

"the term WEB GIS is being tossed all over the place right now but true meaning of the term may be very different than what you understand it to be. Web GIS explained in its simplest form, web GIS can be any GIS that uses web technology to communicate between a server and a client". [SSp,2017]

"Web GIS is a distributed computer application system for storing, managing, analyzing, publishing and sharing geographic information based on the HTTP protocol in the Internet or Intranet environment."

"Web Gis/Internet GIS mainly as a framework of "network-based GIS that uses the Internet to access remote geographic information and geo processing tools". (Peng et alii, 2003)

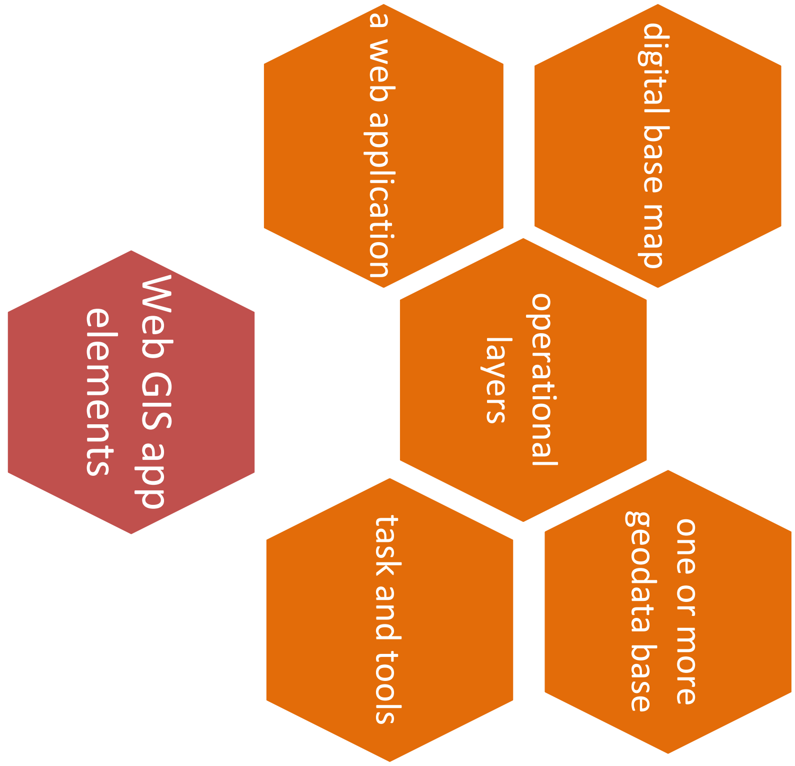
"the application for processing and the database as the different elements that compose the so-called WEB GIS client-server structure". (Penev, 2006)

"Also known as 'Internet GIS'; is a new technology that is used to display and analyze spatial data on the internet. It combines the advantages of both internet and GIS. It offers public a new means to access spatial information without owning expensive GIS software".

The following are the key elements essential to the web GIS:

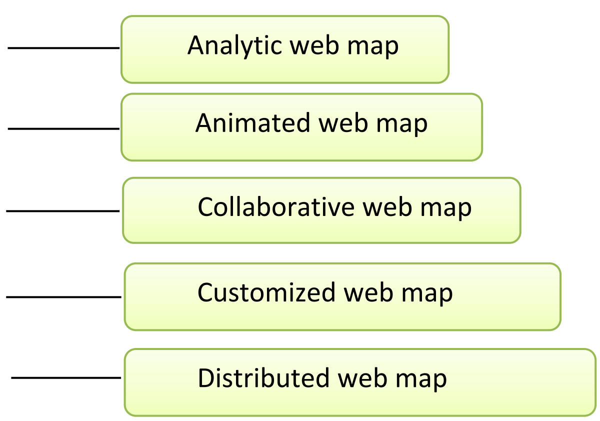
- ✓ The server has a URL so that client can find on the web.
- ✓ The client relies on HTTP specification to send request to server.
- ✓ The server performs the requested GIS operations and sends response to the client via HTTP.
- ✓ The format of response sent to the client can be in many formats, such as HTML, binary image, XML (Extensible Markup Language), or JSON (Java script object notation).

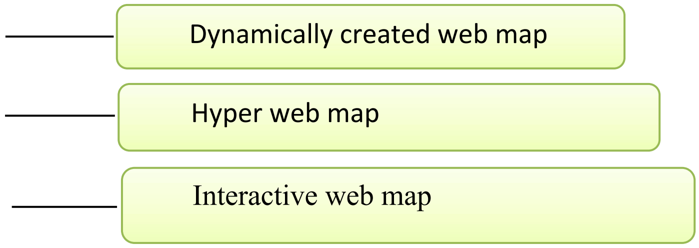
In the document titled "A framework for deploying web GIS application"- it has been pointed out that there are five essential elements in a Web GIS application. These includes-



Development of web and expansion of the internet technology is the duo which brings revolution in the field of GIS, these two provide capabilities which greatly help geoscientist. A first classification of web maps has been made by "Kraak". He distinguished static and dynamic web maps and further distinguished interactive and view only web maps. However, today in the light of an increased number of different web map types, this classification needs some revision. Today, there are additional possibilities regarding distributed data sources, collaborative maps, personalized maps, etc.

Types of web maps-





a). Analytic web map:

These web maps offer GIS analysis, either with geodata provided, or with geodata uploaded by the map user. As already mentioned, the borderline between analytic web maps and web GIS is blurry. Often, parts of the analysis are carried out by a server-side GIS and the client displays the result of the analysis. As web clients gain more and more capabilities, this task of sharing may gradually shift.

b). Animated web map:

Animated Maps show changes in the map over time by animating one of the graphical or temporal variables. Various data and multimedia formats and technologies allow the display of animated web maps: SVG, Adobe Flash, Java, Quick Time, etc., also with varying degrees of interaction. Examples for animated web maps are weather maps, maps displaying dynamic natural or other phenomena (such as water currents, wind patterns, traffic flow, trade flow, communication patterns, social studies projects, and for college life, etc.).

c). Collaborative web map:

Collaborative maps are still new, immature and complex to implement, but show a lot of potential. The method parallels the Wikipedia project where various people collaborate to create and improve maps on the web. Technically, an application allowing simultaneous editingacross the web would have to ensure that geometric features being edited by one person are locked, so they can't be edited by other persons at the same time. Also, a minimal quality check would have to be made, before data goes public. Some collaborative map projects:

- Google map maker (till march 2017)- Google maps
- Open street map
- Wikimapia
- meta: Maps- survey of Wikimedia map proposal on Wikipedia: meta

d). Customized web map:

Web maps in this category are usually more complex web mapping systems that offer APIs for reuse in other people's web pages and products. Example for such a system with an API for reuse is the Open Layers Framework, Yahoo! Maps and Google Maps.

e). Distributed web map:

These are maps created from a distributed data source. The WMS protocol offers a standardized method to access maps on other servers. WMS servers can collect these different sources, re-project the map layers, if necessary, and send them back as a combined image containing all requested map layers. One server may offer a topographic base map, while other servers may offer thematic layers. Web maps in this category are usually more complex web mapping systems that offer APIs for reuse in other people's web pages and products. Example for such a system with an API for reuse is the Open Layers Framework, Yahoo! Maps, these are maps created from a distributed data source. The WMS protocol offers a standardized method to access maps on other servers. WMS servers can collect these different sources, reproject the map layers, if necessary, and send them back as a combined image containing all requested map layers. One server may offer a topographic base map, while other servers may offer thematic layers. Web maps in this category are usually more complex web mapping systems that offer APIs for reuse in other people's web pages and products. Example for such a system with an API for reuse is the Open Layers Framework, Yahoo! Maps and Google Maps.

f). Dynamically created web map:

These maps are created on demand each time the user reloads the web pages, often from dynamic data sources, such as databases. The web server generates the map using a web map server or self-written software. Some applications refer to depictions as hyper maps. One of the examples is- Bhoosampada by Indian Space Research Organizations.

g). Hyper web maps:

Any approach offering the planar presentation of a portion of an n-dimensional orthogonal web map structure with the option to choose the axes for depiction from the dimensions.

h). Interactive web maps:

Interactivity is one of the major advantages of screen-based maps and web maps. It helps to compensate for the disadvantages of screen and web maps. Interactivity helps to explore maps, change map parameters, navigate and interact with the map, reveal additional information link to other resources, and much more. Technically, it is achieved through the combination of events, scripting and DOM manipulations.

In this age of information where thousand tons of data is generated every day, there is need to utilize these data to solve day to day problems encountered every day in our life. The evolution of Web GIS has been steady, and this is in line up of progress with the advancement of hardware and software. The web provides the end user with the ability to interact with the data and make better decisions from it.

A lot of attempts have been made to track the evolution of web mapping. This is usually done in relation to technological advancements, either on a broad level or in relation to specific technology efforts. The vivid history of web mapping all started with the successful implementation of Global Positioning System, thereafter development of internet as well as 1st web map server. The table shows the year along with achievement in the field of development of web GIS: -

Table 5.1

1989	Global positioning system
1991	Internet
1993	1st web map server
1994	Genasys Spatial web broker
1995	Multimap formed (United Kingdom maps on the www)
1996	map Quest formed
1997	Map info launch store locater application ESRI Map Objects IMS MapInfo launched (Map Xtreme)
2000	Open Source starts (Map Server)

2002	Arc Web service
2004	Yahoo maps, Google maps, Open street Maps
2005	Google release maps API
2008	Rest API,s
2009	ArcGIS.com
2010	More than 30,000 sites use google map API

The advent of web mapping is a key new trend in cartography that has just emerged. Cartography was formerly limited to a small number of enterprises, institutions, and mapping agencies, needing expensive and complicated hardware and software, as well as expert cartographers and geomatics engineers. With web mapping, publicly available mapping tools and geo data might allow any knowledgeable individual to create online maps, with expensive geo data and technical complexity (data harmonization, missing standards) remaining as two of the major hurdles preventing web mapping from becoming popular. The cheap and simple exchange of geo data over the internet permits for the integration of disparate data sources, bringing up possibilities that go beyond what fragmented data storage can provide. Anyone with a basic understanding of geo data and the necessary infrastructure may become a geo data provider. The very first appearance of web GIS can be found in the year of 1993 with the introduction of Xerox Corporation's interactive map viewer, allows for the retrieval of geographical data over the internet and user engagement with that data in the form of data browsing, modification, or spatial analysis, depending on the interface. One such application is made possible by a distributed information system, in which a GIS server interacts with a web browser, desktop, or mobile application to transmit geographic information to the user via web technologies. While the first development of Web GIS were limited to simplified user features and functions such as simple zoom and layer selections, with the advent of 2.0, where exchange of information and user-generated content grew significantly, the capabilities of Web GIS grew from primarily spatial data browsing to now include functionality in four different areas. To begin, the most often offered capabilities in a Web GIS are geo visualization and data querying, with the results being shown to the user as maps. Each place on a map is shown through data visualization as attributes that explain the features of the location in question, as is the case with spatial information. Second, a Web GIS can acquire geographic data as a professional or commercial operation as well as through volunteered data collecting. There are several instances of this function in action, with Open

Street Map being likely the most successful example of volunteered geographic information (VGI) initiatives. The distribution of geographic information is the third function of Web GIS, which extends the first two purposes by allowing secondary users of a Web GIS dataset to engage with it further or by allowing data to be downloaded and shared. The fourth role of a Web GIS is to disseminate geographic information using one of these techniques after doing geospatial analysis.

The classification of web by various authors: Web 1.0, Web 2.0, Web 3.0, and Web 4.0 (and beyond) in terms of its evolution (Fu & Sun, 2011; Choudhury, 2014). The given table below shows how to find them. The initial age of reading webs, centered on retrieving information, was defined by Web 1.0. According to Berners lee-concept, they employed fundamental HTTP and HTML technologies to disseminate online linked information (1998).

Web 2.0 (O'Reilly, 2005) is a read-write web in which users can collaborate on content production, modification, and sharing.

Web 3.0, often known as the semantic web, focuses on allowing machines to understand semantically structured material, giving users more context and relevant information (Berners-Lee et. al., 2001; Palmer, 2001).

Web 4.0, according to Choudbury (2014), is an ultra-intelligent agent that makes judgments as powerful as the brain to enable community cooperation. Although several attempts have been made to define Web 5.0, it remains mostly undefined and speculative.

Peng & Tsou (2003), in one of the first textbooks on the Internet and GIS, present online advancements for GIS from a technical perspective (Table 5.2).They associate static map publishing with the use of HTML and static pictures for maps in the beginning. Static web mapping complements HTML Forms and CGI advances. Then there's interactive web mapping, which uses Dynamic HTML, scripts, plug-ins, ActiveX, Java applets, and servlets to allow for limited user interaction.

Table: 5.2

Various authors	**Peng & Tsou, 2003**	**Tsou,2005**	**Plewe,2007, Tsou,2011**	**Hall& Tiropanis, 2012**
Web 1.0	Static map publishing.	GIS awareness.	First generation -static.	Web of document.
Web 2.0	Static web mapping.	Free satellite imagery.	Second generation - dynamic.	Web of people.
Web 3.0	Interactive web mapping.	Disaster response.	Third generation -interactive.	web of data and social network.
Web 4.0	Distributed GIS service.	User response times-AJAX, image tilting.	Forth generation virtual earth globe.	
Web 5.0		Virtual globes and online mapping services.	Fifth generation-cloud computing, RIAs, crowd -sourcing.	

For decades, most digital geographic data was limited to usage on desktop PCs or in-house computer systems, and it was challenging to exchange with other enterprises. GIS analysts used their own workstation computers to access data, which were frequently connected to a common file server in the office. Viewing or manipulating the data needed specialized software, effectively limiting the number of people who could benefit from it. With the widespread adoption of the Internet as we already know during mid-1990s, people began to consider how maps and other geospatial information could be exchanged among computers, both within and outside the business. The first stage was to use HTML sites to display static images of maps; however, people quickly grasped the value of interactive maps. The proliferation of the internet and the development of the web give two major features that can greatly assist geoscientists. For instance, the Web provides for visual data by the help of which other clients can access these updates because the maps and charts were published on the Internet, which helps to speed up the review process. Second, due of the Internet's near-ubiquitous nature, geographical data may be readily accessed. Clients can access it from practically anywhere. Both of these depict how geoscientists will conduct their research in the nearby future. The combination of accessible data availability and visual display overcomes some of the most common challenges in geosciences evaluations.

5.4 SUMMARY

The development and advancement of this dynamic technology brings a revolution in the world. Internet has ability to change the texture of development occurring in the entire world, thanks to web GIS that we can shift from a system of record to a system of engagement, which allows us to do anything from self-service mapping to making smarter judgments, now the more advanced types of software as well as hardware were developing day by day and they very well reefing an enhancing our requirement. It makes GIS more widely available, inexpensive, and accessible. It emphasizes the importance of GIS and the importance of GIS specialists. It has extended GIS power from local servers to the cloud, putting online maps and geospatial intelligence in the hands of millions and billions. As these developments were occurring, the GIS is also getting in more advance phase. GIS has begun to adapt to these developments in software and technology, as well as the expansion of the internet. Devices have gotten smaller, lighter as well as full of power, and can now be utilized on the move. As a result, all GIS development efforts should be carried out in a manner that keeps pace with technological advancements and the ever-increasing needs and wants of GIS users. GIS applications, particularly Web-GIS, must be deployable, scalable, and accessible. Remote material must be integrated with local content. Web services allow us to connect to a wealth of information, ranging from our own organization to the Internet of Things (IoT), Big Data, and many other sources. In fact, Geographic information systems (GIS) have progressed from being desktop-based software to being web-based and available as web services. Many individuals today use GIS as a result of online services delivered through the internet. These operations are continuing to increase, and will definitely continue to provide the broadest range of spatial information services and GIS capability long into the future.

5.5 GLOSSARY

Binary Image - Binary images are images whose pixels have only two possible intensity values.

DOM - Document Object Model

GIS - Geographic information system.

Geographic Data - A Geographic Information System (GIS) is a computer system that analyzes and displays geographically referenced information.

Geospatial - Geospatial data is information that describes objects, events or other features with a location on or near the surface of the earth.

Geospatial Techniques - Geospatial technologies are a term used to describe the range of modern tools contributing to the geographic mapping and analysis of the Earth and human societies.

HTTP - Hypertext Transfer Protocol

HTML - Hyper Text Markup Language

JSON - Java script object notation.

Spatial Data - Spatial data comprise the relative geographic information about the earth and its features.

Server - A server is a piece of computer hardware or software (computer program) that provides functionality for other programs or devices.

Thematic - A thematic map is a type of map that portrays the geographic pattern of a particular subject matter (theme) in a geographic area**.**

Topography - the physical characteristics of an area of land, especially the position of its rivers, mountains, etc.

URL - Uniform Resource Locator

Variables - A variable is any entity that can take on different values.

Web GIS - Web GIS is an advanced form of Geospatial Information System (most commonly known as Geographic Information System) available on the web platform.

Web pages - A web page is a hypertext document provided by a website and displayed to a user in a web browser.

WMS - Web Map Service

XML - Extensible Markup Language

5.6 ANSWER TO CHECK YOUR PROGRESS

Q.1. Define web GIS and its objectives.

Q.2. Discuss the types and elements of web mapping.

Q.3. Web GIS technology is limited to some sort of people having smart and developed technologies do you agree? Elaborate.

Q.4. Discuss the development processes of webs GIS and its components.

Q.5. Does the expansion of GIS limits the growth of web GIS or both are supplementary to each other? Discuss.

5.7 REFERENCES

- Alesheikh, H. Helali and H. A. Behroz, "Web GIS: Technologies and its Applications," in Symposium on Geospatial Theory, Processing and Applications, canada,2002.
- R. Cavaco, R. Sequeira, M. Araújo and M. Calejo, "Rapid GIS Development: a model-based approach focused on interoperability," in 13th AGILE International Conference on Geographic Information Science, Guimarães, Portugal, 2010.
- Li S, Dragicevic S, Veenendaal B (2011). Advances in Web-based GIS Mapping Services and Applications. Taylor & Francis Group, London, UK.
- Nayak, S (2000) GIS Data Dissemination: A New Approach through WEB Technology. Rolta India Ltd.
- Asbury S (2000) GIS Industry Outlook: The Birth of a New Millennium. GeoWorld, December 1999, Vol. 12, No. 12
- M. J. Escalona, A. Torres-Zenteno, J. Gutierrez, E. Martins, R. d. S. Torres and M. C. C. Baranauskas,"A Development Process for Web Geographic Information System-A Case Study," Requirements Engineering, 2008.
- R. Vidgen, "Constructing a Web Information System Development Methodology," Blackwell Science Ltd, p. 247–261, 2002.
- Pascaul M, Alves E, De Almeida T, Sand de França G, Roig H (2012). An Architecture for Geographic Information Systems on the Web.
- Web GIS, GEO Processing: The Fourth International Conference on Advanced Geographic Information Systems, Applications, and Services.

5.8 TERMINAL QUESTIONS

Q.1. Elaborate the initial history of GIS and its features. How the GIS technology helping the human kind.

Q.2. What is the application of web GIS?

Q.3. GIS and web GIS is a greater tool for new millennium. Elaborate.

Q.4. Write down the summary of web GIS.

Q.5. Do you think web GIS is compatible with the modern world technology? Define with suitable examples.

UNIT 6 : SURVEY OF SOFTWARE AND HARDWARE

6.1 OBJECTIVES
6.2 INTRODUCTION
6.3 SURVEY OF SOFTWARE AND HARDWARE
6.4 SUMMARY
6.5 GLOSSARY
6.6 ANSWER TO CHECK YOUR PROGRESS
6.7 REFERENCES
6.8 TERMINAL QUESTIONS

6.1 OBJECTIVES

Since we know that GIS (geographic information system) combines the power of a map with the power of a database to allow people to produce, manage, and analyse data, particularly data about location. For this purpose, use of certain hardware and software are mentioned along with the survey and analytical techniques used to choose a specific hardware or software over another. A survey is gathering and enquiry of information carried out by observations, analysis and measurements.

The study's goals are to find out which institutions have implemented GIS and its survey of software and hardware and which institutions are planning to do so, how institutions have put it into practise GIS concentrating on hardware/software, manpower, service levels, training, financial support, and use.

6.2 INTRODUCTION

Hardware and software are required in order to develop and use GIS.

The software refers to the operating systems, and application programs such as GIS programs, word processors and other third-party programs.

The choice of suitable hardware and software is needed in order to have an efficient GIS that is capable of running the GIS application in terms of suitable response time, adequate results, and such like. The swiss army knives of GIS are desktop GIS platforms. They're used to create, modify, visualise, manage, and analyse geographic information. Desktop GIS software is a complicated programme. Most users can quickly learn the basics of products like Photoshop or Microsoft Excel, but becoming a great master can take years.

A problem/solution approach is the most effective way to learn a desktop GIS system. Rather than trying to learn everything it can do, it's advisable to start with the aspects that are relevant to the issues you're trying to solve.

Growing advancements in VLSI technology have resulted in a rise in the complexity of today's hardware systems. Due to the related time delay and loss of production, late detection of design mistakes often results in greater costs. As a result, it's critical that hardware designs are error-free. Formal verification is becoming a more significant tool for ensuring that hardware designs are valid. In this essay, we review the existing research in this field, with a focus on more recent trends. Based on the forms of the specification, the implementation, and the proof method, we provide a classification system for the various techniques. This framework makes it easier to see the connections and interactions between seemingly disparate approaches.

Hardware is devices or instruments required to run and display the results produced by software.

- Refers to the computer machinery.
- This includes computers ranging from a standalone microcomputer to a large mainframe supporting many users; the devices for handling machinery input and output. Graphics card, printer, hard drive, motherboard, and CPU – which practically every GIS will need – are

some of the most fundamental hardware components used in GIS. Scanners, video cards, transparency adapters, and other advanced geographic information systems may be used.

Software - GIS software, on the other hand, is the actual application or programme that is run on the hardware to perform geographical analysis. Software refers to the programs that tell the computer what to do, and the data the program will use. It also includes a database management system that is used to process and simply large chunks of data that are too extensive to check and edit manually.

- GIS software provides the tools needed to conduct geographic analysis.
- GIS software includes a database management system that is used to process and simply large chunks of data that are too large to manually check and edit.
- The database tools enable querying, editing, and altering of data through the use of scripts that execute commands across large amounts of data in a matter of seconds
  GIS software is commonly used to manage two kinds of data:

  Raster data: is a type of data structure that is based on the tessellation of the two-dimensional plane into cells.
  Vector data: A mathematically discrete data structure that stores data in the graphical x and y format, such as lines, points, and areas.

There are also enormous number of GIS software available and one can choose from the following given list. The requirement of the software depends on the needs of the organization, functionality desired, and the availability of money, and the time period. Comparisons of cost and benefits should be made (both of which are changing rapidly) before making a final decision

6.3 SURVEY SOFTWARE AND HARDWARE

Hardware requirement of GIS

There is a wide range of hardware used in geographic information system some of which are mentioned below:

Theodolite

Digital theodolites consist of a telescope that is mounted on a base, as well as an electronic readout screen that is used to display horizontal and vertical angles. Digital theodolites are convenient because the digital readouts take the place of traditional graduated circles and this creates more accurate readings.

Digital and non-digital theodolites are the two types of theodolites. Theodolites that aren't computerised are becoming increasingly rare. A telescope mounted on a base, as well as an electronic readout screen, is used to display horizontal and vertical angles in digital theodolites. Digital theodolites are useful because they replace traditional graduated circles with digital readouts, resulting in more precise readings.

Fig 1
Source:

https://www.johnsonlevel.com/Content/files/ProductImages/0000000395.png

What Is a Theodolite and How Does It Work?

To find vertical and horizontal angles in surveying, a theodolite uses optical plummets (or plumb bobs), a spirit (bubble level), and graduated circles. The theodolite is put as close to absolutely vertical above the survey location as possible using an optical plummet. The device's internal spirit level ensures that it is level with the horizon. The user can scan for angles using the graduated circles, one vertical and one horizontal.

How to use a Theodolite

1. Mark the point at which the theodolite will be set up with a surveyor's nail or a stake. This point is the basis for measuring angles and distances.
2. Set up the tripod. Make sure the height of the tripod allows the instrument (the theodolite) to be eye-level. The centred hole of the mounting plate should be over the nail or stake.
3. Drive the tripod legs into the ground using the brackets on the sides of each leg.
4. Mount the theodolite by placing it atop the tripod, and screw it in place with the mounting knob.
5. Measure the height between the ground and the instrument. This will be used a reference to other stations.
6. Level the theodolite by adjusting the tripod legs and using the bulls-eye level. You can make slight tunings with the levelling knobs to get it just right.
7. Adjust the small sight (the vertical plummet) found on the bottom of the theodolite. The vertical plummet allows you to do ensure the instrument remains over the nail or stake. Adjust the plummet using the knobs on the bottom.
8. Aim the crosshairs in the main scope at the point to be measured. Use the locking knobs on the side of the theodolite to keep it aimed on the point. Record the horizontal and vertical angles using the viewing scope found on the theodolite's side.

The benefits of using a theodolite are numerous.

When compared to other levelling instruments, theodolites have a number of advantages:

- Increased precision.

- Internal optical system for magnification.
- Readings on a computer.
- Horizontal circles can be zeroed or set to any value in an instant.
- Readings on the horizontal circle can be taken to the left or right of zero.

Reading circles with theodolites is significantly more accurate than with other instruments because of an inside optical mechanism. These measurements may also be made significantly faster using the theodolite because it allows you to take fewer repeat readings. The use of optical instruments with theodolites has advantages over conventional layout tools. They provide more accurate measurements, are unaffected by wind or other weather conditions, and may be used on both level and inclined ground.

