Assessing internet and web services based *webdom* and virtual web-data-centric geographical study

Abhay Sankar Sahu

Accepted: 8 November 2021 / Published online: 18 November 2021 © The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract At present, the traditional geographical study is somewhat shifted to virtual web-data-centric geographical study and it was started almost around last decade of the twentieth century concerning remotely sensed earth surface, digitisation, virtual data storing and their artificial intelligence based machine learning for spatio-temporal analysis, prediction and thematic mapping. Now in the age of webdom—a global digital kingdom of internet and web services, when internet services flooded everything relating to continuous up-gradation of computer-related science and technologies along with increasing number of internet users and GPS enabled smart phones, geographical data are generated by exaflood leading to bigdata. Many online data are now easily available to everyone from anywhere at any time. The objective is to assess the webdom and to understand the scope of contemporary virtual web-data-centric geographical study involving geographical bigdata, machine learning and WebGIS, when virtual world abundantly expands gradually during this period of webdom. Each data in bigdata contains geospatial attributes. Worldwide under the exponential growth and advancement of data science like other scientific disciplines geography is also predicting different phenomenon based on artificial intelligence and machine learning. The three main facets of virtual web-data-centric geographical study are geographical bigdata, geographical machine learning, and WebGIS. It is centrally concerned with digital data and web services. Here data are analysed through machine learning methods. Sometimes reliability of web-data is in question without field verification which is considered as the primary requisition of traditional geography.

Keywords Webdom · Geographical Bigdata · Artificial Intelligence · Geographical Machine Learning · WebGIS · Geographical Study

Introduction

Remote Sensing (RS) and Geographical Information System (GIS) introduce a new dimension in geography after intersecting with computer science (Goodchild, 1999, 2009, 2010, 2019; Goodchild et al., 2007; Wang & Liu, 2009; Sui, Goodchild, & Elwood, 2013; Miller & Goodchild, 2013; Miller & Goodchild, 2014). From the last half of the twentieth century models and data based analysis in different sectors and services are promised to a better output accompanied with data science and technological development. Accordingly, it is also found to develop different types of digital realities like virtual
reality, augmented reality and mixed reality (Farshid et al., 2018; Kaplan et al., 2020) parallel to the actual physical world which exists. All these digital realities are something for physical world representation on virtual platform sometimes applying trend analysis to get what will be the future scenarios. The virtual world after consuming our physical world data is expanding now where from anyone can access any information whenever required and also further experience augmented reality or mixed reality. At present, with continuous up-gradation and increasing use of computers and different kinds of software, people are somewhat tilted from non-digital learning activities to artificial intelligence (AI) based machine learning (ML) methods in researches. It is also true for spatial science like geography to build models for understanding real time and future behaviour of the earth surface phenomenon. Since virtual world is expanding consuming data therefore the digital data footprint is very speedily enlarging under which the real earth seems to be very small and it is continuing to infinity. It is presumed that geographical distance is becoming less to no-distant in-between different places on the earth surface. Cairncross (1995) assumed and wrote on ‘the death of distance’ in The Economist (Vol. 336, No. 7934). Here he raised the issue of worldwide low-internet cost in this twenty-first century as a result huge internet users can easily access internet from anywhere (Moore, 1995). It addressed less importance to the idea of geographical distance (Sui, Goodchild, and Elwood, 2013) and therefore one question was raised—will geography be insignificant in future?

The present time can be addressed as the age of webdom when vast and continuously increasing use of internet and virtual platform of web services conquered the whole world connecting almost every place through virtual communication technology and flooded everything of human livelihoods. Web is now powerfully attached to worldwide all-round human activities and concerns e.g. social, economic and political. Kaplan (2009) wrote on ‘the revenge of geography’ highlighting the increasing significance of the ideas on less distance and location specificity of geography in communication technologies which bound the virtual and physical worlds closer. Geographical knowledge and information communication technologies are thus crucial to business and economic activities, social and cultural affairs. To answer of the early raised question about the less significance of geography in future, ‘the revenge of geography’ has established the significance of geography indicating growing importance of location and geocoding or geotagging in the ocean of bigdata (Sui, Goodchild, and Elwood, 2013) as also Gore (1998) remarked that in twenty-first century we are in ‘digital earth’ where every data is georeferenced.

In this present overwhelming internet and virtual world outburst, the prime objective of this study is to assess the current time as webdom—a global digital kingdom of internet and web services, when virtual world abundantly expands gradually. And, belonging to this time toward understanding the present major systematic change in the study of geography it is also to understand the scope of contemporary virtual web-data-centric geographical study involving technologically concerned webdom content geographical bigdata, machine learning and WebGIS. Here, research questions are—What does practically webdom refer to? What is the content of webdom? In the vast arena of technological applications, digital flooding, internet and web services development where is hard-core field survey based geographical study in this twenty-first century? Is geographical study shifting from its laborious traditional field survey based data acquisition process to easily available huge online collection of data and information? Is geographical study now more concerned on continuously generated and easily available geographical bigdata, artificial intelligence based machine learning and WebGIS? How does today virtual web-data-centric geographical study seem to be a major shift in the system of geographical study? What its future trend is? This study is therefore more specifically to reveal under virtual web-data-centric geographical study present scenario of generation and online availability of huge geographical data and information from different sources as well as more dependency and use of geographical bigdata, use of artificial intelligence based geographical machine learning involving geospatial models and software for clustering spatial and non-spatial geographical data and systematic analysis, and WebGIS as modern-time, user-friendly, problem-specific online service.

