Big Data Visualization: A Game changer in GIS, Geo-analysis and Geo-demographics.

By Prince Ogbonna.

Today, everybody around the world is living and working under the coverage of Geographic Information system (GIS) application and services such as the Google Earth, GPS and much more. Big Data visualization tools are increasingly creating a wonder in the world of GIS. GIS has diverse application, from geo-positioning services to 3D demonstrations and virtual reality. Big Data and its tools of visualization has boosted the field of GIS. This article seeks to explore how Big data visualization has expanded the field of Geo- spatial analysis with the intention to present practicable GIS-based tools required to stay ahead in this field.

Geo-demographics is an explicit classification of socio-economic data, used to describe or analyze individuals and community characteristics. Geo-demographic information is mostly used by the government to properly plan and allocate resources. Commercial application can also be seen in demographic marketing.

Usually geo-demographic information systems are designed to generate area-based census. Big Data technology has helped in improving the accuracy of the information derived from these datasets.

Fundamental changes are experienced when big Data integrates with GIS, from large volume, fast velocity and vast variety. Big data technologies have capabilities to analyze massive datasets in social media and other platforms. Let's examine the big data production to see how it collaborates with Geographic information system. The big Data production process consists of collecting data, storing, batching and computing, analyzing and visualization. Creative advertisement and insightful interpretation can be produced when GIS analyst can effectively interpret the big data production process. Let's discuss the components of Big Data.

Components of Big Data

Big data is characterized by 3V namely, volume, velocity and variety.

1. Volume: Big Data consist of large volume of datasets which are usually in terabytes and originates from Global positioning system (GPS), sensors and social media. The connotation "Big Data" itself refers to a massive dataset.
2. Variety: Big Data deals with many kinds of dataset such as videos, pictures, sound, man and even text in social media fields. It is designed to accommodate structured and unstructured datasets.
3. Velocity: Big Data is very fast in analyzing real world data. Using a data technology call Map-reduce, it can navigate through millions of dataset searching or sorting for a specific data.

Processes and Analysis of Big data techniques

Fig 2. Description of GIS Data processes.

Like we discussed earlier, big data process involves data source, collection, processing, storage with analysis and visualization. Big Data sources originates from internal database of institutions, social media, video streams and search engines. Collection of this data is done by using a crawling method to reap for data on search engines, IOT sensors, social media etc. It then stores the data gotten using a no-sql data management structure, it extracts the data using Map-Reduce algorithm and then implements a distributed parallel processing with Hadoop. For Big Data analysis, researchers explore neurolinguistics programming for natural processing, or explore deep learning for pattern identification, and serialization for assigning

orders among data. Statistic tools such as R-programming language is used to efficiently analyze big data and remain integral for other statistical packages. Big Data visualization is the last part of the process which involves analyzing the datasets in a graphical or tabular format.

Big data and Geographic Information system

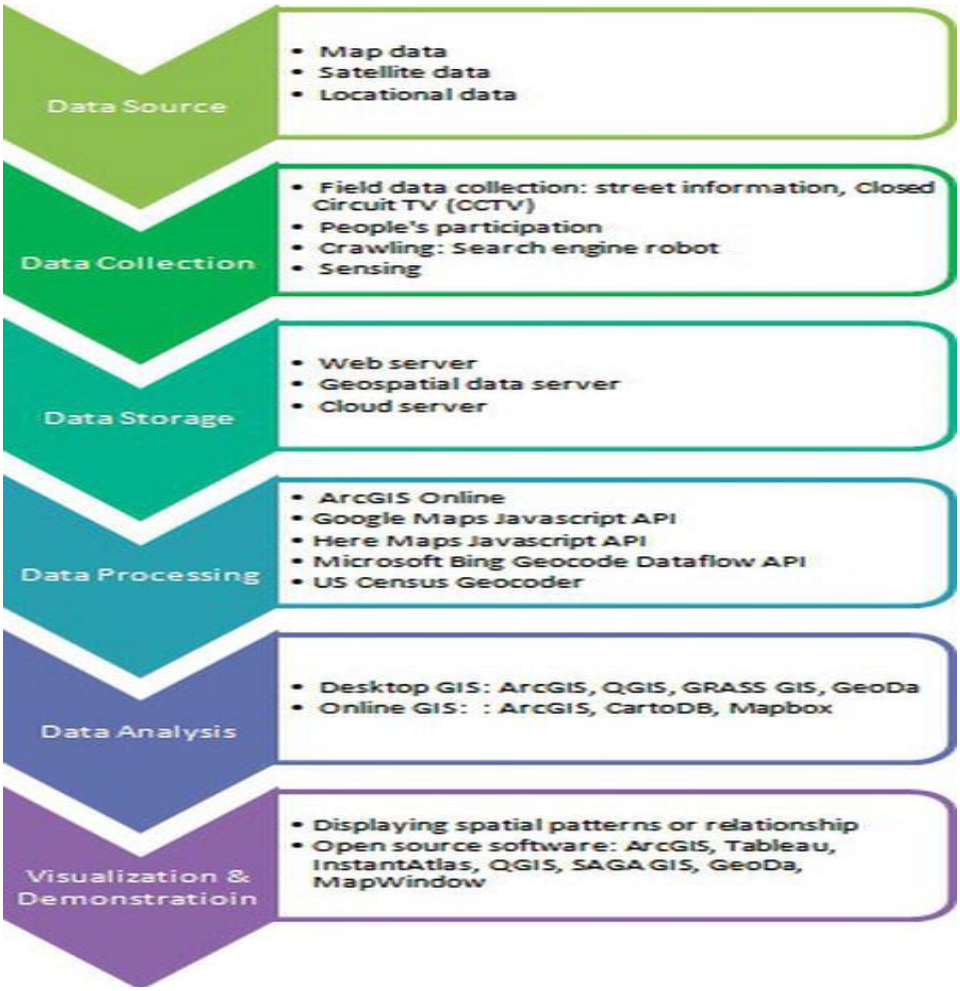

Fig 3. Description of GIS Data processing and software tools.