GPR (ground penetrating radar)

Ground-penetrating radar (GPR) is a geophysical technology for imaging the subsurface that uses radar pulses. It's a non-intrusive approach of surveying the subsurface to look for concrete, asphalt, metals, pipes, cables, or brickwork. [1] This non-destructive method detects reflected signals from subsurface structures using electromagnetic radiation in the microwave band (UHF/VHF frequencies) of the radio spectrum. In addition to rock, soil, ice, fresh water, pavements, and constructions, GPR can be used in a number of media. GPR can be used to detect subsurface objects, changes in material characteristics, and voids and fissures in the right conditions. [2]

High-frequency (generally polarised) radio waves are used in GPR, which typically range from 10 MHz to 2.6 GHz. Electromagnetic radiation is emitted into the ground by a GPR transmitter and antenna.

The energy may be reflected, refracted, or scattered back to the surface when it contacts a buried item or a border between materials with varying permittivity. The differences in the return signal can then be recorded using a receiving antenna. The principles are similar to those used in seismology, with the exception that GPR methods use electromagnetic energy rather than acoustic energy, and energy can be reflected at boundaries where subsurface electrical properties change rather than subsurface mechanical properties, as with seismic energy.

Applications

GPR offers a wide range of applications in a variety of fields. It is used to examine bedrock, soils, groundwater, and ice in the Earth sciences. It can help you uncover gold nuggets and diamonds in alluvial gravel beds by locating natural traps in buried stream beds that have the capacity to collect larger particles. The Chinese lunar rover Yutu carries a ground penetrating radar (GPR) on its underside to explore the Moon's soil and crust.

Non-destructive testing (NDT) of structures and pavements, finding buried structures and utility lines, and researching soils and bedrock are all examples of engineering applications. GPR is used to designate landfills, pollution plumes, and other remediation sites in environmental remediation, and it is also used to map archaeological features and cemeteries in archaeology. In law enforcement, GPR is used to locate hidden graves and buried evidence. Mine detection, unexploded ordnance detection, and tunnel detection are all examples of military applications.

Fig 2

Source:

https://cpasurvey.com/wp-content/uploads/2019/05/Ground-Penetrating Radar.jpg

Bathymetric survey

Bathymetric surveys allow us to measure the depth of a water body as well as map the underwater features of a water body. Multiple methods can be used for bathymetric surveys including multi-beam and single-beam surveys, ADCPs, sub-bottom profilers, and the Eco mapper Autonomous Underwater Vehicle. We use bathymetric surveys for many different types of research including flood inundation, contour of streams and reservoirs, leakage, scour and stabilization, water-quality studies, dam removal, biological and spill, and storage and fill in reservoirs and ponds. Multibeam echo sounding: A boat-mounted multibeam echo sounder puts out a wide array of beams across a "swath" of the waterbody bottom. The data is collected and processed as the beams bounce back from the waterbody floor. During the survey, the processed data can be watched in real time on the boat. Larger bodies of water usually require multi-beam surveying.

Surveying using a single beam: Single-beam bathymetry examines the water depth immediately beneath the boat rather than sending out a large number of beams. For smaller bodies of water, single-beam surveys are commonly utilised.

ADCPs (Acoustic Doppler Current Profilers) are used by the USGS to measure streamflow. ADCPs use sound waves to detect water velocity, which are reflected off silt and other things in the water. Bathymetric mapping can be done with the data collected from ADCPs.

Sub-bottom profilers are most typically used to examine the layers of sediment and rocks beneath the water's surface. A sound wave is sent to the water's surface by a transducer. This sound wave is capable of penetrating the water's surface. The information obtained from sound waves can be mapped to reveal the layers beneath the water's surface.

Eco mapper Autonomous Underwater Vehicle: The Eco mapper can collect detailed bathymetric data, down to one-foot contours, in areas where boats can't go. Side-scan sonar and a Doppler velocity log are used by the Eco mapper.

Fig 3

Source: https://professional-surveys.com/wp-content/uploads/2017/02/Residential.jpg

Total station

The total station is used to measure slant lengths, horizontal and vertical angles, and elevations in topographic and geodetic operations, as well as tachometric surveys and solving application geodetic tasks. The results of the measurements can be saved in the internal memory and sent to a computer interface.

The main qualities are unrivalled measurement range, speed, and accuracy. Total stations are designed with the user's maximum convenience in mind. Electronic tachometers with high efficiency are used to make decisions.

For the sole purpose of industrial difficulties, it has a large readership.

The coordinates (X, Y, and Z or northing, easting, and elevation) of surveyed points relative to the total station position are computed using trigonometry and triangulation.

The total station's data can be downloaded to a computer, where application software can be used to compute results and generate a map of the surveyed area.

In modern surveying, a total station is an electronic/optical instrument. Archaeologists use it to document excavations, while police, crime scene investigators, private accident reconstructionist, and insurance firms use it to measure scenes. The complete station consists of an electronic theodolite (transit) with an electronic distance metre (EDM), as well as internal and/or external data storage.

Control points were a basic necessity for the compilation of these maps, and they were used in almost every survey.

There are a variety of ways for obtaining these control points, including traverse, triangulation, and others.

Fig 4
Source: https://geooprema.com/wp-content/uploads/2020/02/Total-Station_Zoom50_Image5-LR.jpg

APPLICATION

In mine surveying, total stations are the most used survey device.

As the drifts of an underground mine are driven, a total station is utilised to record the absolute location of the tunnel walls, ceilings (backs), and floors. The captured data is then imported into a CAD computer and compared to the tunnel's designed architecture.

Control stations are placed at regular intervals by the survey party. These are little steel plugs that are inserted in pairs in holes drilled into the walls or the back of the house. Two plugs are inserted in opposite walls for wall stations, making a line perpendicular to the drift. Two plugs are inserted in the back of the station, forming a line parallel to the drift.

By applying measurements to the plugs by intersection and resection, a set of plugs can be utilised to find a total station placed up in a drift or tunnel.

Construction of mechanical and electrical systems- For most types of construction planning, total stations have established the gold standard. They're most commonly employed in the X and Y axes to lay out the positions of underground utility penetrations into the foundation, as well as between floors of a structure and roofing penetrations. Because increasingly commercial and industrial construction projects are based on building information modelling (BIM), the coordinates for practically every pipe, conduit, duct, and hanger support are now available in digital form.

The application of a virtual model to a physical construction could reduce labour expenses associated with relocating poorly measured systems, as well as time spent setting out these systems while a full-fledged building project is underway.

Total stations are also used by meteorologists to track weather balloons and determine upper-level winds. The change in azimuth and elevation measurements produced by the total station as it tracks the weather balloon over time are used to compute the wind speed and direction at different altitudes when the average ascent rate of the weather balloon is known or assumed. The total station is also used to measure the height of cloud layers by tracking ceiling balloons. Upper-level wind data is frequently used for weather forecasting in aviation and rocket launches.

GPS (global positioning system)

The Global Positioning System (GPS) is a navigation system for air, sea, and land transport that uses satellites, a receiver, and algorithms to synchronise location, velocity, and time data. The satellite system consists of 24 satellites in six Earth-centred orbital planes, each with four satellites, orbiting about 13,000 miles (20,000 kilometres) above Earth and travelling at 8,700 miles per hour (14,000 kilometres per hour).

While just three satellites are required to establish a location on the earth's surface, a fourth satellite is frequently utilised to confirm the data from the other three. The fourth satellite takes us into the third dimension, allowing us to calculate a device's altitude.

The three parts of GPS are as follows:

- Satellites circling the Earth deliver signals to users based on their geographical location and time of day.
- The Control Segment consists of Earth-based monitor stations, master control stations, and a ground antenna. Tracking and operating satellites in space, as well as monitoring signals, are all part of the control tasks. Monitoring stations can be found on nearly every continent, including North and South America, Africa, Europe, Asia, and Australia.
- GPS receivers and transmitters, such as watches, smartphones, and telematic devices, are examples of user equipment.

Fig 5

Source:https://images.world-of waterfalls.com /GPS_Handheld _iPhone_Fenix_002_ _07032020.jpg

What is GPS technology and how does it work?

Trilateration is a technology used by GPS to work. Trilateration receives signals from satellites to output position information and is used to determine location, velocity, and elevation. It's frequently confused with triangulation, which is used to measure angles rather than distances.

Satellites circling the earth send signals that a GPS device on or near the earth's surface reads and interprets. A GPS gadget must be able to read signals from at least four satellites in order to determine location.

Each satellite in the network does two daily orbits around the planet, sending a unique signal, orbital characteristics, and time. A GPS gadget can read signals from six or more satellites at any given time.

A single satellite sends out a microwave signal that a GPS device picks up and uses to compute the distance between the GPS device and the satellite. A single satellite cannot provide much location information because a GPS gadget only provides information about the distance from a satellite. Because satellites do not transmit information about angles, a GPS device's location might be anywhere on a sphere's surface area.

When a satellite transmits a signal, it forms a circle with a radius equal to the distance between the GPS device and the satellite. When a second satellite is added, a second circle is created, and the location is narrowed down to one of two spots where the circles overlap. The gadget's location can eventually be identified with the help of a third satellite, as the device is at the intersection of all three circles.

There are five primary applications for GPS:

- Determining a position is known as location.
- Navigation is the process of moving from one place to another.
- Monitoring the movement of an object or a person is known as tracking.
- Mapping is the process of creating world maps.
- Timing – Allowing exact time measurements to be taken.

DGPS (differential global positioning system)

A more advanced kind of GPS navigation that delivers more precise positioning than basic GPS. The error correction supplied from a GPS receiver situated at a known location is used by DGPS. This receiver, referred to as a reference station, estimates the error in satellite range data and provides corrections for use by mobile GPS receivers in the same area. The DGPS system removes all measurement inaccuracies in satellite ranges, allowing for extremely accurate position calculations.

The fundamental distinction between GPS and DGPS is that, GPS instruments range is global. on the other hand, DGPS's instruments range is local.

Difference between GPS and DGPS

• GPS stands for Global Positioning System.	• DGPS stands for Differential Global Positioning System.
• In GPS, frequency varies from 1.1 GHz to 1.5 GHz.	• While in DGPS, frequency varies as per agency requirements.
• GPS uses **WGS84**, that is used for time coordinate system.	• While in DGPS, local coordinate system is used.
• The cost of GPS is lower than DGPS.	• While the cost of DGPS is costlier than GPS
• GPS's instruments range is global.	• While DGPS's instruments range is local.
• In GPS, only one receiver is taken place.	• While in DGPS, two receivers are taken place.

Software requirements of GIS

There are enormous number of GIS software available and one can choose from the following given list. The requirement of the software depends on the needs of the organization, functionality desired, and the availability of money, and the time period. Comparisons of cost and benefits should be made (both of which are changing rapidly) before making a final decision.

Most widely used software applications are: -

LiDAR

Light detection and ranging, or LiDAR, is a popular remote sensing technique for determining an object's exact distance on the earth's surface. Despite the fact that it was originally employed in the 1960s when laser scanners were put on planes, LiDAR did not gain widespread acceptance until twenty years later. Only after the introduction of GPS in the 1980s did it become a common tool for determining precise geographical measures. We should learn more about LiDAR mapping technology and how it works since that its application has expanded to include a wide range of fields.

LiDAR employs a pulsed laser to compute an object's varying distances from the earth's surface, according to the American Geoscience Institute. When these light pulses are combined with the data collected by the aerial system, precise 3D information about the earth's surface and the target object is generated.

A LiDAR apparatus is made up of three main parts: a scanner, a laser, and a GPS receiver. The photodetector and optics are two more important components in data collecting and analysis. The majority of government and corporate companies collect LiDAR data using helicopters, drones, and aeroplanes.

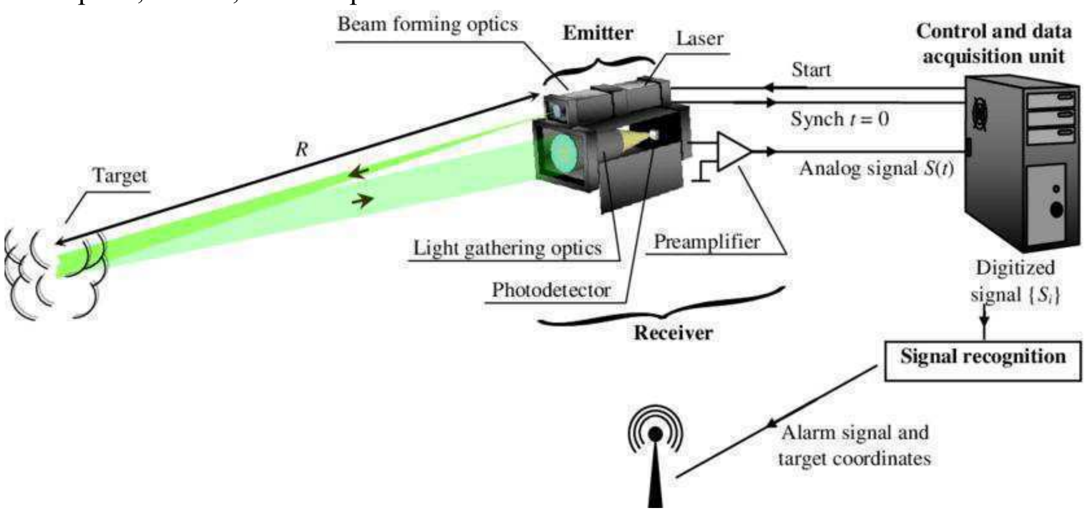


Fig 6

Source:https://www.researchgate.net/profile/AndreiUtkin/publication/236681742/figure/fig3/AS:667856744247297@1536240864054/Lidar-equipment-and-detection-principles.png

Working of LiDAR

LiDAR works on the basis of a basic principle: shoot laser light at an object on the ground and calculate how long it takes for the light to return to the LiDAR source. Given the speed of light (about 186,000 miles per second), the process of measuring the exact distance using LiDAR looks to be extremely quick. It is, however, extremely technical. Analysts use the following calculation to determine the exact distance of an object:

(Speed of Light x Time of Flight)/ 2 is the object's distance.

LiDAR can be used to achieve a variety of developmental goals, including:

Oceanography

LiDAR technology is used by authorities when they need to know the exact depth of the ocean's surface to find any object in the event of a maritime mishap or for research purposes. LiDAR is also used to calculate phytoplankton fluorescence in addition to locating things.

Terrain Model or Digital Elevation

When it comes to building highways, major structures, and bridges, terrain elevations are critical. Because LiDAR technology uses x, y, and z coordinates, it's relatively simple to create a 3D depiction of elevations so that interested parties can more readily draw the necessary conclusions.

IDRISI- is a remote sensing and integrated geographic information system (GIS) developed by Clark Labs at Clark University for the analysis and display of digital geospatial data. IDRISI is a PC grid-based system that provides tools for researchers and scientists who are analysing earth system dynamics in order to make effective and responsible decisions for environmental management, sustainable resource development, and equitable resource allocation.

Features: -

It has the ability to query raster and vector map layers and report basic statistics as well as profiles over space and time, histograms, and tabulation of area and parameter.

A full site of mathematical and modelling tools for deriving new data layers as a function of existing layers.

A method of allocating land that considers contiguity and compactness.

An extensive collection of criteria aggregation processes based on Weighted Linear Combination and Ordered Weighted Averaging, as well as a consensus-seeking procedure for weighting criteria, fuzzy standardisation, and a consensus-seeking procedure for weighting criteria.

The evaluation of decision risk as a result of propagated error, the calculation and aggregation of Fuzzy Sets, and the aggregation of indirect evidence to support a weight-of-evidence conclusion using both Bayesian and Dempster-Shafer methodologies are all examples of error propagation tools.

An implementation of spatial prior probabilities for Maximum Likelihood classification, as well as a soft reclassification technique to map the chance of a site being above or below a threshold (such as sea level rise).

ESRI- includes ArcGIS Desktop which is a foundational piece for GIS professionals to create, analyse, manage, and share geographic information in order for decision-makers to make intelligent, informed decisions. It enables you to create maps, conduct spatial analysis, and manage data.

Features- it captures data and has ability to input geographic data into GIS.

It stores vast amount of data in a variety of ways such as individual files, databases, geodatabases or the cloud.

It manipulates data and ability to make edits to data as the landscape changes.

It assigns locations to significant addresses so they may be placed as points on the map.

Microsoft MapPoint is a Microsoft software tool and service that allowed users to browse, edit, and combine maps. The software and technologies are designed to make it easier to visualise and analyse geographical data, whether it's included data or bespoke data. Some alternatives of map point are as follows-

MAPCITE is a platform that allows users to dynamically display, analyse, filter, animate and consume data, as well as produce new and unique location- specific data, resulting in new insights and perspectives on your organisation.

The value of realtime data from mobile and smart devices has skyrocketed, and the Mapcite p latform is perfectly positioned to help you take advantage of these game-changing prospects.

Features: -

This geofencing technology ensures that actual 'Location Intelligence' is delivered at the time of need. Users may use Mapcite's easy to use location data platform to bring data to life, making it more useful, producible, and profitable. Mapcite is a web-based mapping tool that is completely free to use (for a limited amount of data).

MAPTITUDE comes close to replacing MapPoint, Streets & Trips, and AutoRoute on the desktop. When migrating from MapPoint, Maptitude imports Microsoft map file formats such as PTM, EST, and AXE, saving you time. Maptitude includes updated versions of all MapPoint capabilities, as well as the tools, maps, and data you'll need to study and comprehend how geography influences your business.
Features:
Unlimited routing and directions, territory construction, address and post code matching/geocoding, ring and drive-time analysis, demographic exporting, report generation, custom area modification, measures and distances, and more are all supported by Maptitude. Maptitude is the most user-friendly, powerful, and affordable full-featured mapping software on the market. It gives developers all the programming tools they need to add GIS functionality, mapping, and geographic analysis to Windows desktop programmes built in any.NET language or Python.

iDesktop is one of the most rapidly evolving software platforms in recent decades. Among the various GIS software, it has one of the best 3D GIS technologies. It was created by a business dedicated to GIS software innovation mixed with conventional IT technology (AI GIS, Big Data GIS, etc.) The software is also very user-friendly, with a wide range of functions, making it ideal for beginners who want to learn more about GIS. They currently hold the third-largest market share in the global GIS industry.

QGIS is a geographic information system (GIS) programme that allows users to analyse and update spatial data as well as create and export graphical maps. [3] Vector data is stored as point, line, or polygon features in QGIS; raster data is stored as point, line, or polygon features. The software can georeferenced images and supports multiple raster image formats.

Shapefiles, coverages, personal geodatabases, dxf, MapInfo, Post GIS, and more formats are all supported by QGIS.

Web services, such as the Web Map Service and the Web Feature Service, are also supported, allowing data from external sources to be used.

Other open-source GIS software that QGIS connects with includes Post GIS, GRASS GIS, and Map Server.

Plugins written in Python or C++ increase the functionality of QGIS. Plugins can use the Google Geocoding API to geocode, perform geoprocessing operations comparable to those present in ArcGIS' standard tools, and connect to databases such as PostgreSQL/PostGIS, SpatiaLite, and MySQL.

6.4 SUMMARY

GIS development and use necessitate the use of both hardware and software. Operating systems and application programmes such as GIS applications, word processors, and other third-party programmes are examples of software. The selection of appropriate hardware and software is required in order to have an efficient GIS capable of executing the GIS application in terms of appropriate response time, adequate results, and so on. Desktop GIS platforms are the swiss army knives of GIS. They are used to produce, change, visualise, manage, and analyse geographic data. Desktop GIS software is a complex programme. Most users can rapidly grasp the fundamentals of tools like Photoshop or Microsoft Excel, but becoming a great master can take years.

6.5 GLOSSARY

Analysis The process of identifying a question or issue to be addressed, modelling the issue, investigating model results, interpreting the results, and possibly making a recommendation.

Annotation 1. Descriptive text used to label coverage features. It is used for display, not for analysis. 2. One of the feature classes in a coverage used to label other features. information stored for annotation includes a text string, the location at which it is displayed, and a text symbol (color, font, size, etc.) for display.

ArcGIS A comprehensive desktop GIS software package developed by ESR.

ArcMap Editing and map making module of ArcGIS.

Coordinate A set of numeric quantities that describe the location of a point in a geographic reference system. A coordinate pair describes the location of a point or node in two dimensions (usually x-y), and a coordinate triplet describes a point in three dimensions (x-y-z). A series of points (two or more) is used to describe lines and the edges of polygons or areas. Coordinates represent locations on the Earth's surface relative to other locations.

Coordinate system A reference system used to measure horizontal and vertical distances on a planimetric map. A coordinate system is usually defined by a map projection, a spheroid of reference, a datum, one or more standard parallels, a central meridian, and possible shifts in the x- and y-directions to locate x, y positions of point, line, and area features. A common coordinate system is used to spatially register geographic data for the same area.

Digitize 1. To encode geographic features in digital form as x, y coordinates. 2. The process of using a digitizer to encode the locations of geographic features by converting their map positions to a series of x, y coordinates stored in computer files. Pushing a digitizer button records an x, y coordinate. A digitized line is created by recording a series of x, y coordinates.

Digitizer 1. A device that consists of a table and a cursor with crosshairs and keys used to digitize geographic features. 2. Title of the person who uses a digitizing device.

Geodatabasean object-based GIS data model developed by ESRI for ArcGIS that stores each feature as rows in a table. Personal geodatabases store data in a Microsoft Access .mdb file. Corporate geodatabases store data in a DBMS such as SQLserver or Oracle. This data structure supports rules-based topology and allows the user to assign behaviour to data.

Geodetic datum A mathematical model designed to fit part or all of the geoid (the physical earth's surface). Defined by the relationship between an ellipsoid and a point on the topographic surface established as the origin of a datum. World geodetic datums are typically defined by the size and shape of the ellipsoid and the location of the center of the ellipsoid with respect to the center of the earth.

Geographic data The locations and descriptions of geographic features. The composite of spatial data and descriptive data.

Geographic database A collection of spatial data and related descriptive data organized for efficient storage and retrieval by many users.

GIS Geographic information system. An organized collection of computer hardware, software, geographic data, and personnel designed to efficiently capture, store, update, manipulate, analyse, and display all forms of geographically referenced information. Outputs include cartographic displays, spatial analysis, and tabular reports

Global positioning system (GPS) A system of satellites and receiving devices used to compute positions on the Earth.

Map projection A mathematical model that transforms the locations of features on the Earth's surface to locations on a two- dimensional surface. Because the Earth is three-dimensional, some method must be used to depict a map in two dimensions. Some projections preserve shape; others preserve accuracy of area, distance, or direction. See also coordinate system. Map projections project the Earth's surface onto a flat plane. However, any such representation distorts some parameter of the Earth's surface be it distance, area, shape, or direction.

Raster A cellular data structure composed of rows and columns for storing images. Each unit in the grid is assigned a value associating it with the corresponding attribute data. Selection of grid size forces a trade-off between data resolution (and detail) and system storage requirements. Data can be converted to vector data through the process of vectorization.

Remote sensing Acquiring information about an object without contacting it physically. Methods include aerial photography, radar, and satellite imaging.

6.6 ANSWER TO CHECK YOUR PROGRESS

1. Define GPS.
2. Define Geodetic Datum.
3. Define Map projection.
4. Define Remote Sensing.
5. Define Digitizer.
6. Define GPR.

6.7 REFERENCES

1. https://www.gissensing.com/gis/differences-between-gis-software-and-hardware/
2. http://wiki.gis.com/wiki/index.php/IDRISI
3. https://www.microsoft.com/en-us/maps/mappoint-alternatives
4. https://clarklabs.org/terrset/idrisi-gis/
5. https://www.johnsonlevel.com/News/TheodolitesAllAboutTheodo
6. https://www.geospatialworld.net/blogs/what-is-lidar-technology-and-how-does-it-work/
7. https://gisresources.com/total-station-and-its-applications-in-surveying/
8. https://www.geotab.com/blog/what-is-gps/
9. https://www.wartsila.com/encyclopedia/term/differential-global-positioning-system-(dgps)

6.8 TERMINAL QUESTIONS

1. What does GIS software do and what can it be used for?
2. What is the best GIS software product to use retail site selection?
3. Does ArcGIS mobile support differential GPS?
4. What are the instruments used in theodolite survey?
5. Are LIDAR and Airborne laser scanning the same thing?
6. What are the advantages and disadvantages of GPS over optical equipment like a total station?

UNIT 7 : WEB GIS APPLICATIONS & FUTURE OF WEB GIS

7.1 OBJECTIVES

Mapping/Visualization and query (attribute or spatial) are just a few of the goals of online GIS. Other goals include geospatial analysis, such as measurement, ideal driving path, routing, pollutant dispersion modelling, retail site selection, and so on. In future it will improve a variety of applications, including environmental quality management and the built environment, as well as land-use and transportation planning. As a result, better judgments will be made, efficiency will increase, and communication will improve. Cities will progressively make their data open geospatial services available (maps).

7.2 INTRODUCTION

Web GIS is a blend of web and geographic information systems. Users can now view spatial data in its appropriate format thanks to GIS software. As a result, spatial data interpretation has become much easier and more understandable. Unfortunately, not everyone has access to GIS, nor would he be able to devote the time required to effectively use it. Web GIS has evolved into a low-cost and simple method of spreading geographic data and processing tools. Many companies want to make maps and processing tools available to users without regard to time or location constraints. Many government agencies, as well as many homes, have adopted Internet technology. Because of the capacity to obtain information via the Internet, spatial data providers have begun to investigate Internet tools for spreading spatial data. To deliver a successful web GIS deployment, it is necessary to think of it as a process rather than a single step. In addition, the implementation should take into account the available technology as well as the application. People may now engage with GIS apps globally and obtain information virtually quickly thanks to the web, which has abolished the restrictions of distance in cyberspace. People may now interact with GIS applications from all over the world and acquire information almost instantly owing to the internet, which has removed the barriers of distance in cyberspace. Web GIS enables analytics to be applied to geographical or geographic data in ways that were previously not possible. Previously, spatial data had to be processed, edited, updated, and extracted in order to answer a certain set of queries. Now, the data is translated into web maps or services, which are mashed up with multiple layers into a web GIS, allowing it to answer any type of question. The data is ready to utilize and dynamically respond to inquiries, and it provides a turnkey solution. Data no longer needs to be handled individually for each question. Web GIS is far more adaptable, nimble, adjustable, and capable of facilitating effective workflow. Planning necessitates the integration of a vast amount of spatiotemporal data. Web GIS will enable the integration of data from different sources onto a single platform. This will indirectly lower the cost and effort spent on repetitive data generation.