The base level key ideas and elements behind to understand the concepts of webdom and virtual web-data-centric geographical study are computer science and technological applications toward internet and web service development, creation of virtual world and also digitisation of real earth, bigdata science,
GIScience, GPS enabled smart phones in relation to public involvement and social networking for huge data generation, artificial intelligence and machine learning in geographical analysis, WebGIS over the virtual platform, and concept of neogeography.

**Webdom: A Global Digital Kingdom of Internet and Web Services**

The World Wide Web (www) has transformed all and everything with relation to human activities and concerns. The Webdom refers to the present time of global digital kingdom involving extensive use of virtual data and online internet communication through World Wide Web services by people in their livings and that expands continuously with increasing human dependency on it in parallel. The webdom was started over the developed countries first then the other parts of the world. It covers urban areas first then rural areas get attached with it. Uses of computer and smart phone are certainly primary physical medium for this webdom development with uninterrupted availability of internet connection as well. Nowadays many people always remain internet connected and attached sometimes for whole days with web-based services to meet their nearly every need. In this webdom period people abundantly use web services for simply chatting to video conferencing, research and education, numerous data and information collection, entertainment, travelling, shopping, banking, and so on. Again, during this webdom period, in the age of software as well as model based analysis everything of the natural and social science are presented in the form of digital data and for their analysis we considered scientific data-centric methods. Here for rectification or validation of any type of obtained digital data and for clustering, classification, analysis, and also to get conclusion again in the form of data we use objective methods. Precisely, we collect raw-data then analyse using data intensive methods or models as instructions or training-sets and get again data as research finding in the final output. That key-findings data again helped to make conclusion of another obtained data-set. Nowadays everything is going to be numerically measurable.

The webdom is inherently associated with the applications of computer science and ICT centric internet and web service development and mashups. It is involved with the Internet of Things [IoT] (Xia et al., 2012) that has changed the way of human livings. Mayer-Schönberger (2011) pointed out that information and technology i.e. IT revolution is evident all around us where along with technology data and information acquire much significance in our lives and education. From the year of 2020, during COVID-19 pandemic situation, world-wide uncountable online meetings, academic classes, offices, webinars, live streaming programmes, uploading videos, chatting and so on strongly established the necessity and significance of up-gradation of internet communication network management along with increasing number of internet users and GPS enabled smart phones. González-Bailón (2013) has rightly remarked that at present our ways of observation and experience the world have been changed by digital technologies. Table 1 shows how the webdom developed with time and its present form including some of the several inherent subject matters encompassing science and technology. Both computer science and technology thereby ICT developed and expanded continuously from almost the decade of 1980s. After the year of 1990 when World Wide Web (www) came on digital platform the data management system was reformed with communicational revolution. The whole 1990s decade was to enhance worldwide digital transformation from physical laborious paper works in banks, commercial sectors, governmental and non-governmental offices, and in other areas also for documentation, record, presentation, publication, transfer of data and information as it is said that the web changes. A continuous up-gradation is also found in web searching and communication network management system after coming of the Web 2.0. Applications of numerous high level computer programming languages also gradually were developed throughout the twentieth century practically during 1990s and also now like SQL, Python, Visual Basic, R, Ruby, Java Script, C#, etc. To geographers, GIS first became widely available in 1980s (Goodchild, 2019). From the very beginning of 1990s to present time, the ICT is experiencing massive development also in RS and GIS technology. It is happened through launching new satellites from different countries. Different types of active and passive sensors have been introduced and upgraded continuously with respect to resolution variations for acquisition geospatial information more accurately. Presently, it is also observed that integration of RS technology by
| Time       | Milestones                                                                                                                                                                                                 |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Before 1980| Hewlett-Packard (HP) is founded (1939) and computer thoroughly developed with time<br>The first computerized GIS in the world (Canada in 1963)<br>@ and E-mail started (First e-mail in 1965); ARPANET’s networked (1969); concept of internet begins after e-mails transfer<br>UNIX, an operating system (1969)<br>Ray Tomlinson invented and developed electronic mail (1971)<br>Floppy disk (1971)<br>Landsat I & MSS (USA in 1972)<br>Apple I, the first computer with a single-circuit board (1976)<br>MOSS—a pioneer vector-based geographic information system (GIS) developed (USA in 1978), an open Source GIS Development of computer programming languages e.g., COBOL, PASCAL, C |
| 1980–1990  | Internet service developed throughout the 1980s, Internet Service Providers (ISPs) started connecting people across the world<br>The first IBM personal computer (1981) using Microsoft’s MS DOS<br>ARC/INFO for minicomputers was launched by ESRI (1982)<br>Landsat 5; MSS & TM (USA in 1984)<br>MS Windows (1985)<br>Dot Com (1985)<br>MIDAS—the first desktop GIS operated in MS DOS operating system (1986)<br>IRS 1A (India in 1988)<br>Development of computer programming languages e.g., MATLAB |
| 1990–2000  | Cheaper, faster and more powerful computers developed throughout the 1990s<br>In 1990s digitization of spatial data and information, records, documents, etc. increase<br>Tim Berners-Lee developed HTML, giving rise to the WWW (1990)<br>Web 2.0; AI and machine learning developed; multiple software and data availability; infancy of bigdata<br>MIDAS renamed to MapInfo for Windows (1990)<br>ArcView 1.0 by ESRI (1991)<br>The full GPS works with 24 satellites (USA in 1993)<br>Yahoo (1994) and Google (1996)<br>Digital PDA based phone (Nokia, 1996)<br>Qualcomm released integrated smartphone (1999)<br>Wifi (1999) helped easy internet use<br>ArcMap 8.0 by ESRI (1999)<br>Landsat 7 and ETM + (USA in 1999)<br>The launch of new satellites by different countries and integration of remote sensing technology by different sectors<br>Significant development of high level computer programming languages e.g., Python, R, Java, JavaScript, Ruby, php<br>Grid computing originated<br>Web mapping popularised and developed |
| 2000–2010  | AI, machine learning and deep learning steadily increase<br>RS and GIS technology continuously developed<br>Cloud based bigdata expands<br>Google Earth (2001)<br>Q GIS released (2002)—free GIS software<br>Resourcesat-1 (India in 2003)<br>Android in smartphone (2003) |
different sectors wherever necessary introducing different types of software and models for data analysis and management. In many countries the 1990–2000 decade was for starting journey and developing computer based science and technological services in general and specifically RS and GIS jobs in geography, then the period from 2000 to 2010 was for the second gear up of all digital services in respect to Web 2.0. When Web 2.0 came inside the field of webdom generation of myriad data was started and now its footprint is enormous which is called bigdata. Since around 2005 throughout the world coming after Android in smartphone (2003), Facebook (2004), YouTube (2005), Google Map (2005), and so on in webdom it is experienced that use of web platform for social media as first preference and then for other searching and it exponentially increases with huge consumption of smart phones. Che, Safran and Peng (2013) mentioned an estimation made by Google in 2010 that during that time in every two days the whole world generated data which is equal to the sum of data totally generated up to the year of 2003. As Sui, Goodchild, and Elwood (2013) mentioned all of the geospatial related technologies like GPS, smart phones, wireless sensor networks, cloud computing, and all the technologies loosely called Web 2.0 radically changed the GIS related techniques of data collection, data management, spatial and thematic mapping, and analysis. Artificial intelligence and machine learning have deep root in the system of data analysis and data management part. Their applications are found to be increased with varying rates since the late 1990s and got acceleration thereafter from the very beginning of this twenty-first century along with digital transformation of geospatial environment for especially prediction related purposes in land and water management, ecology and environment, site suitability, hazard forecasting, and so on. The present time is also known as the era of deep learning, a subset of machine learning, where artificial intelligence is used for analysing different categories bigdata to cluster, classify and prediction. Roser, Ritchie, and Ortiz-Ospina (2020) present different types of spatial and temporal data on internet users and usage. According to them the number of internet users increased throughout the world very rapidly from 413 million in 2000 to over 3.4 billion in 2016. Diwanji (2020a) presented an increasing trend of internet users in India from 2015 (259.88 million) to 2018 (483 million) and a forecast until 2023 (666.40 million). Diwanji (2020b) also presented much of these internet users are using social networks and for this purpose in 2015 there were 142.23 million users and it was 326.10 million in 2018 and it is predicted to be 447.90 million in the year of 2023 as well. At present, all social networks every day originate large volume of unstructured bigdata wherefrom using machine learning techniques lot of geographical