Above is a figure 3, that highlights GIS data processing with other components. It is important to note that there are several open source or commercial software that promote the usage of web-based online GIS systems. This tools play a pivotal role in processing and analyzing data.

Data Sourcing: GIS derives it data from several sources as highlighted earlier such as social media, search engines, institutional etc, this data contains space and location which are displayed as a map or picture form. Satellite and aerial data has increasingly been used recently to generate location based data.

Data Collection: collection of data can be achieved through CCTV, street information and several location datasets. When the dataset provided do not show location information, GIS operatives are permitted to perform geo-coding conversions to transform the data into useable GIS datasets. Recent technologies in collection has seen the expertise of crawling search engine robots that obtains GIS data on the web based on specified algorithmic themes.

Data Storage: GIS uses several platforms for storing data such as geospatial data server, web servers and cloud servers. These servers usually serve for specific operations but integrates largely with each other.

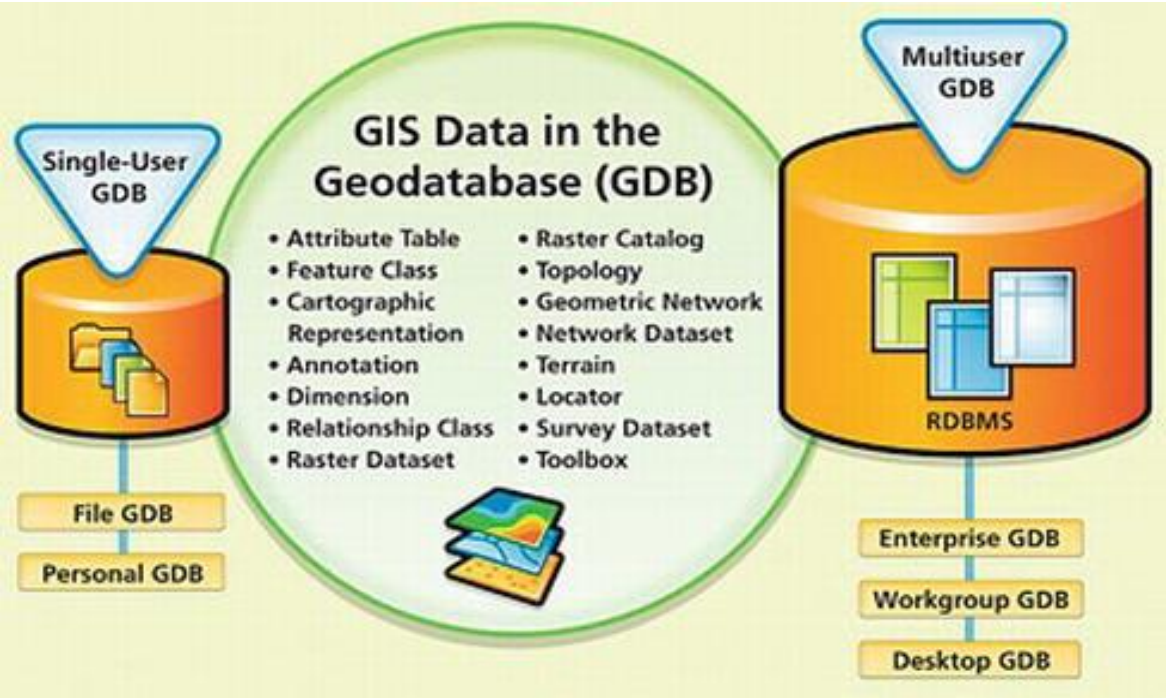

Fig 4. Geo-database for single and multi-users

The above diagram highlights the fundamental principle of geo-database for multiple and single users on ESRI. Understanding Geo-database system is important when handling advanced structured GIS datasets and their characteristics.

Data Processing: The rapid development of GIS software for web and mobile accessibility has improved the processing, analysis, and visualization. Some GIS software tools that has improved the processing of big data are Google Maps Javascript API, Microsoft Bing Geocode Dataflow API, Here Maps Javascript API, ArcGISonline. They all perform tremendous geo-coding and mapping functions in the database.

Data Analysis: As shown in the previous fig , ArcGIS, QGIS and Grass GIS all support desktop application while Mapbox, CartoDB and ArcGIS support online-real-time analysis.

Big-Data Visualization: This process showcase spatial patterns or the relationship among these locations. There are many popular software that are useful tools in big data visualization example are Tableau, InstantAtlas, ArcGIS, SAGA GIS, QGIS and MapWindow.

Big data visualization brings a whole new insight in GIS visualization since it does not just improve spatial context but bring new interpretations not previously seen to GIS maps. The field is growing as data is becoming more abundant, however meaningful insight to data analysis can be derived when data analysts are deliberate towards designing, calling, instructing and allocating big datasets.

By Prince Ogbonna