In 1993, the Xeror Corporation Palo Alto Research Centre (PARC) created a web-based map viewer, which marked the beginning of Web GIS. Users might examine the map in a web browser and then click a link to a function. The browser initiates an HTTP request to the web server. When the server receives the request, it performs mapping operations, creates a new map, and provides it to the browser. The map is shown as a map image by the browser.

Early Web GIS examples:

- 1994, Canadian National Atlas Information Service

- 1995, Alexandria Digital Library
- 1995, National Geospatial Data Clearinghouse (by USGS)
- 1995, U.S Census TIGER Mapping Service
- 1996, Mapquest.

Web development and Internet expansion are two essential characteristics that geoscientists may tremendously benefit from. The Web, for starters, enables visual data engagement. Clients can build maps by setting up a Web Server. Other clients can view these updates because the maps and charts are published on the Internet, speeding up the review process. Second, geospatial data can be broadly available due to the Internet's near-ubiquity. It can be worked on by clients from practically anywhere. Both of these depict how geoscientists will conduct their research in the near future. The combination of accessible data availability and visual display overcomes some of the most common challenges in geoscience evaluations. There are some drawbacks to Web GIS. The main issue is speed, as GIS is heavily reliant on graphics. For users, Internet connection speeds can make extensive use of graphics intolerably slow. In the near future, it will not be able to match the intricacy of dedicated GIS tools like "ArcView & Arc Info" or "MapInfo." Web GIS, on the other hand, does not necessitate the same number of resources as these programmes. A site-wide GIS solution does not necessitate powerful computers, substantial training, or pricey site licenses.

7.3 WEB GIS APPLICATION AND FUTURE OF WEB GIS:

Web GIS application:

1. Query and visualization/mapping (attribute or spatial).
2. Collaborative geospatial data collection
3. Measurement, optimal driving route, routing, pollution dispersion modelling, and retail location selection are all part of geospatial analysis.
4. Web GIS as a novel product and business model Location: particular advertising and branding based on maps. Take Google Maps, for example.
5. The Geographic Information System (GIS) programmed is a fun and productive way to build and plan government activities like flood management, forest mapping, and natural catastrophes.
6. Web GIS technology is used by a geoscience research partnership.
7. Web GIS in everyday life: location-based services are available via mobile web, cellphones, and tablets (LBS). LBS comprises, among other things, services for determining a person's or object's location, such as finding a nearby ATM, restaurant, stores, and hotels. LBS also offer shipment and vehicle tracking services.

Future of Web GIS

Unlike desktop GIS, which is limited to a small group of GIS experts, web GIS can be utilized by anybody in an organization, including the general public. This diversified audience has a wide range of expectations. A few applications include street mapping, locating areas to tag personal images, tracing friends, and displaying Wi-Fi hotspots. Web browsers such as Internet Explorer, Google Chrome, Mozilla Firefox, Apple Safari, and others are the most common Web GIS clients. Web GIS that relies on HTML clients will often support different operating systems such as Microsoft Windows, Linux, and Apple Mac OS X because these web browsers are compliant with HTML and JavaScript standards. Web GIS allows you to share geographic data and maps with the rest of the globe. Anyone, anywhere in the world, can use their computers, desktops, or mobile devices to obtain geographic data. Almost all organizations enable particular network ports on their firewalls to allow HTTP requests and answers to pass through their local network, enhancing client accessibility. A classic desktop supporting GIS application software is typically used by just one person at a time, however a Web GIS can be used by dozens or hundreds of people at the same time. As a result, Web GIS offers far better speed and scalability than a single desktop GIS. End users can access the vast majority of internet content for free, and Web GIS is no exception. In most cases, using Web GIS does not require the purchase of software or payment. Web GIS can also help organizations cut costs by providing GIS services to individual users. Instead of purchasing and installing desktop GIS for each user, an organization can set up a single Web GIS that can be shared by a large number of people, whether they are at home, at work, or out in the field. Desktop GIS is designed for professionals who have had months of GIS training and experience. Web GIS is designed for a wide range of users, including those who are unfamiliar with GIS. Web GIS is created in such a way that it is as simple to navigate as a conventional website. Web GIS was created with simplicity, intuition, and convenience in mind, making it far easier to use than desktop GIS. To upgrade GIS on desktops to a new version, the update must be installed on each computer. For Web GIS, a single update is sufficient for all clients. Web GIS is a fantastic fit for presenting real-time information and making decisions because of its ease of upkeep.

Web GIS includes a lot of cool features that set it distinct from other school-based geographic software. Because web GIS is supplied through a web browser, the interfaces and data can be customized to meet the needs of individual learners, taking into account their developmental level, instructional objectives, and even the hardware technology used to access the maps or analysis. While most geographic desktop software may stream data to a user's computer, these apps do not use the web; rather, they use the Internet to send and receive data, and so are not included in web GIS. For instructional designers and educators, the technical distinctions in data protocols or transport techniques are largely arbitrary. However, the importance of a web browser's ability to display and analyze GIS data cannot be overstated.

The future of Internet GIS is promising and full of possibilities. There will be chances to cut operating costs by, for example, using a cable to access and pay for GIS capability on a "as required" basis rather than investing in the technology overkill of boxed desktop solutions. We'll find ways to boost productivity not only by putting user-friendly GIS interfaces in beginner users' hands, but also by putting it in their palm PCs. Smart businesses with rapid deployment tools will extend GIS-based services to their employees, who will use these applications to better predictably retain consumers.

7.4 SUMMARY

The term "Web GIS" refers to a combination of web and geographic information systems. Thanks to GIS software, users may now access spatial data in its proper format. As a result, interpreting spatial data has become considerably simpler and clearer. Unfortunately, not everyone has access to GIS, and he wouldn't be able to commit the time necessary to use it efficiently. Web GIS has evolved into a low-cost and straightforward technique of disseminating geographic information and processing tools. Many businesses want to make maps and processing tools available to users at any time and in any location. Many government offices, as well as many private residences, have embraced the Internet. Geographic data providers have begun to examine Internet technologies for sharing spatial data as a result of the ability to get information via the Internet. It is vital to think of a successful web GIS implementation as a process rather than a single step. Furthermore, the implementation should consider both the available technology and the application. People may now interact with GIS apps from anywhere in the world and access information almost instantly thanks to the internet, which has removed the barriers of distance in cyberspace. Because of the internet, which has abolished the limits of distance in cyberspace, people from all over the world can now engage with GIS applications and obtain information nearly quickly. Web GIS makes it feasible to apply analytics to geographical or geographic data in previously unimaginable ways. To answer a certain set of queries, spatial data had to be processed, changed, updated, and extracted previously. Now, the data is converted into online maps or services, which are then combined with numerous layers to create a webGIS that can answer any question. The data is ready to use and reply to enquiries dynamically, and it is a turnkey solution. Each question's data no longer needs to be handled separately. Web GIS is much more versatile, fluid, customizable, and capable of aiding efficient workflow than desktop GIS. The integration of a large amount of spatiotemporal data is required for planning. Web GIS will allow data from various sources to be combined on a single platform. This will reduce the cost and time spent on data generation that is repeated.

7.5 GLOSSARY

- **Affine transformation:** An affine transformation scales, rotates, skews, or translates points, polylines and polygons preserving points, straight lines and planes.
- **Arc:** Arcs are lines or polygon boundaries, represented as a series of vertices or coordinate points.
- **ArcCatalog:** ArcCatalog is an application in the ArcGIS suite for managing geographic data – similar to windows file explorer.
- **ArcGIS:** [software] ArcGIS is a GIS software package produced by the Environmental Systems Research Institute (Esri) which allows you to collect, store, manage, visualize, export, analyze and map geographic data.
- **ArcGIS Pro:** [software] ArcGIS Pro is Esri's latest GIS software with a ribbon-based user interface, project files and 64-bit processing.
- **Arc Globe:** Arc Globe is a 3D visualization and analysis environment as part of the Esri ArcGIS suite (3D analyst), specializing in global datasets and larger study areas.
- **Arc Scene:** [software] Arc Scene is a 3D feature and raster viewer part of the Esri ArcGIS suite of applications (3D analyst) specializing in small study area scenes.
- **Shaded Relief:** [remote sensing] A shaded relief map displays the brightness and shadows of terrain reflection given a sun angle and direction of sunlight.

- **Side lap:** Side lap (or side overlap) consists of the overlapping edge areas of photographs between adjacent flight lines.
- **Skeletonize:** Skeletonizing is the process of thinning a raster line to a single pixel width, typically for conversion to vector data format.
- **Skew:** Skewing distorts a feature by arranging its vertices in the x or y-direction.
- **Sliver:** A sliver is a small, spurious gap between polygons often considered as a topology error from imprecise digitization of features.
- **Slope:** Slope is the change in elevation or steepness with respect to change in location measured in degrees or percent slope.
- **Snapping Environment:** Snapping determines the distance that newly digitized points, lines and polygons will occupy the same location of existing features.
- **SSURGO:** Soil Survey Geographic Database (SSURGO): [organization] SSURGO is a fine-"scale, county-level, soil survey database prepared by the National Resource Conservation Service.
- **Spaghetti Data Model:** The spaghetti data model is a simple and dated GIS model where lines may cross without intersecting or topology without attributes.
- **Spatial Relationship:** [data structure] A spatial relationship links features geographically with a table by a unique identifier.
- **Spherical Coordinates:** coordinates is a coordinate system based on a sphere defined by two angles of rotation in orthogonal planes such as latitudes and longitudes in a geographic coordinate system.

7.6 ANSWER TO CHECK YOUR PROGRESS

1. Define slope.
2. Define shaded relief.
3. What do you mean by Spherical Coordinates?
4. Define Arc Catalog.

7.7 REFERENCES

- https://www.vizexperts.com/webgis
- Anon. (1997). Formal reply to the comments of F. Potucek and J. Stejskal. In *Chemical Engineering Science* (Vol. 52, Issue 4). https://doi.org/10.1016/S0009-2509(96)00411-3
- Kong, N., Zhang, T., & Stonebraker, I. (2015). Evaluation of web GIS functionality in academic libraries. *Applied Geography*, *60*, 288–293. https://doi.org/10.1016/j.apgeog.2014.11.017
- Luqun, L., Jian, L., & Yu, T. (2002). ACSG Table of contents Table des matières Authors index Index des auteurs The Study on Web GIS Architecture Based on JNLP. *Dbms*, 1–6.
- Parker, R. C. (n.d.). *Concepts and Principles of Point-Samplin*.

7.8 TERMINAL QUESTIONS

1. What are the applications of the web GIS?
2. Enlist some examples of early web gis applications.
3. What will the status of web GIS in the upcoming years?
4. How is web GIS is useful to a researcher and in the field of education?
5. How we can differentiate between the web and desktop GIS ?

BLOCK 3 : CROWD SOURCING

UNIT 8 : OVERVIEW OF CROWD SOURCING

8.1 OBJECTIVES

Crowd sourcing is the process of gathering information, views, or work from a large number of people via the Internet. Crowdsourcing allows businesses to save time and money while gaining access to people with a variety of skills and perspectives from all around the world.Crowdsourcing serves a variety of purposes that firms can take advantage of.

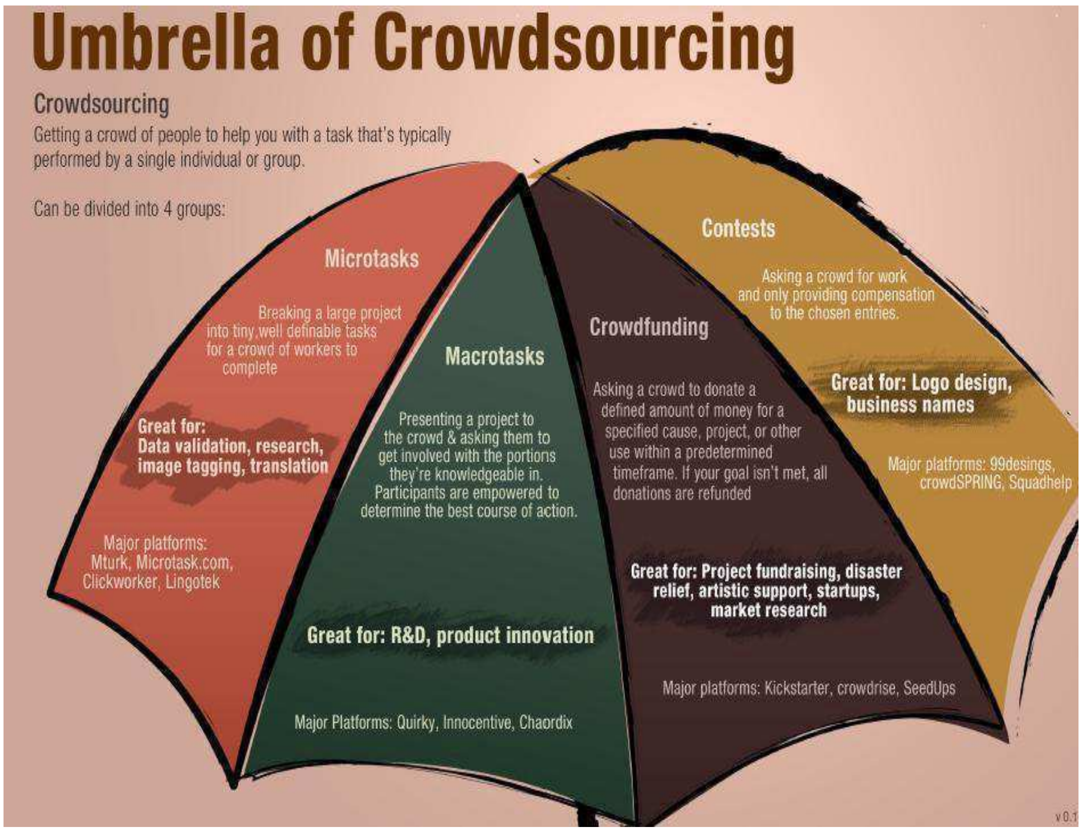
- **To make it more scalable**
  Scaling is a difficult problem for any company to tackle, but it's especially difficult when working on large projects with limited resources. Crowdsourcing, on the other hand, makes it simple to scale up any workforce by farming out small sections of a project that can be performed by remote employees at any time or location. One of the main reasons why firms are interested in crowdsourcing is because of its flexibility.
- **ToFill Knowledge Gaps**
  Unless a company is operating on a massive scale, most of them don't staff all of the resourcesthey need at any given time. Crowdsourcing provides the ability to access people who have skill sets that are unavailable within the company. This can be invaluable for projects or problems that require specialized knowledge or skill sets that are scarce.
- **To Accelerate Processes**
  Crowdsourcing allows businesses to perform tasks more quickly than a single employee. Breaking up a project into a collection of smaller pieces and providing those pieces to a larger group of workers expedites the completion of projects. Overall, crowdsourcing presents a more efficient way to do work.
- **To bridge the gap between what people, know and what they don't know**
  Unless a corporation operates on a large scale, it is rare for it to have all of the resources it requires on hand at any given time. Crowdsourcing allows companies to reach people with skill sets that aren't available within their own organisation. This is especially useful for projects or challenges that call for specialised knowledge or skill sets that are in short supply.
- **Processes are being sped up.**
  Crowdsourcing helps firms to do things faster than they could with a single employee. Breaking down a project into smaller pieces and assigning those pieces to a bigger number of workers speeds up project completion. Overall, crowdsourcing is a more efficient method of accomplishing tasks.
- **To cut down on operational costs**

  Crowdsourcing allows you to finish projects at a lower cost. Businesses can avoid most of the expenditures involved with operations when a group of people collaborates digitally to achieve a task. This covers the price of lodging and paying full salaries to employees, as well as the costs of paying for employees to learn new skills and other expenses. Faster turnaround times might result in higher profitability, depending on the projects being executed.

When a company chooses to seek out customers as part of its crowdsourcing activities, it can get a high level of consumer engagement. The majority of traditional marketing media only keep a consumer's attention for a brief period of time. By encouraging customers to help solve an issue or offer much-desired data about their brand, the company is garnering important attention for which many corporations pay a lot of money.

8.2 INTRODUCTION

Crowdsourcing is a web-based invitation to all members of the public to generate, discuss, revise, and rank relevant ideas, tasks, or contributions.
OUTSOURCING + CROWDSOURCING = CROWDSOURCING



Source: https://miro.medium.com/max/1886/1*mTh3mjwE4pEELEGD6RrIIg.jpeg

In **2006**, writer **Jeff Howe** created the phrase "crowdsourcing." Simply explained, crowdsourcing is the process of delegating a task to a group of people. The crowdsourcing paradigm of mass participation is at the heart of several famous Web 2.0 websites and services, the most well-known of which is Wikipedia. The crowdsourcing principle is used by many types of businesses and organisations to develop innovative ideas and long-term solutions for a common good.
An open call is when a company or organisation takes a function that was previously done by workers and outsources it to an undefined and broad network of people. This can take the form of peer-production when the job is done collectively, but it is also frequently done by

single people. The utilisation of an open call format and a vast network of potential labourers is essential prerequisites.

The word "crowdsourcing" was coined in 2006 by Jeff Howe in his Wired Magazine article "The Rise of Crowdsourcing". Howe defines crowdsourcing as "the act of a firm or institution adopting a function that was previously handled by staff and outsourcing it in the form of an open call to an undefined (and generally huge) network of people". The term "crowdsourcing" is a neologism derived from the words "outsourcing" and "crowd" (Hirth et al. 2011). While outsourcing refers to the outsourcing of internal tasks through bilateral partnerships (Grossman and Helpman 2005), crowdsourcing refers to the outsourcing of activities through an unspecified collection of people known as the crowd. 2012 (Leimeister).

The topic has sparked a lot of attention in both business and research since Howe's publication. This first chapter will present a definition of crowdsourcing in order to establish a shared understanding of the concept for the rest of the book.

There have been a number of attempts to define crowdsourcing in the academic literature. While the majority of them share many characteristics, several of them target different phenomena to varying degrees. Bücheler et al. (2010) and Huberman et al. (2009), for example, consider Wikipedia and YouTube to be examples of crowdsourcing, although other authors expressly reject these platforms (Kleemann et al. 2008; Brabham 2012). Estellés-Arolas and González-Ladrón-de-Guevara made one of the most important attempts to tackle the difficulty of arriving at a widely recognised definition (2012). They found 209 publications with 40 different definitions of crowdsourcing after completing a literature study. They've created an exhaustive and detailed list as a result of this.

Crowdsourcing is a sort of participatory online activity in which an individual, an institution, a non-profit organisation, or a firm proposes the voluntary completion of a task to a group of people with varied levels of knowledge, heterogeneity, and numbers via a flexible open call. The completion of a task of varying complexity and modularity, in which the community is expected to contribute their time, money, knowledge, and/or experience, always results in mutual gain. The user will be satisfied with a specific type of need, such as economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and use what the user has brought to the venture, which will take various forms depending on the type of activity undertaken.

Kietzmann (2017) claims in a recent paper that various technical advancements and their rapid dispersion over the previous few years have resulted in knowledge being accessed considerably more quickly, easily, and efficiently. He derives conclusions that necessitate a broader definition of crowdsourcing from these shifts and the growth of research in the subject. Most of his issues, however, are already covered in Estellés- Arolas and González-Ladrón-de-(2012) Guevara's definition and so do not need to be explored again. Nonetheless, the key distinction is that Kietzmann (2017) assumes that the activity does not have to be completed by humans, but that a mix of humans and machines can potentially act as a crowd.

The wisdom of the masses premise underpins crowd sourcing (Surowiecki 2004). This theory, in turn, is founded on the concept of collective intelligence (Lévy 1997), which explains a group's intelligence as a result of peer interaction. Surowiecki (2004) claims that, under some circumstances, a group of people can make better judgments and create better results than individuals, even though the latter are in theory better equipped to do the job. As a result, crowdsourcing can be utilised to tap into the collective wisdom of the population in order to solve an issue.

Crowdsourcing platforms are classified as "information systems." These information systems, according to Alter, are "a system in which human participants and/or machines do work (processes and activities) using information, technology, and other resources to produce informational products and/or services for internal or external customers."

According to this definition, crowdsourcing-specific information systems are "socio-technical systems that produce informational products and services by harnessing the potential of huge groups of people via the Web," as defined by Geiger and Schader.

The crowdsource or content owner and the crowd are the two main parties involved in crowdsourcing. The content owner is the person who is looking for a solution to an issue, while the audience is made up of the people who are solving it (Leimeister et al. 2015). On an IT-enabled crowdsourcing platform, the crowdsourcing process takes occur. This allows content creators to develop and share problems, which the community can then solve collectively or individually and submit solutions for. If this platform is run by an intermediary, a third function emerges: the crowdsourcing middleman.

Additional roles must be considered when characterising the internal crowdsourcing information system from an intraorganizational perspective. Ulbrich and wedel construct a complicated model explaining the primary, secondary, and tertiary responsibilities required for a successful implementation of internal crowdsourcing (see chapter "Systematization Approach for the Development and Description of an Internal Crowdsourcing System" of this book). There are eight separate positions described: (1) crowd master, (2) campaign owner, (3) crowd technology master, (4) content owner, (5) secondary counterpart, (6) crowd, (7) executive board, and (8) employee union representation.

How to crowd source –

Design the job - After you've opted to use crowdsourced content in your marketing campaign, you'll need to figure out what the job will include. Will you let customers create an advertisement? Perhaps they'll assist you in developing a new product, similar to Lays? Your marketing team must design the job in any case. This means you'll have to identify what you want your audience to do and then provide a way for them to achieve it. You'll need to come up with the rules, regulations, and reward for winning the contest (if it's a contest) after you know what you want them to build. If you're not sure what you want your audience to perform, crowdsourcing can help with a variety of tasks, including:

- App development
- Editing
- Ad design
- Product creation
- Photography
- Copywriting
- Transcription

Create the promotional materials - It's time to get the word out about your employment now that you've developed the position and the terms of service, so to speak. However, before you start posting on social media, you'll need to create the campaign's creative materials.Lays, for example, promoted their "Do Us a Flavor" campaign on Facebook. To promote the campaign, the company updated its profile and cover photo. There were also multiple posts on social media and the company's website that went through the restrictions.You'll need to develop these assets before you can promote your campaign. Write your social media advertising, produce your photos, write your website's press release, and so on.

Choose a promotional strategy - So, you have the job all figured out, you've created your assets, and now you have to decide how you want to promote the campaign. The next step is to choose the social media networks you wish to use to promote the campaign. To figure out where your audience is, use your social media stats and buyer persona. If your target demographic is predominantly Gen Z, for example, you could wish to promote on Tik Tok. Instagram, on the other hand, might be the ideal location to reach out to millennials. Overall, make sure you're communicating with your target audience where they are.

Manage the results - The results should start flowing in now that you've launched your campaign. Make sure you have a system in place to manage the submissions. Will you, for example, have one or more staff in charge of saving and arranging submissions? Alternatively, you might have your audience send submissions to a specific email address.

In any case, having a system in place will assist you in keeping things organised so that selecting a winner is simple when the time comes.

Produce the final project - It's now or never. You've gotten all of the results, and it's time to choose a winner. The winner of a crowdsourced campaign is usually decided by a public vote. If you're going with a vote, make sure you have polls and surveys in place so that your audience may participate. It's time to advertise the final campaign once the winner has been decided. If you're choosing a new logo, for example, you can start utilising it on social media, your website, and advertising materials right away.

Check out the following crowdsourcing sites whether you're searching for work as a freelancer or a brand looking for talent, or if you're looking for crowdfunding.

1. Fiverr.com

Fiverr is a marketplace for freelance services that empowers freelancers. Rather from being a location where freelancers browse for tasks offered by brands, Fiverr is a place where brands look for freelancers with the expertise and abilities they require.

Graphic design, digital marketing, writing & translation, video & animation, music & audio, programming & tech, business, and lifestyle are among the most popular services offered by Fiverr freelancers.

2. Upwork

Similar to Fiverr, Upwork is a freelance service marketplace where freelancers create profiles, and then brands can hire them for short-term tasks, recurring projects, or full-time contract work. Most freelancers have skills and expertise in web development, mobile development, design, writing, administrative work, customer service, sales, marketing, accounting, and consulting.

3. Make use of crowd sourcing

CrowdSource has trained, tested, and qualified a community of over 200,000 freelancers who can supply copywriting, content moderation, data entry, and transcription experience and talents, and is trusted by businesses like Target, Coca-Cola, and Major League Baseball. Brands can also look for freelancers by industry, such as advertising, marketing, publishing, retail, and service providers.

4. Believably

Contently is a content production platform that connects enterprise brands with freelance talent, so they're always on the lookout for freelancers who can meet both their clients' and their own requirements.

If you're a freelance creative searching for work with big companies, you can construct a free portfolio on Contently's platform by registering as a freelancer. Before you can work with any of their clients, you'll need to get approved and finish their training, but once you do, you'll be a part of their freelance network. Check out Contently's platform if you're a brand searching for freelancers to help you generate fresh stories.

5. Skyword

Like Contently, is a content creation platform that also connects businesses with freelancers. You can construct a portfolio that Skyword's clients will have direct access to if you're a videographer, writer, photographer, or designer. Check out Skyword's platform if you're a brand seeking for freelancing talent.

6. Crowdfunding is one of the most common kinds of crowdsourcing.

Consumers sponsor a project or endeavour by donating money to the cause through crowdfunding.Kickstarter is one of the most well-known crowdfunding platforms, allowing any creator to interact with potential backers. This implies they can publicise their project and accept donations from anyone.

People who donate money to a crowdfunding campaign typically become stakeholders and may be allowed to collaborate with the creator as crowdsourced talent in exchange for a cut of the revenue.

8.3 OVERVIEW OF CROWD SOURCING

The word "crowdsourcing" was coined and first used a decade ago to describe a means for expressing the wisdom of the crowd in achieving two types of activities. The first category comprises jobs that require human intelligence rather than robots, while the second category includes tasks that can be completed with greater time and cost efficiency by using the crowd rather than experts. There are five modules to the crowdsourcing process:

The first step is to provide incentives to encourage people to perform the required work. This is followed by four modules for gathering and ensuring quality, as well as verifying and aggregating the information obtained. To eliminate errors from unreliable participants, verification and quality control can be done for the tasks, gathered data, and participants by having multiple participants answer the same question or accepting only answers from experts.