| Time          | Milestones                                                                 |
|---------------|-----------------------------------------------------------------------------|
| 1990–2000     | Mozilla’s Firefox 1.0 (2004); Facebook (2004); OpenStreetMap (2004)         |
|               | YouTube (2005); Google Map (2005)                                           |
|               | iPhone of Apple (2007)                                                      |
|               | Development of computer programming languages e.g., C#, F#                   |
|               | Cloud computing popularised and expanded (2006–2007)                        |
| 2010–2020     | AI, machine learning and deep learning developed continuously and highly     |
|               | accepted in different sectors for apply to solve many problems              |
|               | RS and GIS technology developed and used in many sectors                     |
|               | GIS data has become more ubiquitous and some are free; expanding geographical |
|               | bigdata                                                                      |
|               | Resourcesat-1 (India in 2011)                                               |
|               | Landsat 8; OLI and TIRS (USA in 2013)                                        |
|               | NASA’s World view tool                                                       |
|               | Development of computer programming languages e.g., TypeScript               |
information are accessed by internet GIS or WebGIS platform. From early 1990s WebGIS is available. It is more dynamic and it helps very easily to share geographical information (Fargher, 2018). More precisely, WebGIS is used mostly as information sharing and information dissemination technology. In WebGIS maps are used with user-friendly interfaces where much GIS knowledge is not required to study.

Exaflood of Bigdata

We have entered in an unprecedented age of bigdata (Che, Safran and Peng 2013), a new paradigm (Jiang, 2015). This is the time of ubiquitous data (Graham & Shelton, 2013). A. Gore in 1998 during a lecture has mentioned on his vision ‘digital earth’ for the twenty-first century (Gore, 1998), which requires bigdata (Sui, Goodchild, and Elwood, 2013). Che, Safran, and Peng (2013) mentioned concerning our lives and education, bigdata plays a crucial role. Bigdata refers to large as well as massive volume of heterogeneous types of data (Goodchild, 2019; Graham & Shelton, 2013; Jiang, 2015; Talia, 2013; Taylor, 2012; Zikopoulos et al., 2012) in different formats accumulated from social media, different satellite sensors, online publications, online transactions, and virtual storing and these are increasing continuously day by day exponentially in cloud computing network (Aymerich et al., 2008; Talia, 2013). It differs from small data where bigdata refers to the grand total and small data is a part of it (Jiang, 2015). Bigdata involves exaflood (Goodchild, 2019) that means rapidly increasing uncountable flooding of data transmitted through internet connections. Swanson (2007) remarked on how exaflood of bigdata is coming through communication sites, video sites and others. Bigdata is large volume of online stored data having wide variety and velocity at which data in system generated, collected and processed (Jiang, 2015; Kitchin, 2013; Mayer-Schonberger & Cukier, 2013). These are different from traditional data type and during their management to extract hidden information from the large volume of data presently advanced data mining techniques and associated tools are widely used including Hadoop, MapReduce, and Big Table (Khan et al., 2014; Talia, 2013). Bigdata is the combination of three types of data (Zikopoulos et al., 2012) like structured, much in unstructured, and semi-structured used to machine learning and deep learning for understanding issues, advanced analysis and toward predictions in different sectors and services. For the purpose of bigdata management and to handle new digital data, more and more ICT related infrastructure and servers are needed as Sui, Goodchild, and Elwood (2013) mentioned. According to Khan et al. (2014) structured data are generating through automated data generators including computers and sensors and these can be processed using query languages. Unstructured data are in the form of images and videos generated from social media. These are complex to process due to their different formats though application software can be applied for clustering and analysis. Lastly semi-structural data like XML are also not followed any predefined type therefore they are not processed easily (Khan et al., 2014). Many online resources used XML format for data representation (Abdo & Alali, 2016). Problems in bigdata are therefore unstructured to semi-structured data complexity especially during their extraction and separation as and also between academic, governmental, and commercial data from the large influx of information and validation of all items (Khan et al., 2014; Linnettaylor, 2012) as Talia (2013) also mentioned bigdata often have unstructured digital content and their processing is highly difficult using traditional data management tools and techniques. Present huge bigdata are storing in cloud base and they need large system with processing speed for management.