So i&P =0

Topic expert methods are used to find trustworthy candidates in the audience who have relevant experience in the subject topic. Expert discovery minimises the number of participants required per question, lowering the total cost. The methods used to complete each processing step are summarised and reviewed in this paper. However, the method you use depends on the application.

Historical examples of Crowdsourcing –

- **Competitions in the early stages**

In the past, crowdsourcing was frequently employed as a competition to find a solution. Several of these tournaments, frequently rewarded with Montyon Prizes, were proposed by the French government for poor Frenchmen who had performed noble acts. The Leblanc process, also known as the Alkali Prize, offered a reward for extracting the salt from the alkali, and the Fourneyron's turbine, which was the first hydraulic commercial turbine. Nicolas Appert earned an award for inventing a new method of food preservation that entailed sealing food in air-tight jars in response to a challenge from the French government. In the Longitude Prize, the British government offered a comparable incentive for finding an easy means to establish a ship's longitude. Out-of-work clerks tallied higher mathematical functions in the Mathematical Tables Initiative as an outreach project during the Great Depression.

- **In Energy system models**

It requires extensive and diverse datasets, which are becoming more important as the trend toward higher temporal and spatial resolution continues. As a result, various projects to crowdsource this data have emerged. OpenEI is a collaborative website run by the US government that provides open energy data. It was launched in December 2009. While much of the data comes from US government sources, the portal also welcomes crowdsourcing contributions from all over the world. Enipedia, a semantic wiki and database, also publishes energy system data utilising the crowdsourced open information idea. Enipedia was launched in March of 2011.

- **In the field of genealogy**

Crowdsourcing approaches were used in genealogical research long before personal computers were widely used. The Church of Jesus Christ of Latter-day Saints began encouraging members to submit information about their ancestors in 1942. The information that was submitted was compiled into a single collection. The three-generation initiative was created by the church in 1969 to encourage more individuals to participate in obtaining genealogical information about their ancestors. Church members were asked to complete documented family group record forms for the first three generations as part of this programme. The four-generation programme was created after the programme was expanded to encourage members to explore at least four generations.

Institutes with records of interest to genealogical research have enlisted the help of crowds of volunteers to create catalogues and indices.

Advantages and Disadvantages:

A hybrid of "crowdsourcing" and "outsourcing," "The technique of getting essential services, ideas, or content by asking contributions from a wide number of individuals, particularly from the internet community, rather than from traditional employees or suppliers," according to Merriam-Webster.

This word refers to the practise of enlisting the help of the educated people to assist in the testing procedure. This method has a number of advantages, including the ability to contact a larger number of testers and a potential higher return on investment for the testing procedure. However, there are drawbacks, such as challenges in maintaining secrecy and communicating with all parties involved.

Advantages:

- Cost-effective: the company only pays for bugs that are found, rather than an hourly or salaried rate that professional testers would receive
- Wide range of users provide a wide range of experiences
- Allows for testing with all kinds of different parameters, such as different connection speeds, browsers, and devices that the core testing team may not have access to
- Larger group is more likely to find reproducible bugs than a handful of testers
- Testers are unlikely to be biassed against the company. The monetary value of both of the above= more extensive testing for an equal price range in a shorter amount of time, without the need for a contract or overhead.

Cons:

- Having testing done by a big group that may or may not be confidential compromises confidentialityhaving a low level of brand/product loyalty.
- Due to time or language difficulties, communication amongst testers can be challenging.
- Test coverage can be more difficult to ensure, necessitating greater managerial effort.
- To ensure thorough bug identification, there will be a supervisory role.
- Because testers are frequently paid based on the number of defects detected rather than their severity, they are paid on a per-bug basis.
- Rather than devoting a lot of time and effort to finding and identifying a few large, important bugs, may search out and discover a large number of tiny, unimportant bugs.
- A concerted effort to locate one or two enormous, incapacitating bugs.

- This, paired with crowdsourcing's experimental nature, makes it difficult to assure that the entire product has been fully evaluated to ensure usability.

Some businesses opt to use crowdsourcing in addition to or as a supplement to their own internal testing teams. This allows for a far more thorough testing procedure, but it might make experienced testers feel like their abilities are being discounted and outsourced to less competent individuals. If a crowdsource tester finds no bugs, they are not paid; this might be perceived as tying their hard-earned talents and education to unpaid labour in the eyes of a professional tester. It can also jeopardise a lot of things.

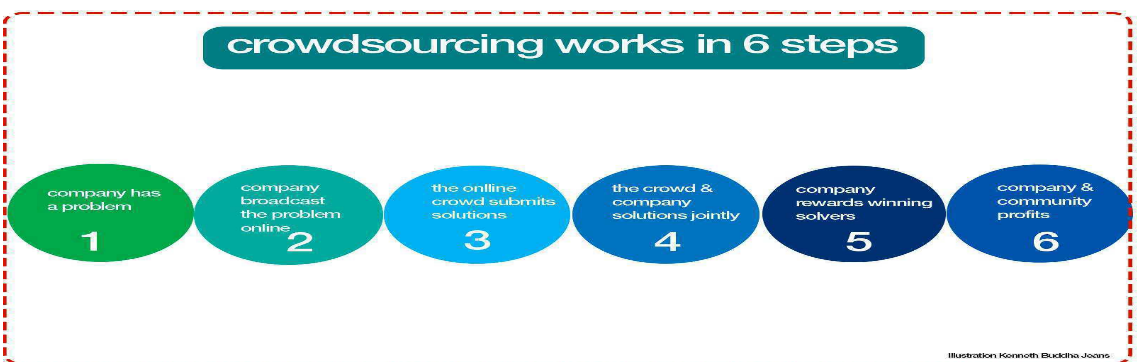
Because the corporation must pay their test teams' wages on top of the per-bug payout, White Paper the cost-effective nature of crowdsourcing.

8.4 SUMMARY

According to the study, crowdsourcing is credited with assisting in the creation of outstanding acts of journalism. It has transformed newsgathering by introducing unprecedented opportunities for attracting new sources with new voices and information, allowing news organisations to unlock stories that would otherwise go unnoticed, and allowing news organisations to experiment with the possibilities of engagement for the sheer joy of it.

In short, it has accomplished exactly what the pundits anticipated a decade ago: it has aided in the transformation of journalism into a discussion rather than a one-way megaphone.

As data or stories come in from contributors, reporters see new possibilities for their journalism—and news organisations see opportunities to incrementally publish those contributions in ways that tease out more—crowdsourcing deserves credit for helping to shape journalism into a more iterative process.



Source:https://i1.wp.com/buddhajeans.com/wp-content/uploads/2016/11/crowdsourcing-1600-1.png?resize=1600%2C960&ssl=1

Crowdsourcing, on the other hand, may be high-touch and high-energy, and not all projects succeed the first time.

For all its promise, crowdsourcing is widespread and systemic at only a few major news organizations—for example, ProPublica, WNYC, and The Guardian. Only a few reporters and editors—not the institutions themselves—are the standard bearers at other prominent news organisations, such as CNN Digital and The New York Times.

Crowdsourcing enterprises are thriving outside of journalism, to be sure. However, further systemic adoption within the news industry may rely on more than seasoned practitioners' enthusiasm and accolades from sources who are ecstatic about the outreach.

We'd like to see more research and evidence on whether crowdsourcing can help journalism gain more support. That support could come in the form of audience participation, such as attention, loyalty, time spent on a site, repeat visits, or personal tales contributed. It could also entail financial support from members or contributors, advertising who want to connect themselves with the profession, or funders who wish to help.

It will also be investigated whether crowdsourced articles have a greater real-world impact than traditional forms of media, such as prompting legislative change.

Some news organisations may be hesitant to join the ranks of long-time practitioners and commit the time and resources required to support crowdsourcing projects until this data is available and a better suite of tools and techniques is built.

The phrase "crowdsourcing" is still in its infancy, and it is constantly evolving as new applications emerge. Following a review of a number of scientific papers, it was discovered that different definitions of crowdsourcing exist, demonstrating a lack of consensus and some semantic misunderstanding.

This article presents a broad definition of crowdsourcing that encompasses the vast majority (if not all) of existing methods. The crowd, the task at hand, the remuneration obtained, the crowdsourcer or initiator of the crowdsourcing activity, what they obtain following the crowdsourcing process, the type of process, the call to participate, and the medium were found to be eight characteristics common to any given crowdsourcing initiative after analysing all of the authors' definitions. For each of these aspects, an analysis was conducted based on the collected definitions, and a conclusion was produced, with the goal of making each element as global as feasible.

While attempting to preserve the highest level of precision The collaboration of these findings has resulted in the formulation of a worldwide definition that encompasses all of the compared crowdsourcing activities.

Each form of particular crowdsourcing activity (crowd voting, crowdfunding, etc.) will require a more exact definition of each of the eight aspects based on this universal concept. For example, in the case of crowdfunding, the crowd's goal will be to donate money, whereas in the case of crowd voting, the crowd's task will be to vote for and provide feedback on specific products.

There is limited agreement in other areas of crowdsourcing, such as the classification of different sorts of crowdsourcing activity. In light of this, some work evaluating, recompiling, and summarising with the purpose of unifying some of the perspectives would be of interest.

8.5 GLOSSARY

- **Crowdsourcing-** Crowdsourcing is the process of enlisting the help of a large number of individuals (the crowd) for a common goal, which is frequently related to problem-solving, efficiency, or creativity. Individuals can pool their ideas, funds, time, and experience to support a cause or initiative. These people are linked by a common goal of achieving greater results through their combined efforts.
- **Governance** – The necessity in this business is for value. So, in this situation, crowdsourcing leads to equity-effective funding by investors and executives, improves ROI, and improves customer experience and valuations.
- **Innovation -** Product management and R&D teams can use Crowdsourcing to gain new ideas, user-driven design, and collaborative partnering to meet their demand for progress.
- **Finance -** Finance requires optimization, which crowdsourcing may assist with through risk mitigation, cost reduction, and outsourcing alternatives.
- **Commercial -** The commercial enterprise requires expansion, and crowdsourcing may assist in this by making it easier to obtain new leads and improving the entire client experience.
- **Operations -** The Operations enterprise has a need for better service. As aresult, using crowdsourcing to manage supply chain efficiencies and quality might be beneficial.

8.6 ANSWER TO CHECK YOUR PROGRESS

Q.1 What is the most crucial information about crowdsourcing that you should be aware of?

Ans. The date is July 8, 2019. The technique of soliciting work, information, or opinions from a wide number of individuals via the Internet, social media, or smartphone apps is known as crowdsourcing. In certain circumstances, crowdsourcing participants work as paid freelancers, while others complete modest jobs for free.

Q.2 What makes crowdsourcing a viable alternative to traditional financing?

Ans. Crowdsourcing, as an alternative to traditional funding, draws into a group's common interest, circumventing the typical gatekeepers and intermediaries required to raise funds. Crowdsourcing usually entails breaking down a huge job into several smaller jobs that a group of people may work on independently.

Q.3 What crowdsourcing platform does Amazon use?

Ans. Mechanical Turk (MTurk) is an Amazon crowdsourcing marketplace that allows businesses to outsource various aspects of their operations, such as data validation, research, and content moderation. Anyone can become a Mechanical Turk Worker by using their Amazon account.

Q.4 What is the best code crowdsourcing platform?

Ans. With the rise of no-code development platforms in 2021, a crowdsourcing platform like uTest provides a clear value to those who participate. Assist in shaping the future of technology (smartphone apps, websites, and so on) to ensure that it meets the needs of those who will use it.

Q.5 What exactly is crowdsourcing, and how does it function?

Ans. Complex business problems are solved using crowdsourcing, which employs an on-demand, scalable workforce. Crowdsource assists you in breaking down large projects into smaller, more manageable microtasks, which are then completed by our virtual workforce of professional individuals.

Q.6 What is the social impact of crowdsourcing?

Ans. Crowdsourcing affects many aspects of social and professional interactions. It is altering the way we work, hire, research, manufacture, and promote our products. Governments are turning to crowdsourcing to empower citizens and offer them a stronger voice. Crowdsourcing can democratize problem solving and expedite innovation in research and health care.

Q.7 In Key Takeaways, what is crowdsourcing?

Ans. The Most Important Takeaways Crowdsourcing is the process of gathering information, opinions, or work from a large number of individuals using the Internet. Crowdsourcing allows businesses to save time and money while gaining access to people with a variety of skills and perspectives from all around the world.

Q.8 Is it possible to generate money via crowdsourcing?

Ans. Crowdsourcing participants operate as paid freelancers in some cases, while others complete little jobs for free. Traffic apps, for example, encourage drivers to report accidents and other road incidents so that app users can get real-time updates.

Q.9 What instances of crowdsourced transportation do you have?

Ans. Uber is an example of crowdsourced transportation, as it connects available drivers with individuals who need trips. While crowdsourcing is frequently used to break up a large project, firms can also use it to evaluate how several people perform at the same time.

8.7 REFERENCES

- https://www.bing.com/search?q=crowdsourcing&cvid=ecdc476a9f6a4d28be18fe00a050b6f3&aqs=edge.0.69i59l2j69i57j69i60l6.5861j0j1&pglt=2083&PC=U531&first=11&FORM=PERE
- https://link.springer.com/chapter/10.1007/978-3-030-52881-2_2

- https://www.researchgate.net/publication/310750128_Crowdsourcing_A_revised_definition_and_introduction_to_new_research
- https://www.bing.com/search?q=advantages+and+disadvantages+of+crowdsourcing&cvid=6b1684bbacdd48be9c52e21f8d825e83&aqs=edge..69i57.16751j0j1&pglt=2083&FORM=ANNTA1&PC=U531
- https://www.bing.com/newtabredir?url=https%3A%2F%2Fwww.researchgate.net%2Fpublication%2F310750128_Crowdsourcing_A_revised_definition_and_introduction_to_new_research
- https://www.bing.com/search?q=overview+of+crowdsourcing&form=ANNTH1&refig=13a21f1d20614b5fbfaed9fd0bc589a9
- https://datasciencecmu.wordpress.com/2014/03/28/introduction-to-crowdsourcing/
- https://searchcio.techtarget.com/definition/crowdsourcing
- https://www.bing.com/search?q=crowd+sourcing&cvid=f90c64c60713444789f97227e9a865fe&aqs=edge..69i57j0l8.13196j0j1&FORM=ANSPA1&PC=U531
- https://www.bing.com/search?q=google&form=ANNTH1&refig=ec93e8b557e344cf83d3cee77cb3bf4b
- https://www.bing.com/newtabredir?url=https%3A%2F%2Fwww.slideshare.net%2Fmirkopresser%2Fcrowdsourcing-an-overview
- https://www.bing.com/newtabredir?url=https%3A%2F%2Ftowcenter.gitbooks.io%2Fguide-to-crowdsourcing%2Fcontent%2Fexecutive_summary%2F

8.8 TERMINAL QUESTIONS

1. How do you crowd source?
2. What are 2 examples of crowdsourcing?
3. Why are most companies interested in crowdsourcing?

4. What are the benefits of using crowdsourcing model?
5. What are the steps included in crowdsourcing?

UNIT 9 : MOBILE APPS FOR CROWD SOURCING

9.1 OBJECTIVES

In order to solve a wide range of problems, crowdsourcing is frequently used. The current unit focuses mostly on applications relating to geographic space, where crowdsourcing is used to collect or even create geographic data from scratch. It is generally combined with GIS in such circumstances, easing the entire data collection process and thereby contributing to the establishment of Integrated Public Participation GIS tools.

According to Goodchild, numerous mechanisms (in a geographic context) for data collecting by distant persons exist, resulting in so-called "volunteered geographic information" (VGI) being generated and disseminated. VGI is increasingly being used as a bottom-up process in geographic applications, such as map generation or elaboration by several users. Data is created "by the users for the users" in a decentralized system.

As a result, crowdsourcing has been integrated into other tools, and it is now a technology that may support spatial and collaborative planning, cartography, and a wide range of other spatial applications that require the "knowledge of the crowd."

9.2 INTRODUCTION

Definition: Crowdsourcing activities that are carried out on smartphones or other mobile devices are referred to as mobile crowdsourcing. Mobile phone users may work on crowdsourcing chores without any additional challenges because to better, modern smartphone features such as dependable GPS, excellent cameras, and constantly updated apps. These jobs now entail more than just site descriptions.

Human computation advocates integrating humans and machines to help solve increasingly difficult issues. As proof, crowdsourcing has shown to be an excellent method of completing jobs that are simple for humans but challenging for machines. The term "crowdsourcing," which is a kind of “peer production,” to describe ways that outsource labor to a wide number of ordinary people? Platforms like Amazon Mechanical Turk have made it possible to build scalable apps for a wide range of activities, from product classification and photo tagging to voice transcription and translation.

Smartphones are now widespread and widely used across the world, with various radios (e.g., Bluetooth, Wi-Fi, and Cellular) and rich sensors (e.g., GPS, accelerometer, and camera) built in. Individuals using smartphones can perceive, gather, analyze, and share data around them at any time and in any location thanks to this trend. Naturally, combining smartphone-based mobile technologies with crowdsourcing provides a lot of flexibility and gives rise to a new computing paradigm known as mobile crowdsourcing (MCS), which may be fully explored for real-time and location-sensitive crowdsourced jobs.

Crowdsensing is a new area of interest for study and applications in MCS, in which the physical world may be felt without constructing a sensor network by utilizing billions of users' mobile devices/phones as location-aware data gathering instruments for real-world observations. Based on the sort of events being observed, MCS sensing applications may be divided into two categories: personal and community sensing. Human phenomena are monitored in personal sensing applications; community sensing applications monitor large-scale phenomena that are difficult to quantify by a single individual. Participatory and opportunistic community sensing are two types of community sensing. Participatory sensing involves the user being personally involved in the sensing action, such as photographing certain sites or activities. The user is unaware of active applications in opportunistic sensing, but the crowdsourcing App for smartphones will take care of everything.

MCS offers a wide range of applications outside of sensing, including environmental monitoring, intelligent transportation, tailored medicine, and epidemiological research.A comprehensive MCS framework and typical MCS workflow are designed, consisting of nine functional modules pertaining to three stakeholders: the publisher of crowdsourced tasks (i.e., crowd source), crowd workers who complete the crowdsourced tasks for monetary compensation, and the crowdsourcing platform that facilitates the interaction between crowdsources and crowd workers. The issues of MCS systems and their solutions, such as task management, incentives, security and privacy, and quality control, are then examined. This document extensively explains the required features that MCS systems should meet from multiple stakeholder perspectives.

9.3 MOBILE APPS FOR CROWD SOURCING

In app stores, crowdsourcing mobile apps are blooming like mushrooms. If you've ever used Uber, Airbnb, or Easy Taxi, you're definitely familiar with how they work. Our digital habits are being shaped by people sharing resources for consumption, whether it's a product or a service. We have become more interdependent users than ever before, from booking a car via mobile to sharing rooms or flats, to open Q&A forums like Quora, to crowdsourced travel information like MiFlight.

1. **Be My Eyes**

Source: https://www.maketecheasier.com/assets/uploads/2015/01/crowdsourcingmobileapps-bemyeyes.jpg

Be My Eyes, from the Lifestyle category, uses a live audio-video connection to help blind persons in need. It allows blind people to request assistance, and a crowd of sighted people will respond (whoever is available) and tell the blind person what they see through the rear-facing camera. Sighted people can accept or reject the call based on their availability, and they don't have to worry if they miss one because it can be picked up by someone else.

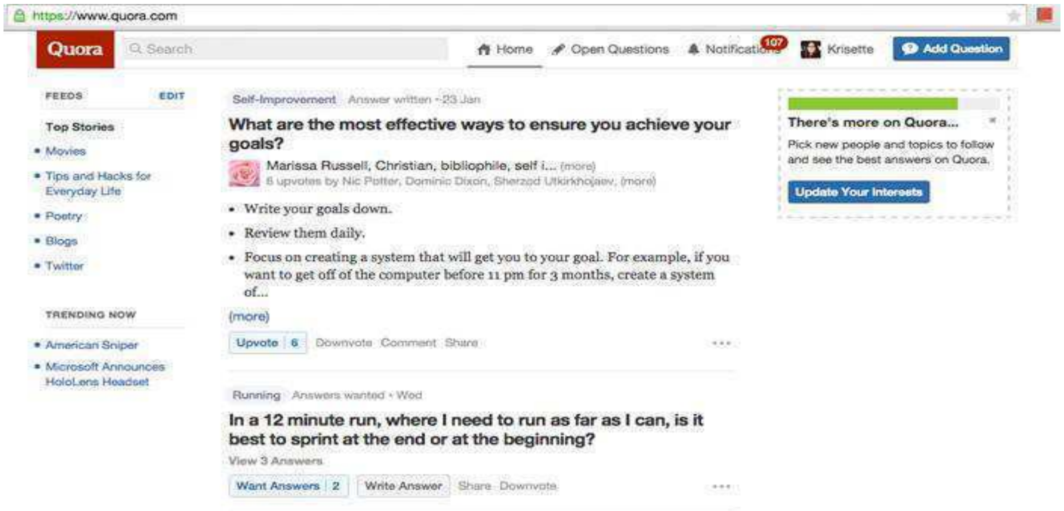
2. FireChat

Source: https://www.maketecheasier.com/assets/uploads/2015/01/crowdsourcingmobileapps-firechat.jpg

This was one of the most popular crowd apps during Hong Kong's pro-democracy movement, allowing users to speak with individuals in their immediate vicinity without the usage of a mobile network or an Internet connection (on board, just turn on Airplane mode and activate Bluetooth or Wi-Fi). Unlike WhatsApp or WeChat, however, FireChat (iOS app | Android app) has public chat rooms where everyone – the public – can read the messages and the crowd can comment in real time and anonymously because it doesn't require real names. You can start your own chat rooms to connect with other Game of Throne’s fans or organize local campaigns and activities.

3. Quora

Source: https://yhoo.it/3EjmqEq



Source:
https://www.maketecheasier.com/assets/uploads/2015/01/crowdsourcingmobileapps-quora1.jpg

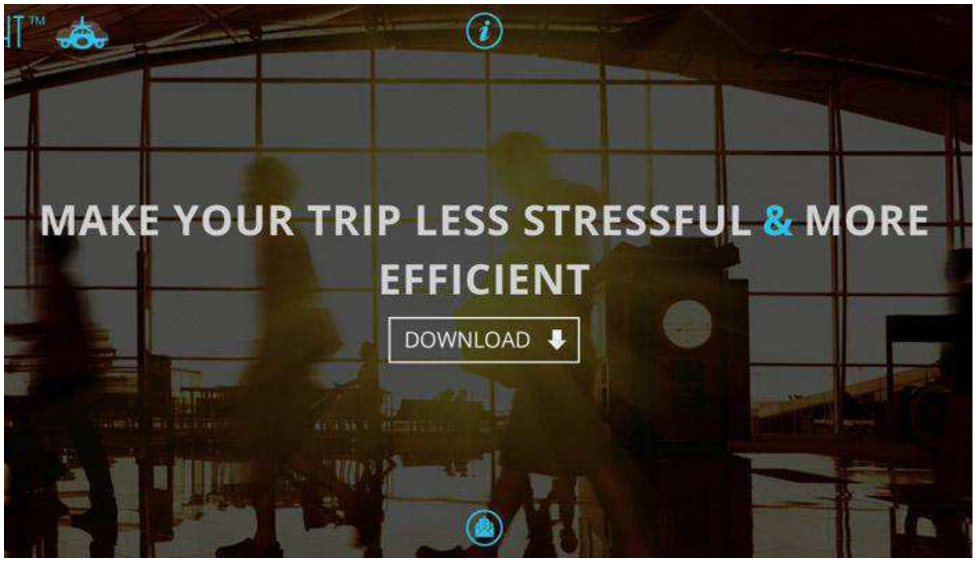
If you ask, you'll get answers from people who are passionate about their profession and are experts in their field. You can ask your queries and the crowd will respond at the same time. This open community promotes learning and invites you to participate in answering questions from other users. It includes social media integration, allowing you to connect your Facebook, Twitter, and other accounts with other followers.

4. **Figure 1**

Source:
https://www.maketecheasier.com/assets/uploads/2015/01/crowdsourcingmobileapps-figure1.jpg

This software is only available to medical professionals, and it allows them to publish and share images as well as connect with others for future collaboration. For doctors and medical workers, think Instagram. They can share their discoveries, access photographs, discuss research, and more while respecting the privacy of the patients.

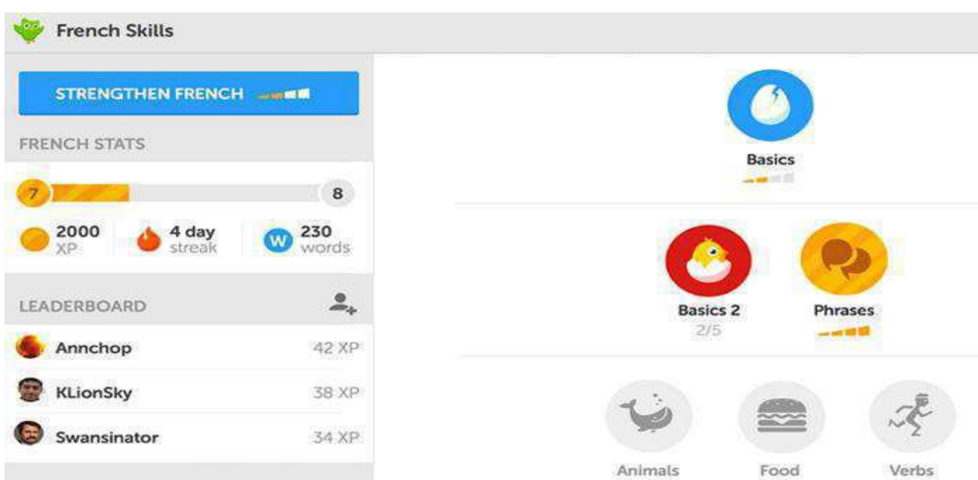
5. **MiFligt**



Source: *https://www.maketecheasier.com/assets/uploads/2015/01/crowdsourcingmobileapps-miflight.jpg*

MiFlight provides you real-time crowdsourced updates – through social sharing from the travelers – on airport lines and security checkpoints at the terminals. How long is the line? Stay informed, and MiFlight provides you maps of the airport, too.

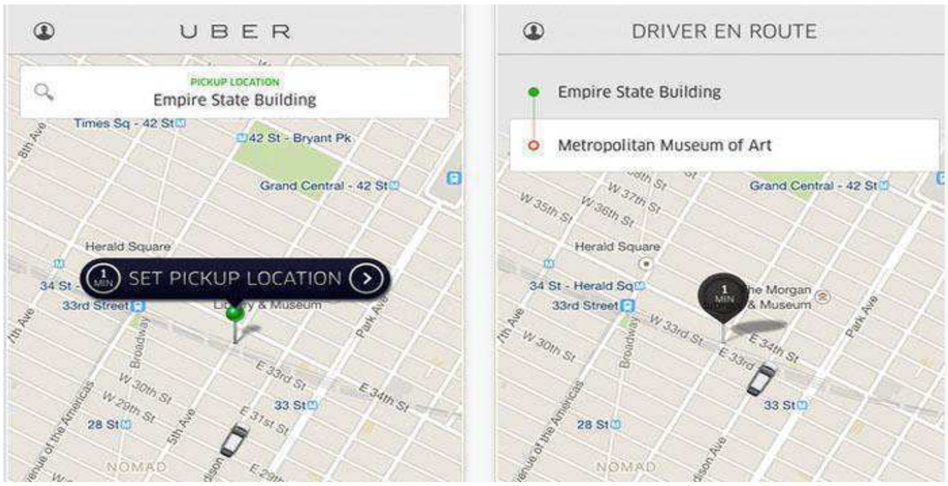
6. **Duolingo**



Source: *https://www.maketecheasier.com/assets/uploads/2015/01/crowdsourcingmobileapps-duolingo.jpg*

Hola! Machine translations are also growing more popular, and Duolingo's crowdsourced language translation platform allows you to learn a new language for free in a fun, engaging, and game-like manner using your smartphone. Keep track of your progress and share it on social media.

7. Uber



Source:
https://www.maketecheasier.com/assets/uploads/2015/01/crowdsourcingmobileapps-uber.jpg

While on the way, Uber provides you with a private automobile. This on-demand service, which is now accessible in over forty countries, allows you to skip the taxi line and save time. Riders can pay using PayPal or credit card, compare fares, and be picked up in a matter of minutes from their current position.

8. Airbnb

Source:
https://www.maketecheasier.com/assets/uploads/2015/01/crowdsourcingmobileapps-airbnb.jpg

It's like having a home away from home. Travelers can use their mobile devices to identify one-of-a-kind places to stay while in transit with Airbnb. Thousands of ads are available on the platform, allowing passengers to compare costs while also allowing hosts to earn money by renting out their space, whether it's a single room, a full apartment, or even a Tesla S car with a sleeping bed inside for $85 per night!

9. Nexar

Source: https://www.fulcrumapp.com/blog/crowdsourcing-apps

Nexar is a crowdsourced dashcam app for smartphones. You're connected to their Nexar vehicle-to-vehicle network whether you use their Nexar app or a Wi-Fi-enabled dashcam, so the more users nearby, the safer we all are. Because vehicles convey proximity warnings to alert drivers of threats beyond their field of view, Nexar makes travel safer and smarter. The software analyses the road and warns the driver of potential hazards. It captures every drive, syncs to the cloud while connected to the internet, and then deletes the video from the device to save space. Nexar creates a complete, printed record of an accident so you may precisely tell your side of the story using crash recreation.

10. Pavemint

PAVEMINT™

Source: https://mma.prnewswire.com/media/564365/Pavemint_Logo.jpg?p=publish

Source: https://i.ytimg.com/vi/j0Zm-iEcRkI/maxresdefault.jpg

Pavemint connects those looking for parking with people who have parking spaces to share, allowing you to locate a spot when you need it and earn money when you don't. Consider Pavemint to be the Uber of parking places. It's a peer-to-peer parking network that links people looking for parking with those who have available parking. You may book a spot on Pavemint or rent out your driveway during events if you live in a high-density region or near a popular event location, and start earning extra money.

11. Ring Doorbell

Source: https://i.ytimg.com/vi/bR9CmKiY7ek/maxresdefault.jpg

Source: https://bit.ly/3odCQIS

A smart video doorbell is the subject of this crowdsourced consumer device. In 2013, it made it debut on the ABC show Shark Tank. It goes above and beyond existing smart home camera systems by connecting you to the footage from your neighbors' cameras. The mobile app will provide you with real-time crime and safety alerts from your neighborhood after you have a Ring doorbell and join up for their free community service. Keep in touch and informed with the latest news from your neighbors, community, and local law enforcement so that you can all work together to make your neighborhood a safer place. In 2016, Ring worked with the Los Angeles Police Department to deploy Ring Video Doorbells in 10% of homes in Wilshire Park, California. The Los Angeles Police Department reported a 55 percent reduction in break-ins in the first six months after the doorbells were installed.

9.4 SUMMARY

Crowdsourcing is the process of delegating duties that would normally be completed by an employee or contractor to a wide group of people via the Internet via an open call. Mixing smartphone-based mobile technologies and crowdsourcing offers significant flexibility and leads to a new paradigm called mobile crowdsourcing (MCS), which can be fully explored for real-time and location-sensitive crowdsourced tasks, thanks to the rapid development of smartphones with rich built-in sensors and multiple ratio interfaces. We provide a taxonomy for MCS applications that is expressly divided into two categories: using humans as sensors and leveraging crowd wisdom (i.e., human intelligence). In addition, two paradigms for mobilizing MCS users are described: direct mode and word of mouth method. A thorough MCS structure and typical MCS application workflow are given, consisting of nine functional modules that relate to three MCS stakeholders: crowdsourcers, crowd workers, and crowdsourcing platform. Then we go into the MCS problems in detail, including task management, incentives, security and privacy, and quality control, as well as the solutions available. We present the required features that an ideal MCS system should satisfy, based on the perspectives of numerous stakeholders. The major purpose of this study is to classify and summarize the MCS framework, difficulties, and potential solutions in order to draw attention to MCS-related research areas and make it easier to design and deploy innovative MCS applications.

In recent years, crowdsourcing has evolved as a distributed problem-solving and business-production paradigm. In the crowdsourcing paradigm, jobs are spread to a network of individuals to accomplish, lowering the cost of production for the crowdsourcer. On the other hand, in the post-PC era, the movement of desktop users to mobile platforms, along with the unique multi-sensing capabilities of modern mobile devices, is projected to unlock Crowdsourcing's full potential. MCS, as a new study and application area, raises a slew of concerns to be addressed, including quality control, task design, motivation, security and privacy, and so on. Academic researchers, industrial designers/developers, and policymakers will have unprecedented opportunity as a result of these issues.

This unit identifies two paradigms for mobilizing rational humans in MCS: direct and WoM modes, as well as a taxonomy for numerous MCS applications. Then, we show a general MCS architecture made up of numerous functional modules, as well as a typical MCS application workflow. Finally, we propose the desired qualities that MCS systems (including WoM-based) should satisfy from the perspectives of various stakeholders (i.e., crowdsourcers, crowd workers, and crowdsourcing platforms). Several integrated MCS systems that meet numerous difficulties and have certain desirable qualities are discussed, as well as their drawbacks. In summary, this unit may help with MCS application research, development, and implementation.

9.5 GLOSSARY

- Crowdsourcing - is used to collect data from a large number of users using a single app on their mobile devices.
- Mobile app - is a computer program or software application designed to run on a mobile device.

9.6 ANSWERS TO CHECK YOUR PROGRESS

1. What is crowdsourcing?
2. What is mobile crowd sourcing? Write down its uses.

9.7 REFERENCES

- https://www.fulcrumapp.com/blog/crowdsourcing-apps
- https://www.maketecheasier.com/crowdsourcing-mobile-apps/?amp
- https://www.google.com/url?sa=t&source=web&rct=j&url=https://www.researchgate.net/publication/294424491_Mobile_crowdsourcing_Framework_challenges_and_solutions&ved=2ahUKEwjpy9OGi8X0AhXxS2wGHQtCBpIQFnoECBwQAQ&usg=AOvVaw0ndwWtkHZ8atiocdrROUgg

9.8 TERMINAL QUESTIONS

1. What is mobile crowdsourcing and how does it work?
2. What do you think of these crowdsourcing mobile apps?
3. Have you tried some of them? Share your experience in the comments below.
4. Do you know some the apps from the government that crowdsource information from the public?
5. What the best crowdsourcing platforms?
6. Does Netflix use crowdsourcing?

UNIT 10 : MODERATION OF CROWD SOURCED DATASETS

10.1 OBJECTIVES

Crowdsourcing involves obtaining work, information, or opinions from a large group of people who submit their data via the Internet, social media, and smartphone apps. People involved in crowdsourcing sometimes work as paid freelancers, while others perform small tasks voluntarily. Its two objectives are:

The first is to inspire machine learning researchers to discover unexpected applications of crowdsourcing in their own research. Because of this, relatively more space is devoted to applications that may be less familiar to the machine learning community, such as hybrid intelligence systems and behavioral studies of users interacting with technology, compared with more familiar applications like data generation and model evaluation.

The second goal is to introduce machine learning researchers to the extensive cross disciplinary literature on crowd behavior and motivate why understanding this literatureis crucial for successfully employing crowdsourcing in research.Several other crowdsourcing surveys have been published, each focused on a differentset of themes and problems.

Readers interested in gaining a broader perspective on crowdsourcing may be interested in recent surveys and guides from the computer vision community the databases community and the marketingcommunity or the older but still widely cited and applicable how-to guides on crowdsourcing user studies

10.2 INTRODUCTION

Crowdsourcing allows us to harness the power of human computation to solve tasks that are notoriously difficult to solve with computers alone, such as determining whether or not an image contains a tree, rating the quality of a website, or verifying the phone number of a business. The machine learning community was early to embrace crowdsourcing as a tool for quickly and inexpensively obtaining the vast quantities of labeled data needed to train machine learning systems. Crowdsourcing has been used to generate the image annotations that are needed to train computer vision systems provide the linguistic annotations needed for common natural language processing tasks and collect the relevance judgments needed to optimize search engines.

10.3 MODERATION OF CROWD SOURCED DATASETS

889This simple idea—that crowds could be used to generate training data for machine learning algorithms—inspired a flurry of algorithmic work on how to best elicit and aggregate potentially noisy labels. Meanwhile, machine learning researchers have begun to put crowdsourcing to use in other ways, most commonly as a tool to evaluate and debug machine learning models. Crowdsourcing has flourished as a research tool outside of the machine learning community as well. In human-computer interaction and related fields, researchers are building “hybrid intelligence systems” with the goal of expanding the capabilities of current AI technology by incorporating humans in the loop. And psychologists and social scientists have increasingly moved experiments that traditionally would have been run in physical labs ontocrowdsourcing platforms While these bodies of research are less well known within the machine learning community, there are countless opportunities for machine learning research to both influence and benefit from these lines of work. For example, human-in-the-loop clustering algorithms have been designedthat produce better clusters by drawing on the common-sense knowledge and experience ofthe crowd whilebehavioral experiments run on the crowd offer insight about how to encourage human trustin algorithmic predictions. The application cans broadly categorizeas:

• **Data generation:** Crowdsourcing platforms are well suited to generating data, but challenges arise since the data supplied by crowdworkers can be prone to errors. It can start with a brief review of two lines of research aimed at improving the quality of crowdsourced labels. The first assumes that data points are redundantly assigned to multiple workers and seeks algorithms for aggregating workers' responses that consider the quality of individual workers. Though there are many notable exceptions, much of this work builds on the influential model and expectation-maximization framework of Dawid and Skene (1979). The second line of work focuses on developing incentive schemes to motivate high quality responses. Much of this work builds on the literature on peer prediction, a framework in which crowdworkers' payments are a function of their own reported labels and the labels of other workers. We then review ways in which crowdsourcing has been applied to generate other forms of data, including transcriptions of printed text translations of sentences from one language to another and image annotations.

• **Evaluating and debugging models:** Crowdsourcing is also commonly used to evaluate or debug models, including unsupervised learning models, which can be difficult to evaluate objectively since there is often no clear notion of ground truth. Along these lines, we discuss the use of crowdsourcing to evaluate the coherence of topic models, generative models used to discover and explore the thematic topics discussed in a set of documents. We also explore ways in which crowdsourcing has been used to evaluate the interpretability of explanations of predictions in supervised learning settings and to debug the components of a pipeline in a complex computer vision system

• **Hybrid intelligence systems**: Hybrid intelligence or "human-in-the-loop" systems advance the capabilities of current AI technology by leveraging the complementary strengths of humans and machines. We explore several compelling examples of hybrid systems that suggest their great potential: hybrid systems for clustering data points. These systems are able to achieve more than would be possible with state-of-the-art machine learning or AI systems alone because they can make use of people's common sense knowledge, life experience, subjective beliefs, and flexible reasoning skills.

• **Behavioral studies to inform machine learning research:** As machine learning becomes a larger part of people's everyday lives, interest in understanding how real people interact with machine learning systems continues to grow. There has been a surge of recent research on questions like how to design machine learning models that are or how to understand the ways that algorithmic decisions impact people's lives. These questions are interdisciplinary in nature and require gaining a better understanding of the underlying principles behind humans' interactions with machine learning and other technological systems. At the same time, psychologists and social scientists have started using crowdsourcing platforms as a fast and easy way to gain access to large pools of subjects for behavioral experiments. This presents a natural opportunity for researchers to conduct studies on crowdsourcing platforms that improve our understanding of how humans interact with technology broadly and machine learning algorithms in particular. Rather than evaluating the way in which people interact with one particular algorithm, this line of research aims to develop an understanding of components of human behavior.

1. Data Generation

Perhaps the most common application of crowdsourcing within the machine learning community is data generation. We first describe techniques for crowdsourcing binary or categorical labels, reviewing the literature on how to improve label quality through redundancy and incentives. We then discuss several examples of ways in which crowdsourcing has been used to generate more complex forms of data, such as image annotations and translations of text.

1.1 Generating Binary or Categorical Labels

Start with the setting in which crowdworkers are presented with unlabeled data instances (for instance, websites) and are asked to supply labels (for instance, a binary label indicating whether or not the website contains profanity). The main challenge arises from the fact that the supplied labels are often noisy or inaccurate, either because workers are imperfect or because workers are unmotivated to put high effort into the labeling task. There are two primary lines of work aimed at improving the quality of crowdsourced labels. The first assumes that each instance is presented to multiple crowdworkers and explores algorithmic techniques for aggregating the workers' responses.

1.2 Generating Transcriptions, Translations, and Image Annotations:

Crowdsourcing is also used to generate more complex and free-form labels, such as transcriptions, translations of language, or image annotations. Perhaps the best-known example of a crowdsourcing system for transcription is re-CAPTCHAs (or Completely Automated Public Turing tests to tell Computers and Humans Apart) are security tools designed to prevent bots from accessing online services. People attempting to access a website or create an account are asked to perform a task that is difficult for computers toperform but that humans find easy, such as reading and transcribing distorted characters. CAPTCHAs are used to stop ticket scalpers from using bots to buy out popular shows and to prevent spammers from opening arbitrary numbers of email accounts. On Ahn et al. (2008) found a way to put the vast quantities of human effort exerted on solving CAPTCHAs to use, harnessing this effort to digitize old books that current optical character recognition (OCR) systems were unable to handle. Their reCAPTCHA system presents two images of words, both taken from scanned text on which state-of-the art OCR systems have failed. One of these images is a gold-standard data point for whichthe correct transcription is already known. This is the image used to test whether or not the transcriber is human. The true label of the second image is unknown. By completing the CAPTCHA, the human is essentially entering a label for this data. Since reCAPTCHA was acquired by Google, similar techniques have been used to annotate images and build other large-scale machine learning data sets. The crowd-generated translations collected by this and other systems can then be used as training data for machine translation tasks. Other data generation and labeling tasks appropriate for crowdsourcing include object classification, attribute (or feature) generation, and image segmentation.

2. Evaluating and Debugging Models

Aside from data generation, the most common use of crowdsourcing within the machine learning community is to evaluate or debug models. Crowdsourced evaluation is especially common for unsupervised models, which generally cannot be evaluated in terms of simple metrics like accuracy or precision because there is no objective notion of ground truth. More recently, crowdsourcing has been used to evaluate human-centric properties of supervised models, such as model interpretability. In this section, we review several examples of applications of crowdsourcing to model evaluation and debugging. This list of examples is not intended to be exhaustive, but to give a flavor of different ways in which crowdsourcing can be used in this context.

2.1 Evaluating Unsupervised Models

It is increasingly common to see crowdsourcing used to evaluate unsupervised models, such as topic models. Topic models are widely used to discover thematic topics from a set of Documents. Topic models are often used for data exploration and summarization, especially in the social sciences (Boyd-Graber et al., 2017). In order to be useful in these contexts, the inferred topics must be meaningful to end users. For example, if the set of words that appear with high weight in an individual topic are not coherent, the topic will not be useful to end users trying to understand the content of their documents. However, "meaningfulness "is hard to measure analytically, leading many researchers to instead evaluate topic models in terms of easier to quantify criteria, such as predictive power. The researchers designed a word intrusion task in which a crowdworkers is presented with a randomly ordered list of the most common words from a topic.

3.2 Evaluating Model Interpretability

In supervised learning, models are often evaluated in terms of objective performance metrics such as accuracy, precision, or recall. However, even if a model performs well in terms of these criteria, users may hesitate to rely on the model if they do not understand the model's predictions, especially in critical domains like health or criminal justice. Because of this, there is now wide interest in developing models that are human-interpretable. Because of the subjectiveand inherently human-centric nature of interpretability, it is natural to use crowdsourcingto evaluate the interpretability of models.

3.2 Evaluating Model Interpretability

In supervised learning, models are often evaluated in terms of objective performance metricssuch as accuracy, precision, or recall. However, even if a model performs well in terms ofthese criteria, users may hesitate to rely on the model if they do not understand the model'spredictions, especially in critical domains like health or criminal justice. Because of this, there is now wide interest in developing models that are human-interpretable. Because of the subjectiveand inherently human-centric nature of interpretability, it is natural to use crowdsourcingto evaluate the interpretability of models.

3.3 Debugging Components of a Pipeline

In fields like computer vision, speech recognition, translation, and natural language processing, systems often consist of several discrete components linked together to perform a complex task. For example, consider the problem of semantic segmentation, which involves partitioning an image into semantically meaningful parts and labeling each part with a class.

There are promising approaches to this problem that use machine learning models such asconditional random fields (CRFs) to integrate feedback from independent components that perform various scene understanding tasks like object detection, scene recognition, and seg-mentation. If a system designer wants to improve performance, it is not always clear whichcomponent to focus attention on.

4. Hybrid Intelligence Systems

Despite the current hype around AI and the great technological advances that have beenmade in recent years, AI systems are still far from perfect. In some cases, AI systems canbenefit by involving humans in the loop to perform tasks that rely on life experience, judgment, or domain knowledge. Such hybrid intelligence systems can leveragethe complementary strengths of humans and machines to accomplish more than would bepossible using humans or machines alone.Hybrid systems have been designed to perform tasks from grading students' work

4.1 Hybrid Clustering

Hybrid intelligence systems can be put to use to solve traditional machine learning problems like clustering in scenarios in which data points are easier for humans to understand and categorize than they are for machines. Many hybrid clustering techniques have been proposed approaches vary in terms of the types of queries given to human judges, the algorithms used to aggregate their responses, and whether or not additional features of each object are available to the algorithm. Some researchers focus primarily on the entity resolution setting, in which the goal is to cluster together objects that refer to the same entity

4.2 Hybrid Speech Recognition

Quickly and reliably converting speech to text requires a level of contextual understanding beyond the capabilities of machines and a level of speed beyond the capabilities of most humans. Closed captioning systems that rely on automatic speech recognition work very well under ideal circumstances (for example, when the voice recording is high quality and the system has been trained on data from the particular speaker), but can fail when presented with low quality audio, speakers with novel accents, or language with technical jargon that falls outside the vocabulary on which the system was trained. In these scenarios, professional stenographers produce the best results, but high-quality stenographers can be prohibitively expensive and are not available on demand.

As soon as a user starts recording, the recorded audio is sent simultaneously to several crowdworkers. These workers are not expected to fullytranscribe the speech, which is generally not possible without a specialized stenotype keyboard. Instead, each worker transcribes sentence fragments. Scribe adjusts the speed andvolume of the speech adaptively for each worker in order to focus workers' attention on distinct, overlapping components. The crowdworker views the transcription output by a speech recognition system whilelistening to the corresponding audio. She then types out corrections to mistakes in the transcription as she notices them. These corrections are automatically incorporated into theappropriate spot in the transcription.

As new breakthroughs continue to improve automatic speech recognition (Yu and Deng,2014), this hybrid approach may eventually become unnecessary. However, the ability for speech recognition systems to achieve true human parity under nonideal conditions likely remains far in the future (Bigham, 2017). This example shows that crowdsourcing can be an effective way of compensating for a lack of sufficient machine learning or AI solutionsuntil a time when the technology improves.

4.3 Hybrid Scheduling

Several researchers have explored the potential use of hybrid intelligence systems to solve complex tasks with global constraints or consistency requirements. Examples of such tasksinclude itinerary planning, taxonomy creation; Chilton et al., 2013), and writing.

The problem of scheduling conference sessions can be viewed as a constrained optimization problem in which the solver has no direct access to the constraints. The goal of conference organizers is to group similar talks together in sessions while minimizing conflicts between talks that are scheduled at the same time, but conference organizers generally do notknow which sets of talks attendees want to see.

4.4 Hybrid Forecasting

Significant resources are devoted to producing forecasts about geopolitical events and economic indicators. Humans are flexible in their ability to reason about arbitrary events, but human forecasts can be limited by cognitive biases or the inability to digest and process in-formation at scale. Statistical and data-driven models, on the other hand, are able to takeadvantage of vast quantities of available data, but are difficult to design and train for one-of-a-kind events. Hybrid forecasting systems aim to combine the computational power ofmachines with the flexibility of humans to produce accurate forecasts.In the most basic hybrid forecasting systems, algorithmic techniques are used primarilyas a way to elicit and aggregate human-generated forecasts. One common example is aprediction market, a financial market inwhich traders can buy or sell securities with payoffs that are linked to future events. For example, in an election market, traders might buy or sell a security that is worth $1 if theincumbent candidate wins and nothing otherwise. If a trader believes that the probability ofthis candidate winning is p and wants to maximize her expected payoff, then she should bewilling to buy this security at any price less than $p, since with probability p she would get$1. Similarly, she should be willing to sell at any price greater than $p. For this reason, wecan think of the current market price of this security as capturing traders' collective beliefsabout how likely it is that the incumbent will win. Prediction markets can be operated ascontinuous double auctions, much like the stock market, requiring very little algorithmicingenuity. However, when the level of trade is low, there can be advantages to operatingprediction markets using algorithmic market makers that automatically set prices based onthe history of trade. Prediction markets haverecently gained more attention in the machine learning community due to the discovery ofstrong mathematical connections between these algorithmic market makers and no-regretonline learning algorithms.

4.5 Hybrid Intelligence Systems in Industry

So far, we have focused on hybrid intelligence systems that have come out of the researchcommunity, but it is worth mentioning that human-in-the-loop systems are widely used in industry as well.

5. Behavioral Studies to Inform Machine Learning Research

Within the past few years, there has been an increased interest, both from within anoutside the machine learning community, in understanding how humans interact with machine learning systems. Researchers are striving to make machine learning models humaninterpretable. Psychologists and social scientists have increasinglyturned to crowdsourcing platforms to run behavioral experiments that traditionally wouldhave been conducted on undergraduates in a physical lab. The developments open up the opportunity for interdisciplinary research that uses behavioral experiments conducted on crowdsourcing platforms to improve our understandingof how humans interact with machine learning and AI systems. This line of work goes beyond the idea of using crowdsourcing to evaluate one particular machine learning model user interface, and instead seeks a general understanding of thecomponents of human behavior that could inform the use and design of machine learning systems more broadly. Such experiments can help us develop better models of human behaviorthat could be used, in turn, to develop better algorithms and interfaces. It is worth noting that behavioral experiments and other crowdsourced user studiescould benefit other subfields of computer science as well.

5.1 Understanding Trust in Predictive Models

There is a large body of work cutting across psychology, management, and other researchcommunities that studies human trust in algorithmic predictions and models. This body of work provides invaluable insights about how real end users interact withmachine learning systems in practice that can be put to use immediately in the design ofmodels and user interfaces.

This has immediate implications about which strategies might help gain user trust in machine learning models. In particular, users might be more willing to trust amodel if they have the ability to intervene. Even if this human intervention leads to a worseprediction, allowing the intervention may still be beneficial because the user will be morelikely to use the model in the first place.It is natural to imagine that similar ideas could be applied to study model interpretability.

6. Understanding the Crowd

In the previous section, we argued that crowdsourced studies of human behavior can bevaluable for understanding how lay people interact with machine learning systems. In this section, argue that such studies are also useful for understanding the behavior of thecrowd itself. This understanding helps us better model the crowd and allows us to defineconcrete recommendations of best practices that can be put to use whether using the crowdfor data generation, model evaluation, hybrid intelligence systems, behavioral research, orany other purpose.

The studies described in this section help us understand how real crowdworkers respondto incentives, yielding immediately applicable guidelines for setting payments in additionto more accurate ways of modeling crowd behavior in theoretical work on incentive design.They tell us how to most effectively gamify crowdworker and provide other sources of intrinsic motivation for workers. They help us get a grip on the question of how widespread dis-honesty is on crowdsourcing platforms, and how dishonest behavior can be mitigated. Andthey show us that crowdworkers are not independent and isolated workers, but have a rich social network.