Artificial intelligence, machine learning and deep learning

Artificial intelligence (AI) is involved with computer software that simulates human-like intelligence processes. It is to solve different kind of real world physical and human related problems. It detects one from massive data and information. It is largely helpful for future planning and also to predict future trend which for AI has been programmed. In this webdom period significance of AI and its areas of application are continuously increasing accordingly with human-less technological upgradation and web services. Starting after Turing test (Turing, 1950) the concept of artificial intelligence continuously has been progressed with time and in recent time it is a part and parcel of the modern computer centric IT civilisation, our webdom. There are several approaches like logical, probabilistic to artificial intelligence for the conceptual framework wherein possibility of an output
can be predicted following complex interconnecting networks (Neapolitan & Jiang, 2018). It is associated with machine learning (ML) and deep learning (DL). Artificial intelligence, machine learning and deep learning now radically have changed the world of virtual space and virtual reality. These are helpful to make a better world for us involving agricultural development, hazard management and social well-being (Singh, 2019). Artificial intelligence enables the machine like a human brain to come up with results what for that has been designed and installed. Machine learning is thus a computer based system including computational methods or data-driven algorithms using past available information which typically in the form of electronic data either digitised human-labelled training-sets or other types collected and made available for analysis as ‘experience’ toward improve performance or to make accurate solutions and predictions (Mohri et al., 2018; Shalev-Shwartz & Ben-David, 2014). Simply machine learning consists of the association of artificial intelligence efficient and accurate prediction digital algorithms. The deep learning is considered as a sub-set or a type of machine learning and it is associated with much sophisticated more complex computer-generated deep neural networks loosely resembling the human brain as deep AI to solve target problems and make precise solutions and predictions (Neapolitan & Jiang, 2018; Singh, 2019).

Scope of virtual web-data-centric geographical study

In geography, geographical data are inherently spatial in nature (Voženilek, 2009) and those create model based GIScience. Goodchild (1999, 2010) remarked that GIScience is an intersecting discipline of geography and computer science or information science. During this webdom age it is appropriate to mention like a new paradigm a major systematic change or modification in geographical study is virtual web-data-centric. It is basically referred to the process as well as somewhat present form of geographical study using computer and smart phone along with internet connection and web browsing environment fully. It is the geographical study from anywhere of the world in anytime searching for anything globally related to geospatial and human ecological. It is worked over the virtual space involving human-technological interactions. It is thus related with all connecting aspects of digital data generation, archiving, sharing, searching, displaying, querying, and spatial analysis and GIS modelling services. It includes both science and technological applications. It is the study of spatial science inherently associated with the knowledge of computer science and technological applications, information science, bigdata science, and GIScience. Broadly, it consists of three categorised elements—computer related science and technologies, virtual data and information management, and geospatial analysis and GIS modelling services (Fig. 1). The virtual web-data-centric study is thus developed due to the development of internet and communication network management, web science and services for its wide level users. In the developed countries it was developed first then in other parts of the world. It is also related to artificial intelligence, machine learning and deep learning. It focused on bridging between real or physical world and virtual world. It is associated with the knowledge and applications of virtual reality, augmented reality and mixed reality. It contains web-data science with the agenda of GIScience as mentioned by Goodchild (1999, 2010). It is thus associated with advanced cartography and geoinformatics. Artificial intelligence and machine learning are inclusively part of GIScience. Navigational tool like GPS, Differential GPS (DGPS) to get more spatial accuracy, UAV (drone)—a digital data generating advanced survey instrument wherever and whenever required, remote sensing and GPS enabled smart phones are nowadays advanced technologies for spatial data acquisition. More devastating remote sensing technologies and applications generate vast geographical bigdata. The cloud database is ex-flooded also through crowdsourcing, social networking, user generated themes, digitization, geospatial information, and through many other ways. These all are considered as virtual sources of huge data in the virtual web-data-centric geographical study. The WebGIS, an advanced form of GIS, available on web platform is a part of it. Knowledge of different model handling and computer based GIS techniques and methods for systematic analysis and presentation are essential in the content of virtual web-data-centric geographical study. It is as Gore (1998) remarked on georeferenced ‘digital earth’ in twenty-first century engaged to understand all georeferenced information.
of the earth surface thereafter classify and management. Virtual web-data-centric geographical study specifically highlights the digital transformation of geographical information in large volume, with large variety and by high velocity to all people. It is centrally thus concerned with the management of geographical as well as geospatial digital data and web services.

The three main facets of the virtual web-data-centric geographical study are geographical bigdata, geographical machine learning, and WebGIS (Fig. 1).