6.1 Crowdworker Demographics

Over the years there have been several studies published that examine the demographics ofworkers on Mechanical Turk.That said, these demographics are more or lessin line with those reported. One of the first questions that many researchers have when they decide to incorporate crowdsourcing into their work is how much to pay per task. When crowdsourcing first began to gain popularity among researchers, part of the appeal was the ability to generate data or run experiments cheaply A natural question is whether paying higher wages increases the quality of work. There is evidence that, at least in some scenarios, the answer is no. Several behavioral studies examining the impact of payments on quality found that setting higher payments increased the quantity of work that crowdworkers were willing to do, but not the

10.4 SUMMARY

Crowdsourcing can be utilized to moderate online user generated content in multiple ways. First, crowd moderation could be implemented as a standalone moderation mechanism. Second, crowd workers can contribute to create accurate and rich training labels for supervised learning based automated approaches. Third, by combining crowdsourcing and automated approaches, a human-in-the-loop moderation system could be implemented. In such a system, the content will be reviewed by crowd workers, when automated approaches fail to make a confident judgement. By analysis of participant comments reveals interesting insights that can help create an effective crowdsourcing mechanism for content moderation. We note participants often labelled unsure whenthey found it difficult to decide on borderline content. This could beimproved by providing specific guidelines or illustrative examplesfor participants. Many crowdsourcing studies have highlighted theimportance in providing clear instructions to obtain high qualitydata. Further, the use of comparative judgements inplace of discrete labels can lead to more reliable, also notice that individual opinion couldhave a considerable impact on the moderation process as certain workers lean more towards labellingcontent either as appropriate or inappropriate compared to the rest. For multiple labels for each item, the influence of personal opinion is expected to be reduced when using crowdsourcing.However, this is also valid for dedicated moderation teams and canhave broader implications when only one-person reviews content.Therefore, it is important to set out clear policies, educate and traincommercial content moderators to limit the influence of personal.

10.5 GLOSSARY

- Crowdsourcing- The practice of obtaining information or input into a task or project by enlisting the services of a large number of people, either paid or unpaid, typically via the internet.
- Data generation- Data generation refers to the theory and methods used by researchers to create data from a sampled data source in a qualitative study
- Model Evaluation- Model Evaluation is the subsidiary part of the model development process.
- Hybrid Intelligence- Hybrid Intelligence combines human and Artificial Intelligence to collectively achieve superior results and learn from each other.
- Behavioral Experiment- Behavioral experiments are an information gathering exercise, the purpose of which is to test the accuracy of an individual's beliefs (about themselves, others, and the world) or to test new, more adaptive beliefs.

- Incentives- A thing that motivates or encourages someone to do something.

10.6 ANSWER TO CHECK YOUR PROGRESS

1. What are categories of 'big data' applications in examining urban landscape?
2. Data set about cellphone usage in an urban area?

10.7 REFERENCES

- Jacob Abernethy, Yiling Chen, and Jennifer Wortman Vaughan. Efficient market makingvia convex optimization, and a connection to online learning.
- ACM Transactions on Economics and Computation, 1(2): Article 12, 2013.Ittai Abraham, Omar Alonso, Vasilis Kandylas, Rajesh Patel, Steven Shelford, and Alek-sandrsSlivkins.
- In SIGIR, 2016. Arpit Agarwal, Debmalya Mandal, David C. Parkes, and Nisarg Shah. Peer prediction with heterogeneous users. In ACM EC, 2017. Omar Alonso. Implementing crowdsourcing-based relevance experimentation: An industrial perspective.
- Information Retrieval, 16(2):101–120, 2013.Omar Alonso, Daniel E. Rose, and Benjamin Stewart. Crowdsourcing for relevance evaluation. ACM SigIR Forum, 42(2):9–15, 2008.
- Vamshi Ambati, Stephan Vogel, and Jaime Carbonell. Collaborative workflow for crowdsourcing translation. In CSCW, 2012.
- Paul Andr´e, Haoqi Zhang, Juho Kim, Lydia B. Chilton, Steven P. Dow, and Robert C.
- Miller. Community clustering: Leveraging an academic crowd to form coherent conferencesessions. In HCOMP, 2013.Pavel D. Atanasov, Phillip Rescober, Eric Stone, Samuel A. Swift, Emile Servan-Schreiber,Philip E. Tetlock, Lyle Ungar, and Barbara Mellers. Distilling the wisdom of crowds: Prediction markets versus prediction polls. Management Science, 63(3):691–706, 2017.Bahadir Ismail Aydin, Yavuz Selim Yilmaz, Yaliang Li, Qi Li, Jing Gao, and Murat Demir-bas. Crowdsourcing for multiple-choice question answering. In AAAI, 2014.Solon Barocas and Andrew Selbst. Big data’s disparate impact. California Law Review,104, 2016.
- Jonathan Baron, Barbara A. Mellers, Philip E. Tetlock, Eric Stone, and Lyle H. Ungar.
- Two reasons to make aggregated probability forecasts more extreme. Decision Analysis,11(2):133–145, 2014.

- Joyce Berg, Robert Forsythe, Forrest Nelson, and Thomas Rietz. Results from a dozenyears of election futures markets research. Handbook of experimental economics results,1:742–751, 2008.
- Michael Bernstein, Greg Little, Rob Miller, Bjoern Hartmann, Mark Ackerman, David Karger, David Crowell, and Katrina Panovich. Soylent: A word processor with a crowd in side. In UIST, 2010.
- Anant Bhardwaj, Juho Kim, Steven P. Dow, David Karger, Sam Madden, Robert C. Miller, and Haoqi Zhang. Attendee-sourcing: Exploring the design space of community-informed conference scheduling. In HCOMP, 2014.
- Jeffrey P. Bigham. Reaching dubious parity with hamstrung humans. Blog post accessed at http://jeffreybigham.com/blog/2017/reaching-dubious-parity-with-hamstrung-humans.html, 2017.
- David M. Blei and John D. Lafferty. Topic models. Text mining: Classification, clustering, and applications, 10(71):34, 2009.
- David M. Blei, Andrew Y. Ng, , and Michael I. Jordan. Latent Dirichlet allocation. Journal of Machine Learning Research, 3(Jan):993–1022, 2003.
- Jordan Boyd-Graber, Yuening Hu, and David Mimno. Applications of topic models. Foundations and Trends in Information Retrieval, 11(2–3):143–296, 2017.
- Jonathan Bragg, Mausam, and Daniel S. Weld. Crowdsourcing multi-label classification fortaxonomy creation. In HCOMP, 2013.
- Michael Brooks, Saleema Amershi, Bongshin Lee, Steven Drucker, Ashish Kapoor, and Patrice Simard. FeatureInsight: Visual support for error-driven feature ideation in text classification. In IEEE VAST, 2015.Alexandra Chouldechova. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. Big Data, Special Issue on Social and Technical Trade-Offs2017.
- Gianluca Demartini, Djellel Eddine Difallah, and Philippe Cudr´e-Mauroux. Zencrowd: Leveraging probabilistic reasoning and crowdsourcing techniques for large-scale entity linking.
- In WWW, 2012. Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. ImageNet: A large-scale hierarchical image database. In CVPR, 2009.
- Berkeley J. Dietvorst, Joseph P. Simmons, and Cade Massey. Algorithm aversion: Peopleerroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General, 144(1):114, 2015.Berkeley J. Dietvorst, Joseph P. Simmons, and Cade Massey. Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. Management Science, 2016.
- Djellel Eddine Difallah, Michele Catasta, Gianluca Demartini, Panagiotis G. Ipeirotis, andPhilippe Cudr´e-Mauroux. The dynamics of micro-task crowdsourcing: The case of Amazon MTurk. In WWW, 2015.
- Dominic DiPalantino and Milan Vojnovic. Crowdsourcing and all-pay auctions. In ACMEC, 2009.
- Finale Doshi-Velez and Been Kim. Towards a rigorous science of interpretable machinelearning. CoRR arXiv:1702.08608, 2017.
- Mary T. Dzindolet, Linda G. Pierce, Hall P. Beck, and Lloyd A. Dawe. The perceived

utility of human and automated aids in a visual detection task. Human Factors, 44(1):79–94, 2002.

- Robert C. Edgar and Serafim Batzoglou. Multiple sequence alignment. Current opinion instructural biology, 16(3):368–373, 2006.
- Ju Fan, Guoliang Li, Beng Chin Ooi, Kian-lee Tan, and Jianhua Feng. icrowd: An adaptivecrowdsourcing framework. In SIGMOD, 2015.

10.8 TERMINAL QUESTIONS

1. Is it possible to illustrate crowdsourcing as a graph? and is it possible to monitor activities of members of this graph?
2. How can we build labeled data set as training data by crowdsourcing?
3. What are standard performance metrics or index for "Crowdsourcing data" systems?
4. What tests (if any) can be used on convenience sample survey data so that the results are able to inform experimental design later on?

UNIT 11 : UPLOADING AND VISUALIZATION OF CROWD SOURCING DATASETS LIKE BHUVAN PORTAL

11.1 OBJECTIVES

After reading this unit the student will understand:

- The data access and visualization in Geoportals specially ISRO Bhuvan;
- Popular Geo-portals and their pros & cons; and
- Crowdsourcing data uploading and access from Geoportals.

11.2 INTRODUCTION

Advancements in Information and Communication Technologies (ICT) has facilitated a new way for sharing and dissemination of Geospatial data and information. The online data repositories and web applications are providing various means of data access by using internet and related technologies. Today, the users of geospatial data and information can use internet platform for various geoscientific activities such as spatial queries, geo-visualization and simple to complex computations for decision-making and virtual reality. Considering the importance of spatial data for humanitarian response during natural disasters many agencies and individuals are hosting their data sets online, which has enhanced the outreach of geospatial data many folds. These open geo-data sets can be used for various thematic applications either as a geo-web service or as a data product(s).

The open online data repositories and geo-web services are providing data and information by using web service standards published by Open Geospatial Consortium (OGC). The websites are available either as geo-portal or online data archive. The web portals are dynamic web applications which serves data and information to its user (s) by using database server technology. In web portal applications, the data is accessed and processed using an additional middle tier at server end using any web programming language such as PHP, C#, JAVA, Python, etc. This middle tier is also known as business logic or application server. In case of GIS data, the database servers hosts the raster and vector data in addition to attribute data sets in a database server. The geo-portals are developed based on GIS servers which are also known as Map server (s). The GIS servers typically act as middle tier in software application architecture to make geospatial data compatible with internet client software applications such web browsers.

The most of online data repositories are providing free and open geo-data in public domain. Some of the popular geoportals and online data repositories are discussed in this chapter. One of the important source of data in GIS is crowdsourcing or public participation. The popular geo-portals such as Open Street Map (OSM) and ISRO Bhuvan has successfully demonstrated the power of public participation in GIS data creation and sharing. Today, you can download large scale maps in original GIS format of entire globe from OSM website. OSM also provides online mapping utility to its users in web platforms where many individuals can be engaged in mapping activities simultaneously so that the task is completed in very short time. The ISRO Bhuvan geoportal provides various mobile apps and online mapping utilities to create data repository of ground data. In India, the Bhuvan geoportal is used by many government departments and ministries to generate the geo-tagged inventory of their assets. In this chapter some of the examples are discussed.

11.3 UPLOADING AND VISUALIZATION OF CROWD SOURCING DATASETS LIKE BHUVAN PORTAL

Google Maps and Google Earth

The mapping applications (https://maps.google.com) from Google (www.google.com) are one of the most popular GIS application available in public domain. The map data from google can be accessed through variety of ways like as web application in web browser, Google Earth as standalone desktop application, mobile application and Application Programing Interface (API) for custom web mapping. The mapping applications from google provides satellite imagery, street maps, and Street View perspectives, as well as many GIS tools such as a route planner for traveling by foot, car, bicycle, or with public transportation, locator for urban businesses and other organizations in numerous countries around the world. In some cities, Google Maps offers street views comprising photographs taken from moving vehicles in different direction.

The maps and satellite images available in google mapping applications are not up-to-date or real time data. However the maps are being updated on frequent basis by using public participation in data creation. The satellite imageries are available at various spatial resolution which allows geo-visualization at street level with spatial resolution of less than a meter. Google Maps uses a close variant of the Mercator projection, and therefore cannot accurately show areas around the poles. A related product is Google Earth, a stand-alone program which offers more globe-viewing features, including showing polar areas. Google Maps for mobile is the world's most popular app for smartphones, with over 54% of global smartphone owners using it at least once during the month of August 2013 (Business Insider, 2013). Some of the popular services available in google maps are:

- A route planner offers directions for drivers, bikers, walkers, and users of public transportation who want to take a trip from one specific location to another. This route planner application is one of the popular mapping application used by majority of google users across the globe.
- The Google Maps application program interface (API) makes it possible for Web site developers and administrators to embed Google Maps into their websites. The mapping applications developed using google map API allows overlay of site specific data i.e. user's data on google maps. The specific and custom mapping application can be build using these API.As per google developers website (googlegeodevelopers.google.com, 2013) approximately 1,000,000web sites using Google Maps API which makes it the most heavily used web application development API (www.programmableweb.com, 2013).The Google Maps API is free for commercial use, provided that the site on which it is being used is publicly accessible and does not charge for access, and is not generating more than 25 000 map accesses a day (Google Maps API FAQ, 2014). Web Sites that do not meet these requirements can purchase the Google Maps API for Business (Google Maps API for Business, 2014).
- The map services of Google for mobile user offers location-based services using Global Positioning System (GPS) of mobile devices and data from wireless and cellular networks.
- Google Street View enables users to view and navigate through horizontal and vertical panoramic street level images of various cities around the world.
- In addition to the above services google maps also offer images of the moon, Mars, and the heavens for hobby astronomers. These services are available as additional components of google map service.

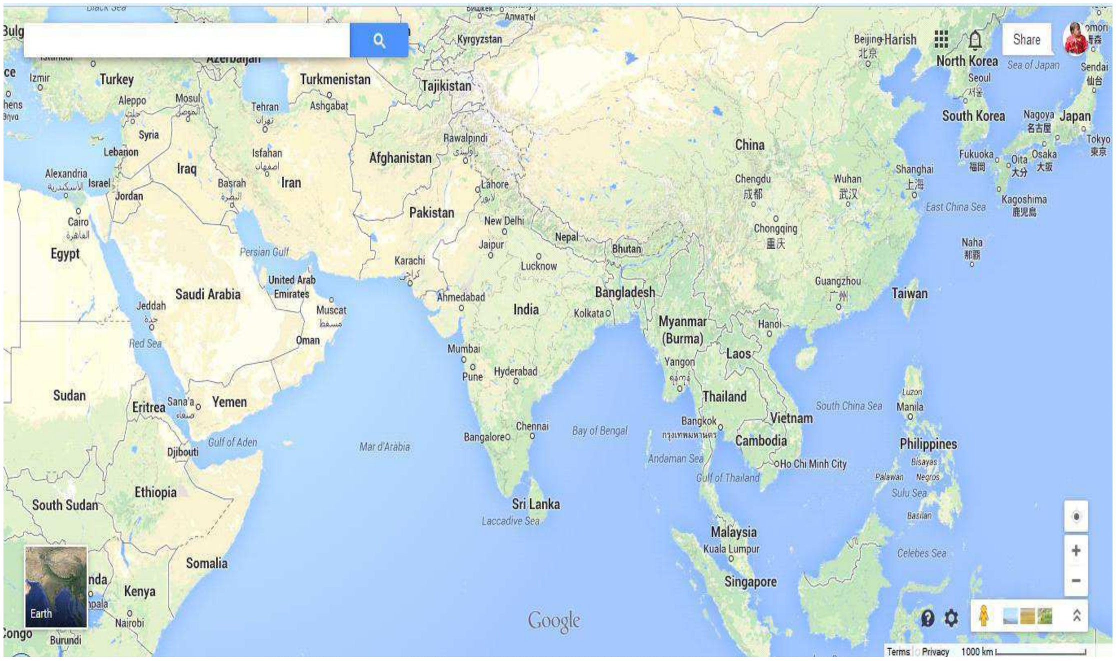


Figure 11.1- Map view of http://maps.google.com

The development and implementation of google map application is based on web 2.0 technology where AJAX based architecture is implemented for interactive mapping applications. The google map application uses JavaScript extensively. As the user drags the map, the grid squares are downloaded from the server and inserted into the page (Gautham, 2012). When a user searches for a business, the results are downloaded in the background for insertion into the side panel and map; the page is not reloaded. Locations are drawn dynamically by positioning a red pin (composed of several partially transparent PNGs) on top of the map images. A hidden IFrame with form submission is used because it preserves browser history. The google map website also uses JSON for data transfer rather than XML, for performance reasons. Google maps has also introduced Maps GL for 3D rendering of geospatial data using WebGL. The WebGL based mapping applications allows 3D visualization of geo-spatial data using java script without having any additional plug-in or software components at client end.

Bing Maps

Bing map is a web-based mapping service provided by Microsoft. Bing is a search service from Microsoft, which includes the 'Bing Maps' platform which was previously known as Microsoft Virtual Earth. The Bing map platform includes map tiles, map embedding APIs, routing, and many more map applications. It makes use of proprietary datasets, often licensed from third party geo-data providers, and its use is therefore bound by copyright restrictions. The Bing map provide global coverage of maps and satellite data for geo-visualization and development of custom web mapping applications.

This mapping service provides browse and search the locations in topographically-shaded street maps for majority of cities worldwide. The maps include certain Points of Interest (POI) such as metro stations, stadiums, hospitals, and other amenities. It also provides facility to locate and add new POI by the users. Searches can cover public collections, businesses or types of business, locations, or people. Bing map provides five primary types of street map views:

- **Road View**: Road view is the default map viewer which allows display of vector maps of roads, buildings, amenities, and geography features. The data from which the default road map is rendered is licensed from NAVTEQ which is an American Chicago-based provider of GIS data and a major provider of base electronic navigable maps. In addition to this the road data from OpenStreetMap is also allowed to overlay and visualized.
- **Aerial View:** Aerial view is a map viewer provided by Bing to allow overlay of satellite imagery onto the map and highlights roads and major landmarks for easy identification. The satellite data of different spatial resolution are available which can be used for identification of geographical features at street level. The satellite images of Bing aerial view are also used by OpenStreetMap for its mapping applications.
- **Bird's Eye View:** Bird's-eye view displays aerial imagery captured from low-flying aircraft (http://www.bing.com). According to Bing Maps Imagery Release in April 2010 the Bird's-eye images are taken at an oblique, 45-degree angle. They show the sides of buildings, not just the roofs, and give better depth perception for geography. Bird's-eye view is available in selected locations across the globe (Bing Maps Imagery Release, 2010). Bird's-Eye images for a location may be viewed from all four cardinal directions.
- **StreetSide View**: StreetSide of Bing provides 360-degree photos of street-level scenes taken from special cameras mounted on moving vehicles. It is similar to Google Streetview but have less number cities covered for 360-degree visualizations.
- **3D View:** The 3D View of Bing provides ability to rotate and tilt the angle in addition to panning and zooming in satellite imagery and map data. To attempt to achieve near-photorealism, the 3D buildings available in this view are textured using composites of aerial photography and ground photograph. To view the 3D maps, users must install a plugin

available freely in the website, then enable the "3D" option on "Bing Maps". In addition to exploring the maps using a mouse and keyboard, it is possible to navigate the 3D environment using an Xbox 360 controller or another game controller in Windows operating system.

Another important initiative taken by Microsoft under Bing Ma is generation of Global Ortho data. In July 2010, Microsoft and Digital Globe, a leading global content provider of high-resolution earth imagery solutions, announced the collection of the first imagery from the company's Advanced Ortho Aerial Program. Through a special agreement with Microsoft, the Advanced Ortho Aerial Program will provide wall-to-wall 30 cm aerial coverage for entire globe.

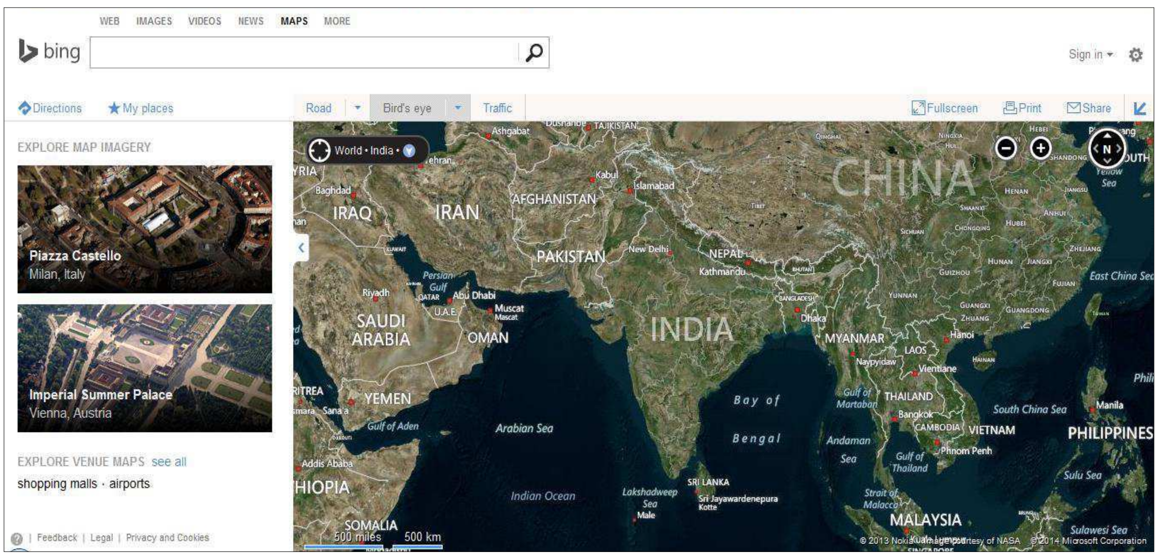


Figure 11.2- Bird eye view from Bing- http://www.bing.com/maps/?mkt=en-in)

Open street Map- An approach using crowdsourcing

OpenStreetMap (OSM) is a collaborative project to create a free editable map of the world using concept of crowdsourcing. The project was initiated by Steve Coast of UK in 2004. Since then, it has grown to over more than one million registered users who can collect data using GPS devices, aerial photography, and other free sources (Neis et al, 2012). The data generated through public participation is made available under the Open Database License. The site is supported by the OpenStreetMap Foundation, a non-profit organization registered in England.

In the initial phase the map data were originally collected from scratch by volunteers performing systematic ground surveys using a handheld GPS unit and a notebook, digital camera, or a voice recorder. The data were then entered into the OpenStreetMap database. Now OSM provides variety of tools and mode to create and update map data directly through internet by the registered users. Editing of maps can be done using the default web browser editor called iD which a HTML5 application written by MapBox. Another OSM editor which is Flash-based application and known as Potlatch is also available for web users. The desktop version of map creator and editor are known as JOSM and Merkaartor which are more powerful desktop editing applications and are better suited for advanced users.

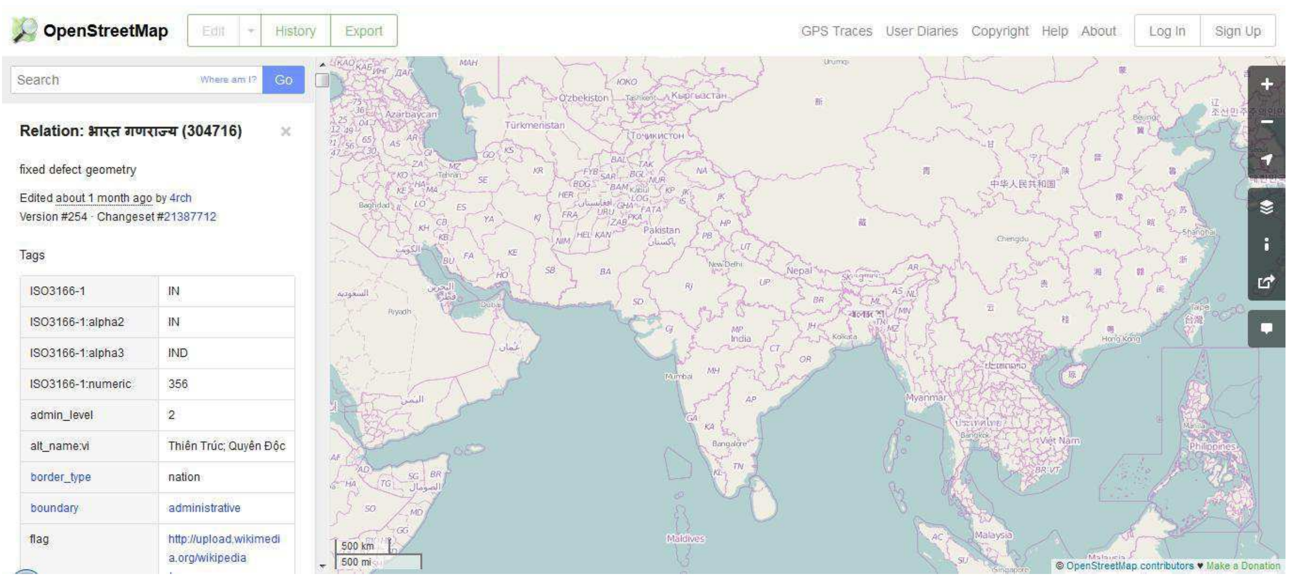


Figure 11.3- Home page of OpenStreetMap- http://www.openstreetmap.org

The satellite imagery from Bing Maps is available freely as background image for all the mapping applications of OSM by using which the geographic features are identified by the map creator.

The online mapping platform using concept of crowdsourcing or public participation is a key element of Open Street map. Any user can create his/her account in the website and can start contributing the map creation as shown in Figure 4. In the online map editing tool, the three basic geometry types are provided to create the vector map. The spatial data structure is designed to create various map features and their attributes. The online mapping application provides the mapping utilities in multi-user environment so that many users can do mapping simultaneously. This participatory approach seems to be one of the best approaches of mapping at large scale.

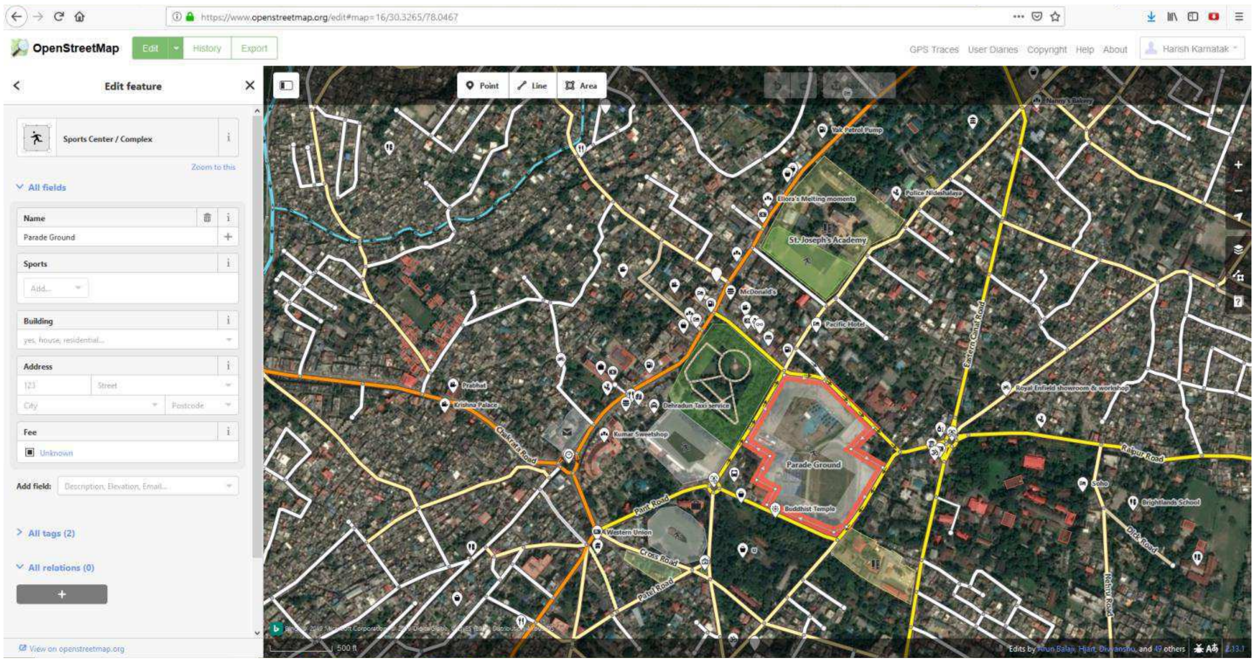


Figure 11.4- Map creation/editing for Dehradun city.

The OSM has designed a data model for different map features. The maps can be created for these predefined data models. For detail structure of OSM map feature please visit- http://wiki.openstreetmap.org/wiki/Map_Features .

The large-scale maps can be downloaded from OSM website as .osm file which is a XML document and can be converted as vector data by using any third party software product like QGIS. Some of open service providers such as https://download.geofabrik.de/ are providing cloud-based services to download the OSM data.

ArcGIS Online

ArcGIS online (www.esri.com/software/arcgis/arcgisonline) is a mapping service from ESRI. According to ESRI website, it is a collaborative, cloud-based platform that allows members of an organization to use, create, and share maps, apps, and data, including authoritative base maps published by ESRI. In ArcGIS online portal the user can use and create maps, access GIS layers and tools, publish data as web service, collaborate and share, access maps from any device, create spatial data using attribute data available in MS Excel, customize the ArcGIS Online website, and view status reports of their data. The ArcGIS Online can also be used as a platform to build custom location-based applications for mobile devices.

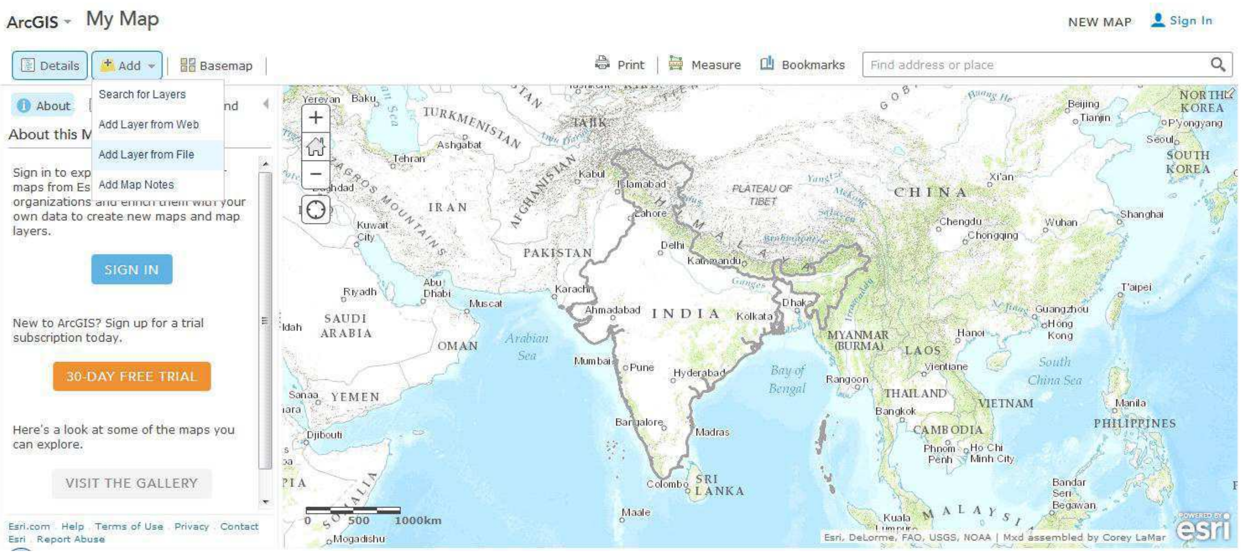


Figure 11. 5- Map viewer of ArcGIS Online- http://www.arcgis.com

This mapping service is commercial; however, the trial access for 30 days is available free of cost for any users. The users can access ArcGIS Online through web browsers, mobile devices, and desktop map viewers, as well as directly through other components of the ArcGIS system, for example, the web APIs and ArcGIS for Desktop.

GIS Cloud

GIS Cloud (www. giscloud.com) is the web-based GIS developed based on cloud computing. It provides features and capabilities of desktop GIS through internet. GIS Cloud offers easy and efficient visualization, analysis and exploration of geographic information. The primary goals of the GIS Cloud platform are as follows; to simplify exchange of geographical information between users and offer an easy way to analyze this information regardless of the location of its users (www.giscloud.com). By using this mapping service, the users can access the capabilities of desktop GIS such as geospatial analysis, spatial intelligence, the creation of customized mapping reports and publishing geographic analysis on the Web. The main differentiating characteristic of GIS cloud website is its unique vector visualization engine and vector-based data analysis in web browser environment. It also supports wide range of vector and raster file and database formats.

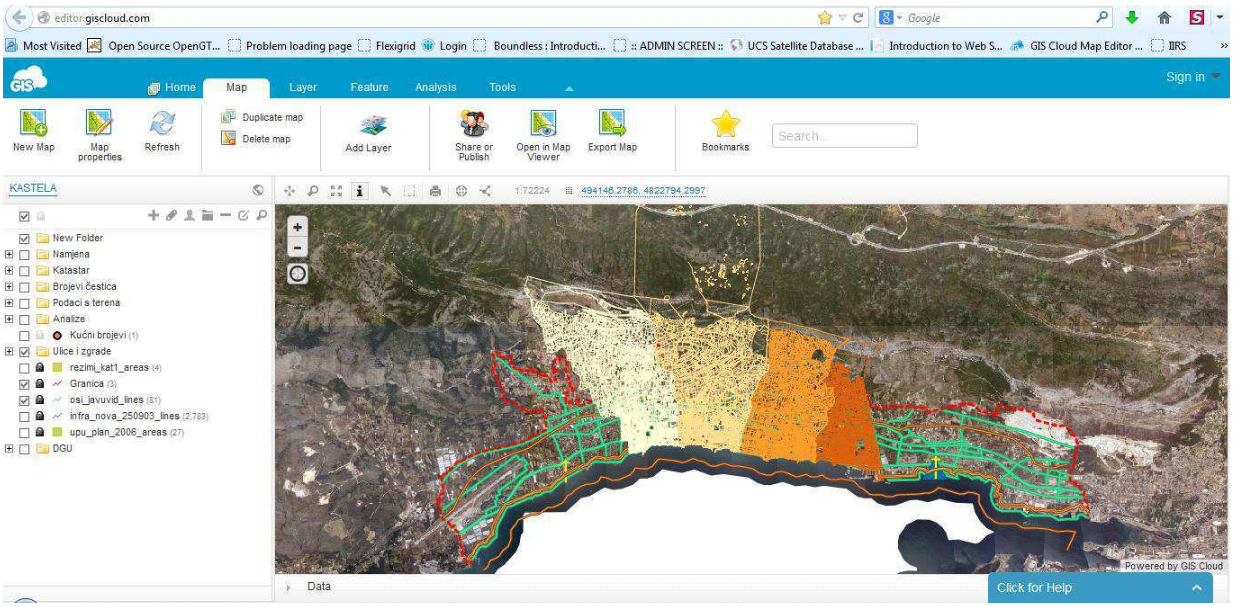


Figure 11.6- GIS cloud environment- http://www.giscloud.com

The Graphical User Interface (GUI) provides various tools and functionalities for GIS operations as shown in Figure 11.6. The various tabs are available like ‘Map’, ‘Layer’, ‘feature’, ‘Analysis’, and ‘Tools’. The GUI looks like a standard desktop-based GIS software. In addition to GIS data creation, analysis and processing, the GIS cloud also provide various utilities for location-based services and mobile GIS. Mobile GIS based data collection system is freely available to the users however many other features are available under commercial mode. This mapping service also provide free trial version for 30 days for all the features of GIS cloud.

The complete operating manual of GIS cloud is available at http://www.giscloud.com/docs/GIS_Cloud_User_Manual.pdf.

Indian Geoportals and Geo-data repositories:

ISRO Bhuvan

Bhuvan (the name is derived from the Sanskrit word which means Earth), a Geoportal of ISRO and Gateway to Indian Earth Observation Data Products and Services (http://bhuvan.nrsc.gov.in or www.bhuvan.nrsc.gov.in), is an initiative of Indian Space Research Organization (ISRO), Department of Space, Government of India, to evince the Indian Earth Observation capabilities from the Indian Remote Sensing (IRS) series of satellites. Bhuvan is hosted at National Remote Sensing Centre (NRSC) Hyderabad India. The satellite images showcased on Bhuvan are from multi-sensor, multi-platform and multi-temporal domains with capabilities to overlay thematic information, derived from such satellite imageries, as vector layers on virtual globe for the benefit of user community. Apart from its unique visualization capabilities, Bhuvan also facilitates the users to download the satellite data and products through its Open EO Data Archive (NOEDA).

The unique features of Bhuvan are availability of uniform high resolution data (6m spatial resolution from LISS IV and 2.5 meter spatial resolution from Cartosat) for entire Indian territory, multi-sensor temporal data from IRS series of satellites, rich thematic data and information (Soil, Land use Land Cover (LULC), wasteland, water resources etc), visualization of ISRO's AWS (Automatic Weather Stations) data/information in a graphic view and use tabular weather data of user choice, Ocean Services, Disaster Services (timely support and services from space systems), Collaboration/Sharing/Community Participation (Volunteered Geographic Information), OGC Web Services, Mobile Compatibility (supports Android, Symbian, iOS and Windows Operating Systems). The major components of Bhuvan geo-portal are:

Visualization

This component focused on geo-visualization of satellite images and thematic maps.

- **Bhuvan 3D:** This map viewer is available for geo-visualization of three-dimensional data sets including digital elevation model (DEM) and other 3D features. Developed based on commercial software product known as skyline Globe and freely available for download as plugin at user end. This map viewer provides various tools for spatial analysis such as terrain profile, terrain analysis, horizontal and vertical distance measurement, flood analysis, urban design etc. As on April 2014, this 3D map viewer is compatible with windows operating system, however NRSC has also released beta version of 3D map viewer using JavaScript based WebGL which is compatible with all the operating systems.
- **Bhuvan 2D:** This map viewer is available as browser-based web GIS application which developed using open-source GIS software solutions. The OGC web service specifications are published for satellite images and thematic data sets. The map viewer provides seamless visualization of map layers at different zoom level. The spatial data up-to 1-meter spatial resolution is available in open public domain. Bhuvan2D also provides various tools such

draw tool for map feature creation, measurement tool for area and distance, navigation tools, WMS manager etc.

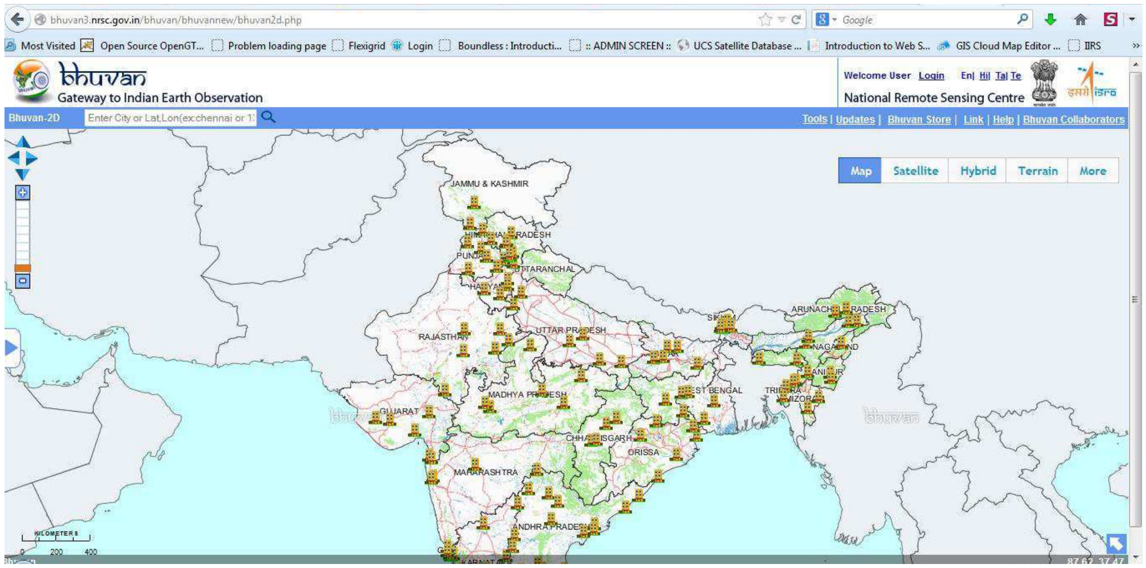


Figure 11. 7- Bhuvan 2D map viewer- http://bhuvan.nrsc.gov.in

- **Pocket Bhuvan (Mobile Version of Bhuvan):** This is a mobile version of Bhuvan 2D geo-visualization. This application is compatible with all the mobile based web browsers. However, Bhuvan also provides specific map application for Android and windows mobile operating system.

Services

This component provides access to various specialized services like:

- **Open Data Archive -**Allows download of free satellite data and products of specified period and resolution. This component is developed based on Remote Sensing Data Policy (RSPD)-2011 of ISRO.
- **Thematic Services -** To display and analyze thematic maps using WMS / WMTS (OGC web Services). The maps for different themes like LULC, wasteland, urban land use, Geomorphology, Lineament, Flood inundation etc., are available with various tools and services.

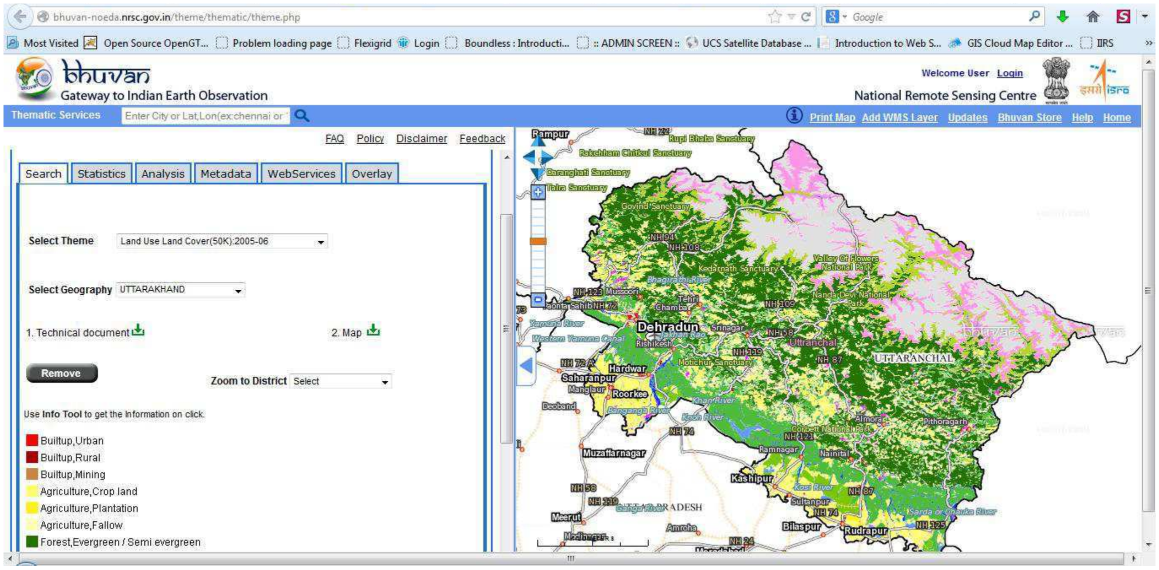


Figure 11.8- map viewer for Bhuvan thematic services- http://bhuvan-noeda.nrsc.gov.in

- **Disaster Services-** To provide timely information on various disasters for better decision making. This component is developed based on ISRO disaster management support programme which address six natural disasters viz. Flood, Drought, Cyclone, landslide, Forest fire and Earthquake. The disaster event based historical data for above disasters are available under disaster services.
- **Weather Services- This component provides v**isualization of data from Automatic Weather Station (AWS) of ISRO in near real time as well as archived mode.
- **Ocean Services -** Visualization of Potential Fishing Zone, Sea Surface Temperature, Chlorophyll information.
- **Mapping Services-** Comprehensive application facilitating user to embed their local knowledge and create value added maps through association.

Application Sectors

This component provides platform to create, visualize, share, analyze Geospatial data products and services towards various applications areas such as Agriculture, E-Governance, Events based spatial data and information, Forestry, Irrigation, Tourism, Urban etc.

Figure 11.9– WebGL based 3D map viewer of Bhuvan

In the recent past, Bhuvan geoportal is used as g-governance dashboard by various government departments and Ministries in India. ISRO has developed quite good number of mobile apps and online mapping applications to generated geo-tagged attribute data and thematic maps using crowdsourcing.

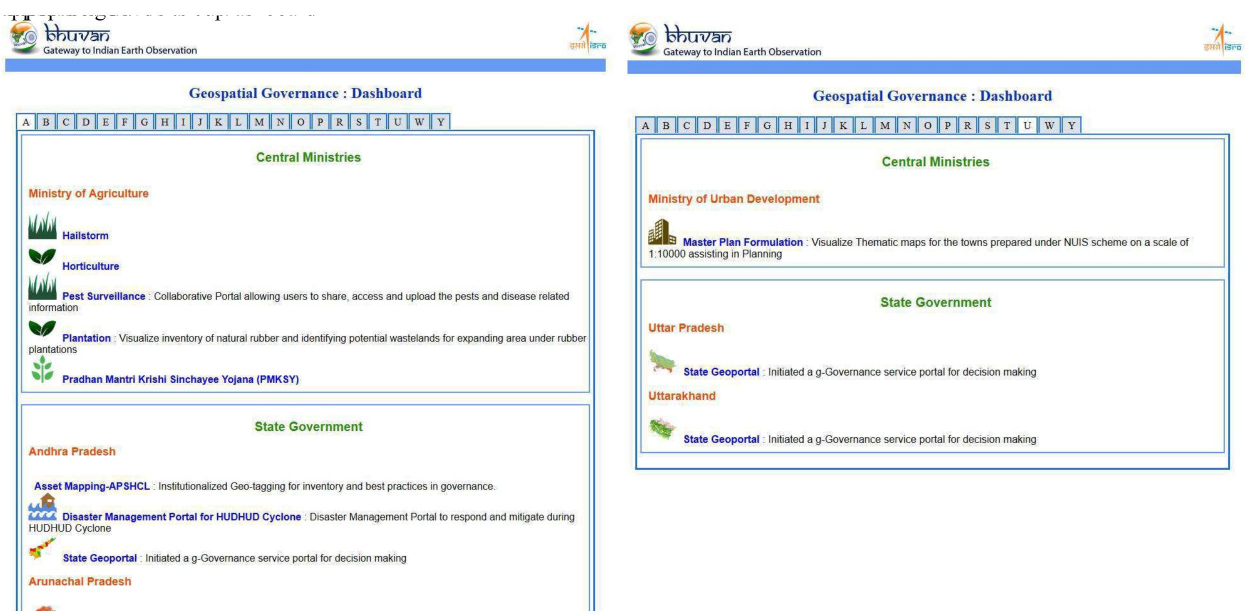


Figure 11.10- Geospatial Governance Dashboard in Bhuvan Geoportal

The crowdsourcing has become one of the important sources of data and information Bhuvan Geoportal. The geo-tagging in Bhuvan platform under various government flagship programme is successfully implemented in India. The Figure 11.11 shows the various geo-tagged assets under Mahatma Gandhi Rural Employment Guaranty Scheme (MGNREGA) of Ministry of Rural Development for Suyalbari Village Panchayat in Nainital District of Uttarakhand state:

Figure 11.11- Geo-tagged assets under MGNREGA programme in Bhuvan Geoportal

Similarly, you can also search for your own village in Bhuvan geoportal. You can also look for any other flagship programme of government of India where the crowdsourcing successfully implemented.

National Spatial Data Infrastructure (NSDI)

In India the geo-spatial data is available with diverse organizations like Survey of India, National Remote Sensing Centre, Geological Survey of India, Forest Survey of India, etc in different data standard and formats. Department of Science and technology (DST) in collaboration with Department of Space Government of India (DOS) has taken up an initiative to establish a Indian National Spatial Data Infrastructure (NSDI) for public domain in participation of other government organizations, private sectors, academia, research centers and NGO's. The Indian NSDI is available as a web portal under URL *https://nsdiindia.gov.in*. The India NSDI Portal makes data access and sharing of geo-spatial data easier, faster, and less expensive for all levels of government and the public. Indian NSDI focuses on de-centralized approach with the major emphasis on:

- Development and maintenance of standard digital collections;
- Development of common solutions for discovery, access and use of spatial data;
- Building relationships among organizations;
- Increase the awareness and understanding of the vision, concepts and the benefits.

In the NSDI framework the data providers of geo-spatial data provide their data under single web portal as a web service. There is a NSDI clearing house for browsing and accessing the geo-spatial data.

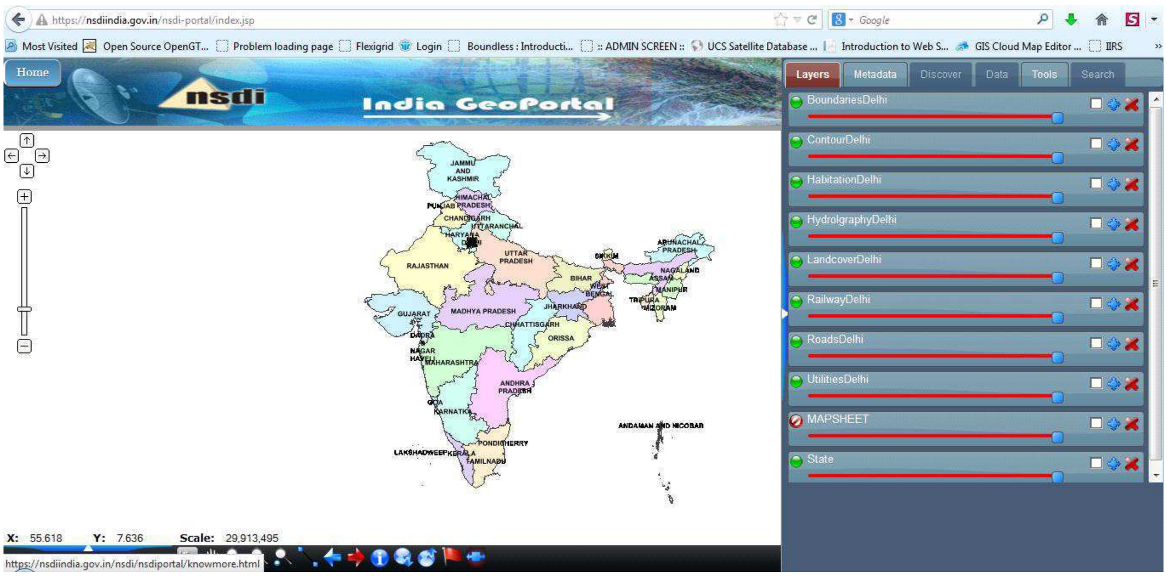


Figure 11.12– Map viewer and metadata search engine of Indian NSDI

The NSDI portal provides browser-based map viewer to visualize and search the spatial data available with different NSDI partners. The main focus is on metadata creation and development of metadata search engine. The metadata for various spatial layers is created based on NSDI metadata standard. The NSDI portal acts as a gateway or clearing house for searching the spatial data available in the country with different mapping organizations.

Biodiversity Information System (BIS)

National Biodiversity Characterization at Landscape Level, a project jointly executed by Department of Biotechnology and Department of Space, was implemented to identify and map the potential biodiversity rich areas in India. This project has generated spatial information at three levels viz. Satellite based primary information (Vegetation Type map, spatial locations of road & village, Fire occurrence); geospatially derived or modeled information (Disturbance Index, Fragmentation, Biological Richness) and geospatially referenced field samples plots. This relatively large spatial information on the above-mentioned facets of biodiversity has been organized in a web-based Biodiversity Information System (BIS) (http://bis.iirs.gov.in) for prioritization, conservation and bio-prospecting.