**Geographical bigdata**

In the vast ocean of generally called bigdata, geographical bigdata are all data as Gore (1998) remarked georeferenced digital information those refer to specific places on the earth surface to create our ‘digital earth’. Simply geographical bigdata are spatial in character. These refer to location specific and/or captured as well as geocoded or geotagged huge cloud based digital data in different formats generating after digital transections and online interactions existing as a sub-set in the vast ocean of bigdata. Much of the data having georeferencing as well as spatial attributes (Kitchin, 2013) of bigdata are geographical in nature. Satellites through their numerous sensors continuously produce much heterogeneous structured and unstructured bigdata (Che, Safran, and Peng 2013) which are inherently geographical bigdata. Originating from social media as described as organic data and transactional data by Linnettaylor (2012) massive geographical bigdata are concerning to human geography (Kitchin, 2013). All geotagged videos and photographs are also geographical bigdata available and streaming in YouTube, Facebook, Whatsapp, National Geographic, and so many. Different publications and research sites generate geographical bigdata e.g. Wikipedia, Springer, Elsevier, Taylor & Francis, HydroSHEDS, ISRO-NICES, Google Scholar, ResearchGate, Academic.edu. Remote sensing satellite data increases geographical bigdata volume available in USGS EarthExplorer, ESA Copernicus, ISRO-BHUVAN, etc. All the geographical bigdata sources can be categorised into three sections (Kitchin, 2013)—directly technologically captured digital data of a place by human operators like advanced data using modern survey instruments (e.g., UAV), cameras, etc., automated generating data through different devices or systems like sensors, and volunteered data uploaded by users and others through publications, social media, crowdsourcing system such as OpenStreetMap (OSM).

**Volunteered As Well As Crowdsourced Geographical Data.** Goodchild (2007) widely analysed the volunteered geographic information (VGI) highlighting citizens as sensors. From the very beginning of twenty-first century as internet users increased very rapidly throughout the world a popular form of GIS in the name of VGI developed, a crowd-sourcing version (Goodchild & Li, 2012) in which general people create and contribute georeferenced facts and geospatial data to websites where those are synthesized into databases. According to Sui, Goodchild, and Elwood (2013) in the age of Web 2.0 VGI is associated with a radical shift in the content, characteristics, and modes of geographic information creation, sharing, dissemination, and use. Haklay (2010) remarked that as VGI develops highly trained surveyors would be replaced...
in future by technicians engaging high accuracy GPS and receivers manufacturing, mapping industries, and so on. This twenty-first century is presumed as spatial century where there is a geographer in most people (Gould, 1999) and this potentiality to be a geographer in everybody manifested through the vast technological use of VGI (Sui, Goodchild, and Elwood, 2013).

The OpenStreetMap (OSM) project was launched in 2004 (Jokar Arsanjani, Zipf, Mooney, and Helbich, 2015), created by Steve Coast of UK. The OSM is the free online and editable form of the map of whole world on web interface. Initially it was started for mapping roadways but now its aim and scope increased as per Jokar Arsanjani, Zipf, Mooney, and Helbich (2015). The OSM is the geographical web database of the VGI system where interested people voluntarily contributed geospatial information on physical world and cultural landscapes. Landuse/land-cover, communication lines, buildings, etc., of an area are marked by point, line and area. The OSM project is a revolutionary work in this time of webdom and virtual web-data-centric geographical study concerning the way in which geodata is collected, virtually stored and management though it has obvious commercial benefits. Haklay (2010) remarked on achievements and difficulties of VGI in the context of OSM study in England. The process of gathering of geographical information is being very much easy through VGI system where the GPS and the receivers like smart phones allow general untrained people to acquire their exact position automatically and accurately.

Geographical machine learning

In the context of geography, from the early 1990s, scope of remote sensing and GIS related science and technology was enhanced with increasing passion of artificial intelligence and machine learning. When we consider geographical machine learning (GML) in relation to geographical as well as geospatial data used, digital algorithms and training-sets designed to geospatial analysis, theoretical and field validation of the results then it is what we call GIS, a hardware and software based system for geographical analysis and prediction. Inherently the GML refers to the geographical data analysis with ideas from mathematical fields of statistics, information theory, probability and optimization. Considering the numerous practical applications and tasks of machine learning as indicated by Shalev-Shwartz and Ben-David (2014) these are too complex to control. To Mohri et al. (2018) these are text or document classification, natural language processing (NLP), speech processing applications, computer vision applications, statistical regression, ranking, and clustering. The GML task applies to spatial features identification, distribution, shape, density, contiguity, extraction, supervised and unsupervised classification, ranking, clustering, measurement, site suitability, relationship, predictions, integration with others, digital mapping and presentations like 3D, super imposing, and so on. The artificial intelligence, machine learning and deep learning are closely associated with the advancement of GIS technology and in the context of GML, at present, deep learning very much is used to geospatial analysis and digital mapping, especially for real-data analysis in climate and weather forecasting, agriculture, environment and hazard management. In the context of GML there is a vast scope of digital realities like virtual reality, augmented reality and mixed reality as these bridges between virtual world and physical world e.g., for clustering people based on shopping category, present urban areas and growth thereafter future prediction as dissemination technology. To analyse satellite imageries, drone captured data and other any type of digital spatial or non-spatial data, GML is progressing rapidly especially in the context of landuse/landcover classification, site suitability analysis, natural resource management, hazard management, regional and town planning. GIS software can automatically extract agricultural lands, forest lands, water bodies, road networks, buildings as per training-sets or designed models like decision tree or random forest for analysis.

All GIS software use several predefined algorithms and also give opportunity to frame own algorithm to user. There are numerous standards bodies and specification programmes like Open Geospatial Consortium Web Services (OGC Web Services) provided models for metadata documentation; Web Feature Service (WFS) provides service-based vector information as feature collections; shapefile of ESRI used in spatial data domain; GeoRSS (RSS: Really Simple Syndication) used in Google Map; GeoJSON (JSON: JavaScript Object Notation) used in mapping and GIS packages for relating objects with location information; Tiled Maps used in small images; GML...
(Geography Markup Language) used in modelling and storing spatial and non-spatial geographic information; GDAL (Geospatial Data Abstraction Library) used as a translator library for geospatial raster data formats (Pasquale et al., 2012). It is therefore there are AI and ML methods/programmes for remotely sensed spatial data labelling and clustering, supervised/unsupervised classification, change detection, GIS analysis and like these for so many others. From the 1990s, artificial intelligence meets GIS mainly through Artificial Neural Networks (ANNs) and fuzzy logic for image classification and feature extraction, to find out spatial solution to decision making, for predictive modelling of geographic phenomena distributions, network analysis and predictions, etc., though the potential of artificial intelligence in GIS is not fully developed (Voženilek, 2009).

WebGIS

With the advancement of data science and internet communication technology from the last few years of the twenty-first century advance form of cartography or GIS available on the web platform is WebGIS, also known as Web Mapping. Simply Web based GIS is WebGIS for sharing data and observation. The Internet GIServices as Kim and Tsou (2013) mentioned is somewhat synonymous to WebGIS. The WebGIS is at present widely used in academic fields from school level to universities and also in several non-governmental sectors for data storing, analysis and mapping on requirement basis for commercial uses. Following remarks on internet GIS as made by Peng and Tsou (2003) it can be said that the WebGIS is a powerful virtual platform to organise geospatial data, web maps, and online geospatial services for studying and understanding the digital earth. As Tsou (2004) mentioned web based GIS is related with sharing data, information and knowledge providing three levels of services like data archiving and searches, information displays and queries, lastly spatial analysis and GIS modelling services. It is associated with components like client, internet service, web server for handling the requests from web browser (user) and return to the web page, and data from GIS server, data server and map server. Thus the WebGIS involves web browsing based client–server interaction system, at least a GIS server on the one end and on the other side the client is a web browser, computer application, or mobile application, including cloud based software as a service. As Alesheikh, Helali and Behroz (2002) mentioned, it is the visual interaction with web bigdata involving many geographical datasets where client using web server prepares maps, diagrams, charts, and so many things and those productions can be easily accessed to anyone, anytime, from anywhere. Pasquale et al. (2012) prepared the development of WebGIS platforms for tourism promotion and heritage conservation. ESRI ArcGIS Online and QGIS are very much remarkable WebGIS platforms today. EDINA Digimap OpenStream is a source of free web maps services to the academicians (Faragher, 2018). The WebGIS is one form of modern GIS refers to the pattern of web services delivering data and capabilities and connects components (ESRI, 2016). Based on ESRI (2016) it can also be said that it is a technological agglomeration of geographical location, bigdata, cloud computing, smart devices, internet of things, virtualisation, consumerisation, distributed processing, and faster computing. Grid computing and cloud computing are high-performance computing resources for digital earth simulation models to solve complex geospatial problems (Kim & Tsou, 2013). Maximum web based GIS models are implemented through grid computing (Kim & Tsou, 2013; Wang, 2010; Wang & Liu, 2009). Cloud computing is generally used by web service providers and businesses (Kim & Tsou, 2013). Kim and Tsou (2013) analysed internet GIServices and parallel computing under different GIS models e.g. terrain triangulation, local spatial interpolation, terrain analysis. The present growing acceptance of WebGIS than desktop GIS is due to several advantages of it like worldwide access by a large number of trained or untrained public users easily at a time using computers and mobiles with web browsers like Internet Explorer, Mozilla Firefox, Google Chrome and so on with low cost as averaged by the number of users, delivering real time information and updates with diverse applications like mapping places and features of interest, tagging photos, and so on as described by ArcGIS Enterprise in its website (ArcGIS Enterprise, 2020). In WebGIS programming different languages are used e.g. HTML, CSS, and JavaScript as client-side languages, whereas, there are object-oriented server-side programming languages like ASP.NET, C#, php, Java, Python and Ruby (Swift & Goldberg, 2019).
Present trend of geographical study