The spatial data available in BIS web portal is as following:

- Spatial Data on 1:50,000 scale for entire India
    - Vegetation Type map
    - Fragmentation map
    - Disturbance Index map an
    - Biological Richness map
- Species Database: Phytosociological database for 16,000+ sample plots for entire India

The portal provides information of high disturbance and high biological richness areas which are suggesting future management strategies and formulating action plans.

The BIS web portal provides data and information services under two major categories i.e., 'Biodiversity Spatial Viewer' and "Data Download.'. The biodiversity spatial viewer provides the utility to visualize all the spatial maps available in the BIS portal along with many base layers and administrative boundaries. The data is available in original scale. The WMS and WFS based quires are available which can generate many user defined map outputs. The portal also integrates the satellite data from Bhuvan geo-portal as background image directly accessed as WMS layer from Bhuvan geo-portal.

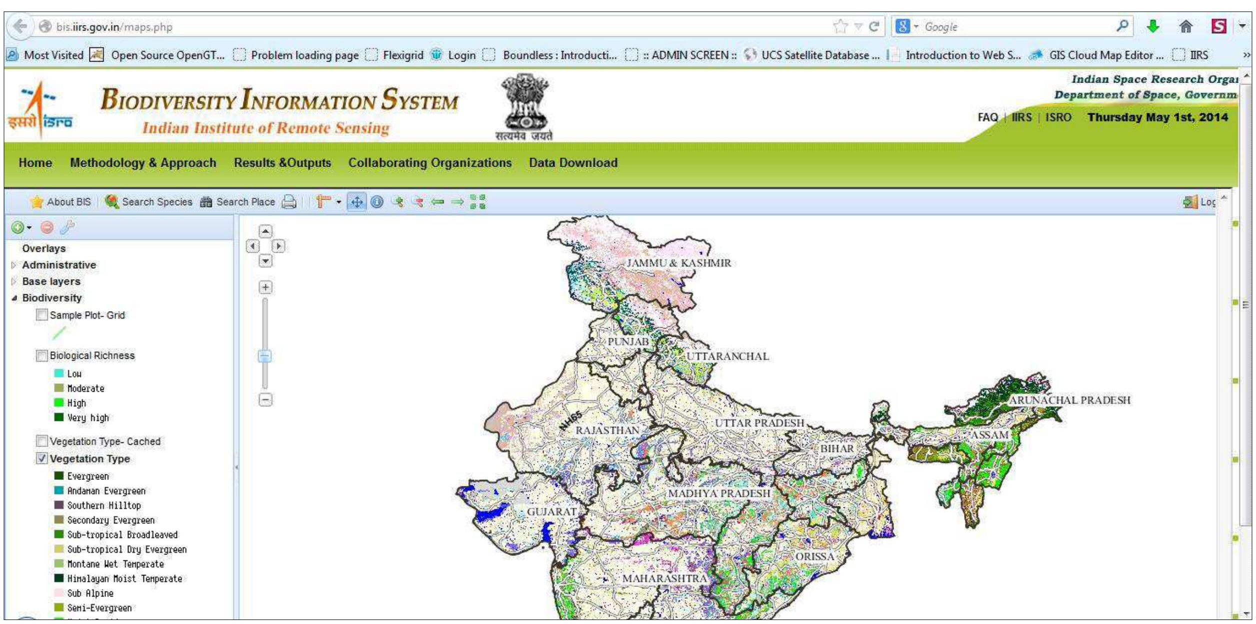


Figure 11.13– Biodiversity spatial viewer of BIS- http://bis.iirs.gov.in

Many unique features are available in biodiversity spatial viewer like overlay of local and remote maps through WMS URL, layer swiping utility for maps, WFS based query builder and filter utility etc. In addition to spatial data viewer BIS web portal also provide a data download utility in public domain. All the biodiversity data sets available in this website are freely downloadable through data download section. The user can define an Area of Interest (AOI) and select the map layer to download. The data download application will extract the data online and generate a quick data analysis report in real time mode.

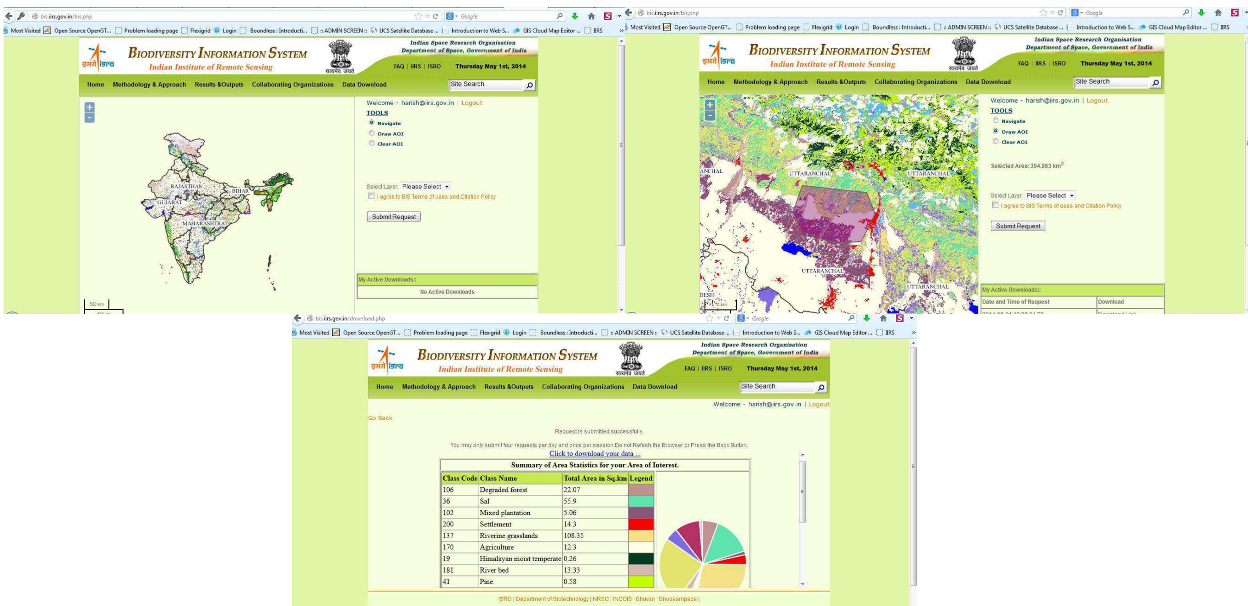


Figure 11.14– Data download and online analysis - http://bis.iirs.gov.in

The online data extraction and analysis for raster data are some of the unique features of Biodiversity Information System. This web portal is hosted at Indian Institute of Remote Sensing (IIRS), ISRO Dehradun India.

Indian Bioresource Information Network (IBIN)

Indian Bioresource Information Network (IBIN) (*www.ibin.gov.in*) is developed as a distributed national infrastructure to serve relevant information on diverse range of issues of bio-resources of India. IBIN web portal aims to offer a platform for all the data holders in the country to host their data at the same time maintaining their ownership on it. Its major goal is to network and promote an open ended, co- evolutionary growth among all the digital databases related to biological resources of the country and to add value to the databases by integration. IBIN portal provides platform to the diverse data providers in a mutually sharable environment with full ownership protocols. It aims at expanding the IBIN program through:

- Creation of an inter-communicating network of databases,
- Promoting an open ended and collaborative growth of participant digital databases and,
- Serving value added information on issues related to biological diversity and associated knowledge in multiple languages.

Data and information available in IBIN portal can be categories in two major categories:

- **Spatial Data Sets:** Under this category the maps generated in various projects on Bio-resources of India are hosted. The spatial data is available is an interactive map viewer with various queries and services developed on top of OGC web service specifications. The spatial data node provides access to primarily spatial data services. The browsing of different spatial and species data services is through biodiversity metadata catalogue.
- **Species Data Sets**: The species data is an attribute data available from various sources and different categories including plants, animals, microbes etc. Based on a common but flexible and indigenously developed inputting system, information was compiled from secondary sources. Suitable querying and outputting systems are also developed to retrieve different linkages among these databases. The data size of the compiled databases of all the groups till date accounts to about 6.8 GB on 39,000 species with 82,00,000 records.

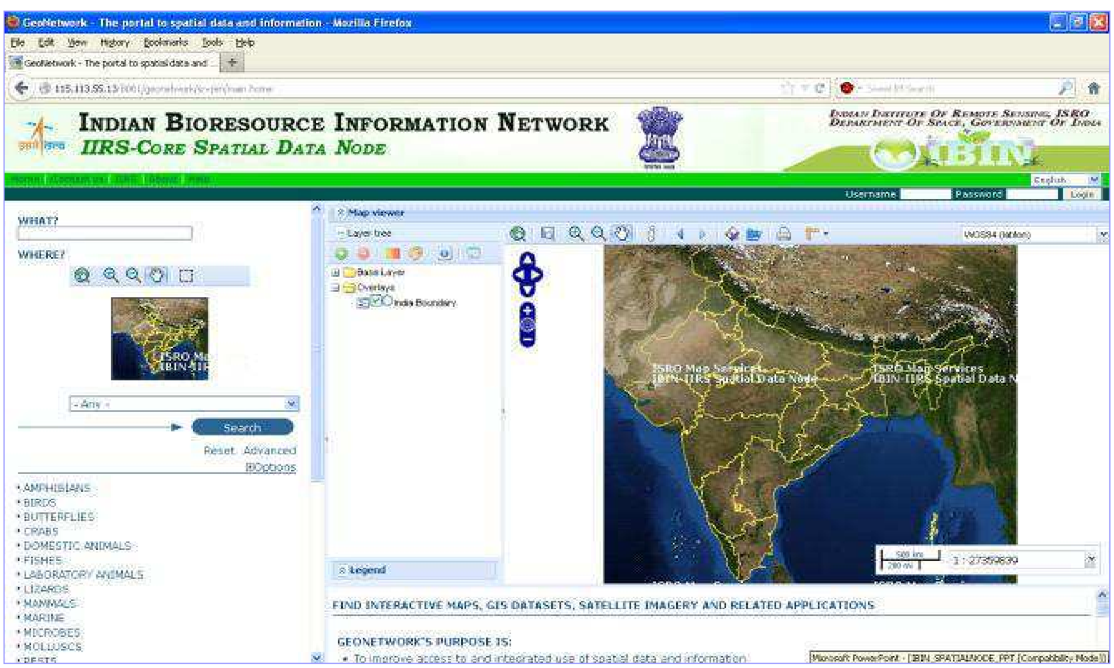


Figure 11.15- IBIN spatial node- www.ibin.gov.in

One of the unique feature of IBIN web portal is availability of different types of data related to Indian bio-resources in a single window. The spatial data is also available as a GIS data catalogue which allows various data search mechanism with good documentation in standard national framework.

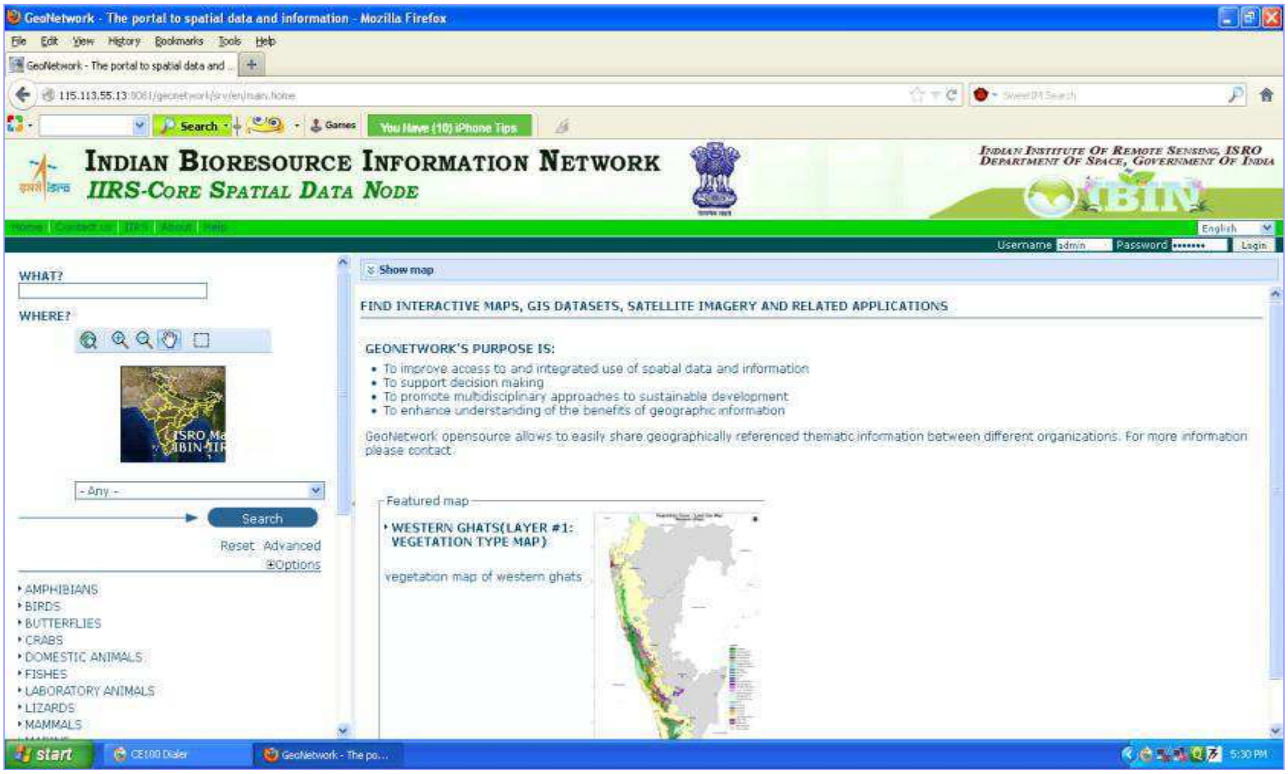


Figure 11.16- Metadata catalogue of IBIN geo-portal- www.ibin.gov.in

India-WRIS

India-WRIS Web GIS (http://india-wris.nrsc.gov.in) is geo-portal developed jointly by Central Water Commission (CWC) and Indian Space Research Organization (ISRO). The main objective of India-WRIS is to provide a Single Window solution for comprehensive, authoritative and consistent data & information of India's water resources along with allied natural resources in a standardized national GIS framework. It also provides tools to search, access, visualize, understand and analyze the data for assessment, monitoring, planning, development as an Integrated Water Resources Management (IWRM) in India. The current version India-WRIS Web GIS has spatial layers and attributes as per data collected from concerned state Govt. departments, CWC offices and Govt. of India departments (http://india-wris.nrsc.gov.in).

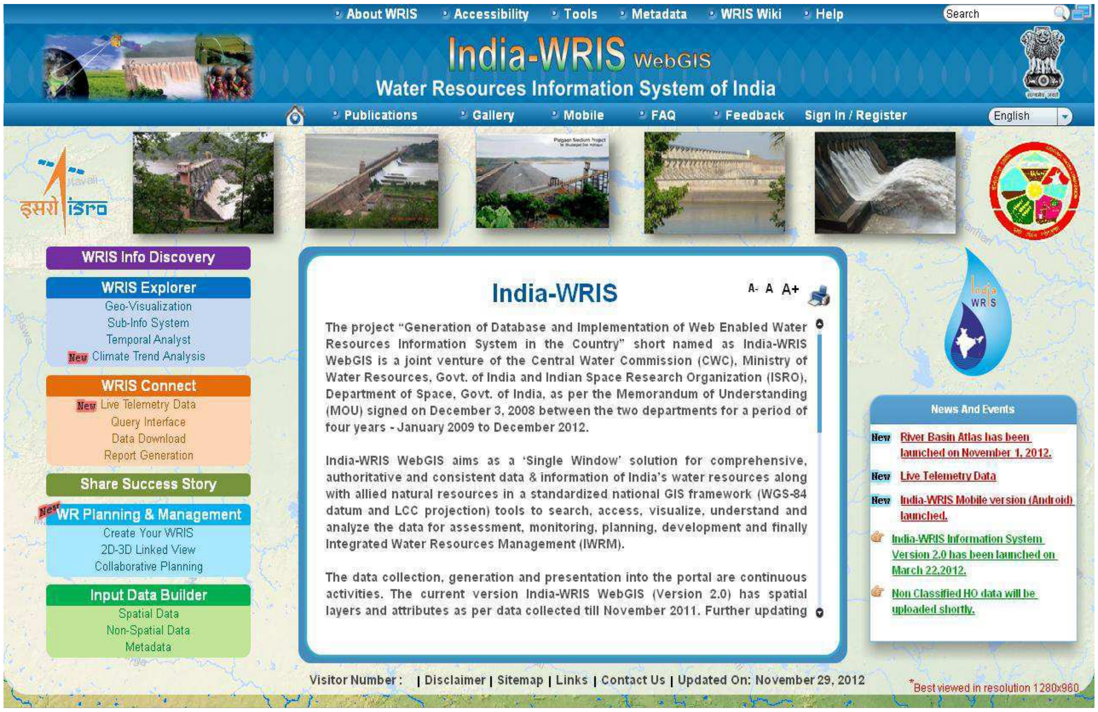


Figure 11.17 Home page of India-WRIS Web GIS portal-http://india-wris.nrsc.gov.in

Based on the type of data and its availability, the portal contains 12 major information systems, 35 sub information systems having 95 spatial layers along with large attribute data of the water resources assets and temporal data of 5-100 years. India-WRIS has six major sections:

- WRIS Info Discovery,
- WRIS Explorer,
- WRIS Connect,
- Share Success Stories,
- Water Resources Planning and Management and
- Input Data Builder.

The development and implementation of India-WRIS geo-portal is based on Flash API for ArcGIS. The system is developed using commercial software solution but the services are freely available in public domain. India-WRIS is a big repository of spatial data related to water resources of India.

11.4 SUMMARY

The GIS based websites or web portals has attracted the attentions of GIS users in last one decade. The mapping organizations have started to publish their data as web service which allows various exciting new applications at user end. The GIS portals display geo-data on maps and gives tools to interact with the data. There are several mapping portals available in internet either in public domain or in closed domain. Most of the web portals are compatible with majority of web browsers such as Mozilla Firefox, Google chrome, Internet explorer, Safari etc. These web portals also allow GIS users to connect to one another and to share their geographic knowledge. The popular GIS portal is summarized here:

S. No.	Web Portal	URL	Specialty
1	Google maps	http://maps.google.com	Global satellite images and maps (2D and 3D) at street level with various applications
2	Bing Map	http://www.bing.com/maps/	Global satellite images and maps (roads and POI).
3	Openstreet Map	http://www.openstreetmap.org	Open and free vector data and collaborative mapping.
4	Wikimapia	http://wikimapia.org	Crowdsourcing approach for tagging ground information.
5	ISRO Bhuvan	http://bhuvan.nrsc.gov.in	Rich contents and seamless availability of multi-temporal, multi-resolution satellite data for entire Indian region. Thematic and disaster services.
6	MapMyIndia	http://www.maps.mapmyindia.com	Rich POI and detailed maps of India.
7	Indian NSDI	https://nsdiindia.gov.in	Metadata catalogue and policy document.
8	Biodiversity Information System	http://bis.iirs.gov.in	Biodiversity spatial viewer and data download utility. Rich data contents on plant biodiversity of India.
9	Indian Bio-resource Information Network	http://ibin.gov.in	Rich data contents for Indian Bio-resource information. System of distributed database.
10	India-WRIS	http://india-wris.nrsc.gov.in	Rich data contents for water resources of India.
11	NIC GIS	http://gis.nic.in	Village level tagging of information in GIS domain.

11.5 GLOSSARY

- GIS Web portal - The ArcGIS enterprise portal that allows you to share maps, scenes, apps and other Geographical information with other people.
- Web GIS services - Web GIS services is an advanced form of Geospatial Information System.
- Thematic data - Thematic data is a qualitative data that involves reading through a data set and identifying patterns in meaning across the data.

11.6 ANSWER TO CHECK YOUR PROGRESS

1. What is google map?
2. Define GIS web portal.
3. Define Bing map.
4. Define ISRO Bhuvan.

11.7 REFERENCES

- GangothriRajaram, Harish Chandra Karnatak, SwaminathanVenkatraman, Manjula KR and Kannan Krithivasa, 2018 “A novel Computational Knowledge-Base Framework for visualization and quantification of Geospatial Metadata in Spatial Data Infrastructures”, Geoinformtica, Springer Science Publication, DOI: 10.1007/s10707-018-0317-6.
- Gautham A. S. "Google Revises Their Map, Ads Offline Version and 3D Imaging". TechGau.org. last accessed on 30th December, 2018.
- Harish Chandra Karnatak, Hariom Singh, Arijit Roy, RD Garg and Kamal Pandey, 2018 “Web-enabled scientific geoprocessing framework for online spatial modelling and algorithm development- A case study for spatial biodiversity modeling”, 38th INCA International Congress, October 29-31, 2018, Hyderabad.
- Harish Chandra Karnatak, Sameer Saran and P. S. Roy, 2005 “Spatial services, a click away” article in Geospatial Today Volume 3 Issue 5 pp-42-46 January-February 2005.
- Harish Karnatak, PLN Raju, YVN Krishnamurthy and A. Senthil Kumar, (2015), “E-learning Based Capacity Building in Geoinformatics”, ISG Volume 21, Issue 1-2, pp 4-13, ISSN: 0972-642X.

- Kamal Pandey, HC Karnatak, VaibhavKhulbe , 2017 "A framework of web processing services (WPS) for spatial data processing", Asian Confine on Remote Sensing (ACRS-2017), 23-27th Oct. 2017, New Delhi.
- Karnatak Harish, Karamjit Bhatia and Sameer Saran, (2008) "Multi-criteria decision analysis using Spatial Compromise Programming", Proceedings of the 2nd National Conference-INDIACom-2008, on Computing for Nation Development, ISBN No-ISSN 0973-7529, ISBN 978-81-904526-2-5, pp-77-62. February 08 – 09, 2008, New Delhi India.
- Neis, Pascal,; Zipf, Alexander, (2012), "Analyzing the Contributor Activity of a Volunteered Geographic Information Project — The Case of OpenStreetMap", ISPRS Int. J. Geo-Inf. **1** (2): 146–165, doi:10.3390/ijgi1020146
- Sameer Saran and Harish Karnatak (2012), "Spatial Decision Support System for Biodiversity Conservation Prioritization using geospatial modeling approach", NNRMS Bulletin- 36, Geospatial Applications and Decision Support Systems, ISRO, June 2012, pp-57-67.
- Sayar. A., 2006. Integration AJAX Approach into GIS Visualization Web Services, ICIW 06 Proceedings.
- Singh, H., Garg, R.D., Karnatak, H.C. , 2019 "Online image classification and analysis using OGC web processing service", Earth Science Informatics, Springer Berlin Heidelberg, https://doi.org/10.1007/s12145-019-00378-z Print ISSN 1865-0473 Online ISSN 1865-0481, pp-1-11
- Zhong-Ren Peng, Ming-Hsiang Tsou, (2003). Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks, ISBN: 0-471-35923-8m March 2003.

Web references:

- "New Bird's Eye Imagery &Streetside Coverage on Bing Maps!". Bing Maps Blog. Microsoft. May 8, 2012. Last accessed on 30thJanuary, 2019.
- A fresh new look for the Maps API, for all one million sites, http://googlegeodevelopers.blogspot.in/2013/05/a-fresh-new-look-for-maps-api-for-all.html, last accessed on 30thJanuary, 2019.
- ArcGIS online official website- www.esri.com/software/arcgis/arcgisonline
- Bing Map official website- http://www.bing.com. Last accessed on 10th February, 2019

- Bing Maps Imagery Release April 2010, http://www.bing.com/blogs/site_blogs/b/maps/archive/2010/05/07/bing-maps-imagery-release-april-2010.aspx, last accessed on 30th January, 2019.
- Bing StreetSide: Appeal is possible until end of September, Netzwelt, http://www.netzwelt.de/news/87752-bing-streetside-widerspruch-gegen-kartendienst-ab-sofort-moeglich.html. Last accessed on 10th February, 2019.
- Biodiversity Information System, Indian Institute of Remote Sensing- http://bis.iirs.gov.in. Last accessed on 1st May, 2014.
- GIS cloud manual- http://www.giscloud.com/docs/GIS_Cloud_User_Manual.pdf
- GIS cloud official website- www.giscloud.com. Last accessed on 30th January, 2019
- Google map official website- https://maps.google.com. Last accessed on 30th January, 2019
- Google Maps API FAQ", https://developers.google.com/maps/faq?csw=1#tos_commercial. Last accessed on 30th January, 2019
- Google+ Smartphone App Popularity-http://www.businessinsider.com. Business Insider. Last access on 1st February, 2019
- India Water Resource Information System, National Remote Sensing Center- http://india-wris.nrsc.gov.in. Last accessed on 1st February, 2019.
- Indian Bio-resource Information Network, official website- www.ibin.gov.in. Last accessed on 1st February, 2019.
- ISRO Bhuvan Geo-portal, National Remote Sensing Center, - http://bhuvan.nrsc.gov.in. Last accessed on 1st February, 2019.
- MapMyIndia official website- http://maps.mapmyindia.com. Last accessed on 1st February, 2019.
- National Informatics Centre, GIS Website- http://cen.gisserver.nic.in. Last accessed on 30th January, 2019.
- National Spatial Data Infrastructure official website-, https://nsdiindia.gov.in. Last accessed on 30th January, 2019.
- Wikipedia- http://en.wikipedia.org. Last accessed on 30th January, 2019.

Acknowledgements

This document is a compilation specifically intended for enhancing knowledge on various geoportal and online repositories. The data content is compiled from various web sources and all the resources considered are gratefully acknowledged.

11.8 TERMINAL QUESTIONS

1. What is web-based GIS?
2. What is the major source of data in google map services?
3. What are the important components of ISRO Bhuvan Geo-portal?
4. What do you mean by Spatial Data Infrastructure? Define major functionalities of Indian NSDI?
5. What is OpenStreet Map? What is the approach adopted in OSM for spatial data creation?