Haden (2008) prepared a brief history of the use of the term neogeography. He mentioned that this term had been used in 1922 to define something new emerging field of geography. Based on the definition of neogeography as Rana and Joliveau (2009) mentioned, the present time is characterised by a new geographical field which brings together several ideas and tools to represent some geographical information. These are internet development and stage like Web 2.0, web development and applications, public involvement and participation, social networking system, VGI, crowd-sourced big data, user-generated theme and content, open source maps API and appropriate navigational devices, smart phones, principally GPS data loggers. Today in many location based services tools and techniques of neogeography are being used (Rana & Joliveau, 2009). Considering the term neogeography and its history of use it can be said that it is basically to understand as a new emerging field. In 1960s and 1970s, the maps were generated with CAD (Computer Aided Design) tools (Subasinghe & Wijesekera, 2008). Presently in geography a drastic change is found in cartography especially to map making and broad scale regional analysis based on remotely generated surveying data, and GIS which introduced in 1980s as a second generation geographic data model (Subasinghe & Wijesekera, 2008). GIS simulation using high-performance computing is regarded as a significant concern in digital earth (Goodchild, 2008). Sui, Goodchild, and Elwood (2013) discussed a detail on VGI and exaflood of bigdata. Bigdata sources transfer geography to a dynamic, more complex, interrelated, extensive data-driven discipline. Virtual web-data-centric geographical study is concerned to geographical bigdata which are generating after digital transactions and online interactions helped to build various types of maps of the world as well as improving maps to present where now we are and where we would like to be (González-Bailón, 2013). As Jiang (2015) described in this twenty-first century bigdata has enormous influence on social science though social science has joined by a little late to bigdata (Linnettaylor, 2012) and especially bigdata in relation to geographic features, their forms, distribution and heterogeneity. As Kitchin (2013) remarked that bigdata drives data-scare situation to data-rich studies with low-cost, detailed, interrelated, and timely available massive amount of unstructured data and information. As positivism describes bigdata opens a new paradigm of computational geography in the present time of data-driven webdom. After proper data curation as well as validation, management and analysis of massive acquired bigdata from various sources, geographers may deal spatial issues under environmental geography, urban geography, social and cultural geography, economic geography, political geography, and so on. Statistical analytical tools use both structured and unstructured bigdata to understand pattern, trend and relationship among different spatial parameters today e.g. business and marketing trend of different things and economic environment of an area, socio-cultural convergence and divergence using social media created data, online shopping originated data to understand urban environment, rural–urban dichotomy using social media created images and videos. In this context as Farshid et al. (2018) described the concept of mixed reality helped to understand future possibilities of current real world scenarios with the help of technological applications. Nowadays automatically generated sensors data having high spatial resolution and technologically modified are to analyse landuse/landcover, regional pattern of different concerns, physical environment and human establishments. Much structured directly uploaded bigdata coming from myriad sources of publications, newspapers, documents, websites, and so many other sources. These at present transferred collection pattern of secondary data replacing hardcopy sources while there are many disadvantages regarding sometimes reliability. Any kind of current data based spatial analysis and mapping at present easy to all dealing with digital data collecting through crowdsourcing processes. All directed, automated and volunteered data transform spatial analysis processes from laborious man induced processes to artificial intelligence based machine learning processes using digital data. If data source is bigdata type of cloud based then analysis would be frequent, low-cost, user friendly, custom based, very easy to present before all. Thus bigdata are undoubtedly becoming a key part of geographic study as Kitchin (2013) once rightly concluded.

In reality to collect geographical data and information there are different governmental and non-governmental departments, cartographic agencies, libraries, and several commercial agencies also. Today mainly
to study physical space of the earth surface people are not only depend on different offices for geospatial data and information collection in reality when this type of job is very much complex following through several rules and obligations and finally time consuming. In the context of digital transformation of real world and its management Pasquale et al. (2012) opined that development and adoption of standards for spatial data acquisition and exchange offers advantages like portability, inter-exchange and easy maintenance. Though, many geodata in soft version from different national and international level agencies are not free of cost rather sometimes high charges along with a long process of permission and copyright restrictions. Now it is the webdom era of the ubiquitous internet, social media and free software along with exponentially increasing number of smart phone users. From the very beginning of the twenty-first century access to spatial data and cartographic products has been changed radically (Jokar Arsanjani, Zipf, Mooney, and Helbich 2015). Today different types open source data have wide impact on geospatial study and applications (Fargher, 2018). With the advancement of internet development as Web 2.0, cloud computing, machine learning and several cyber infrastructures a new trend in geographical study can be ascertained as described by Sui, Goodchild, and Elwood (2013) “ranging from crowdsourcing bigdata to user-generated content, from geoweb to the semantic web, from volunteered geographic information to neogeoography, postGIS, citizen science, and eScience, the general idea coalesces around the use of the internet to create, share, and analyse geographic information via multiple computing devices/platforms (traditional desktops, iPads, or smart phones)””. Batty (1997) wrote on a new dimension to geography i.e. ‘virtual geography’, a transforming form of real geography invocating different types of it as cspace—the space within computer, cyberspace—use of computers for communication purposes, and cyberplace—the infrastructure of the digital world.

The twenty-first century is considered as the time of bigdata and data-driven geography (Miller & Goodchild, 2014) where based on Hey, Tansley, and Tolle (2009) it is presumed that the present exaflood of bigdata drives geographical researches far away from depending on traditional careful observation by individuals during field survey, theory development and statistical simulation to new abundance of digital data. Present virtual web-data-centric geographical study is a drastic change in the system of geographical education and research involving abundance of digital data and web services like ‘socialising spatial information’ (Gordon and de Souza e Silva, 2011) as in the cases of OSM, Google Map, Bing Map, and ‘spatialising social information’ (Gordon and de Souza e Silva, 2011) as in the cases of the maps of the 2020 Covid-19 outbreak throughout the world.

In the context of telecommunication technological rising and its relation with geography from socio-cultural perspective Gordon and de Souza e Silva (2011) revealed the concept of ‘net locality’ (or networked locality) which describes the relationship between increasing location aware technologies and the changing conditions of contemporary societies (Jason Vincent, 2014). Nowadays all mobile phones have location specific settings enabling GPS. Using different mobile apps users can share others their current position. Enabling geographical latitude and longitude based location of anybody transportation apps like Uber and Ola make socio-economic revolution to a city. Google Maps are widely used to view several/alternative journey routes with lot of navigational information at real time position. WhatsApp is also used to share current location to others by messaging along with all navigational information as well. There are location based social networking applications like Google Latitude, Loopt and Bright-Kite to broadcast users location and other contents wherever they move to friends, relatives, and others. These are also used for social surveillance and monitoring people. Fusco, Michael, Aloudat, and Abbas (2011) made a study on monitoring real time position of people using location-based social networking anytime and anywhere and its negative impact on trust which is significant in geographical study combining physical world, virtual world and socio-cultural aspects of human society. The Google Earth is widely used now from its inception (2001) for georeferencing to non-referred maps, collecting geographical data and information like current location and boundary of an area, change detection of drainage system, forests, transport lines, and so on. From anywhere anyone can detect different types of physical information of the earth surface in anytime without physical survey the area. It brought a major change in geospatial data collection process and
hazardous field survey, comparison, analysis, trend and prediction along with mapping as well.

**Maintaining quality of geographical information**

When huge and variety of geographical information are coming from different sources through online in the form of Google Earth, Google Earth Engine, OSM, Google Map, Bing Map, videos, photographs, and so many then question arises how much these information are reliable? Kresse and Fadaie (2004) discussed the role of the International Organization for Standardisation (ISO) toward standardisation in GIS. The ISO 19100 family of standards stands for geographic information/geomatics. Haklay (2010) remarked on methods of evaluation of the quality of geographical information based on the work of Kresse and Fadaie (2004) and thereafter those were synthesised by Van Oort (2006) as lineage, positional accuracy, attribute accuracy, logical consistency, completeness, semantic accuracy, usage, purpose and constraints, and temporal quality. The quality of VGI is in question in terms of its consistency (Haklay, 2010).

**Conclusion**

Throughout the world from the very beginning of twenty-first century a major systematic change is observed in the geographical study practically in the field of geographical education and research system. Geographical study especially known to all for its hard-core field survey based analysis and map making. But nowadays taking the advantages of digital technologies and online web sources of data and information all are easily moved to geographical study avoiding laborious traditional field survey based data acquisition process and also hazardous non-digital mapping as well. Geographical study somewhat shifts from non-digital form to virtual web-data-centric geographical study with all-round digital transformation. Virtual web-data-centric geographical study is centrally concerned with digital data and web services. It depends on easily available huge online collection of spatial and non-spatial data and information. It involves digitised real world data and their virtual storing, their artificial intelligence based machine learning for spatio-temporal analysis, trend and prediction, their different spatial and thematic mappings in different geospatial web service compartments. In digital earth everything is almost numerically tends to be measurable. Here data (training-set) analyses data (raw-data) and again creates data (research finding) in conclusion which further can be used again as data for measuring another data-set. At present in the age of webdom—a global digital kingdom of internet and web services, through continuous up-gradation of web searching and internet communication network management along with increasing number of internet users and GPS enabled smart phones geospatial data and information are generated by exaflood leading to enormous cloud based geographical bigdata. Today virtual web-data-centric geographical study offers many opportunities to academic and social fields along with different challenges. Throughout the world, under the exponential growth and advancement of digital science and digital technologies like other natural and social sciences geography is also now predicting different phenomenon based on virtual geographical throughputs. The three main facets of it are geographical bigdata, geographical machine learning, and WebGIS. Geographical bigdata is highly crowded now. Geographical machine learning helps to geographical data management and analysis. The WebGIS is modern-time, user-friendly, problem-specific online service of complex geospatial data management and web mapping. Sometimes reliability of available virtual web-data is in question without physical field verification which is considered as the primary requisition of traditional geographical study.

**Acknowledgements**

The author thanks the editor(s) for acceptance of the paper and also reviewers for their valuable comments. The author is also greatly thankful to the DST-PURSE II programme, University of Kalyani, Kalyani, Nadia, West Bengal, India, for also some financial assistance in the processing of writing of the research paper.

**References**

Abdo, A.M., & Alali, S.H. (2016). Comparing common programming languages to parse big xml file in terms of executing time, memory usage, CPU consumption and line number on two platforms. *European Scientific Journal, 12*(27), 325–334. https://doi.org/10.19044/esj.2016.v12n27p325
Alesheikh, A.A., Helali, H., & Behroz, H.A. (2002). Web GIS: technologies and its applications. Symposium on geospatial theory, processing and applications. Available at https://www.isprs.org/proceedings/XXIV/part4/pdfpapers/422.pdf. Accessed on 30-04-2020.

Aymerich, F.M., Fenú, G., & Surcis, S. (2008). An approach to a cloud computing network. *IEEE*. https://doi.org/10.1109/ICADIWT.2008.4664329. Accessed on 02-05-2020.

Batty, M. (1997). Virtual geography. *Futures*, 29(45), 337–352.

Cairncross, F. (1995). The death of distance: A survey of telecommunications. *The Economist*, 336(7934), 5–28.

Che, D., Safran, M., & Peng, Z. (2013). From big data to big data mining: challenges, issues, and opportunities. In B. Hong, et al. (Eds.), *DASFAA Workshops 2013: LNCS* (vol. 7827, pp. 1–15). Springer.

Diwanji, S. (2020a). Number of internet users in India from 2015 to 2018 with a forecast until 2023. Available at https://www.statista.com/statistics/255146/number-of-internet-users-in-india/. Accessed on 27–04–2020.

Diwanji, S. (2020b). Number of social network users in India from 2015 to 2018 with a forecast until 2023. Available at https://www.statista.com/statistics/278407/number-of-social-network-users-in-india/. Accessed on 27-04-2020.

ArcGIS Enterprise. (2020). https://enterprise.arcgis.com/en/server/latest/create-web-apps/windows/about-web-gis.

ESRI. (2016). Web GIS, Simply. Available at https://www.esri.com/about/newsroom/insider/web-gis-simply/. Accessed on 01-05-2020.

Farshid, M., Paschen, J., Eriksson, T., & Kietzmann, J. (2018). Go boldly!: Explore augmented reality (AR), virtual reality (VR), and mixed reality (MR) for business. *Business Horizons*, 61(5), 657–663. https://doi.org/10.1016/j.bushor.2018.05.009

Fusco, S.J., Michael, K., Aloudat, A., & Abbas, R. (2011). Monitoring people using location-based social networking and its negative impact on trust: an exploratory contextual analysis of five types of “friend” relationships. In: *ISTAS 2011: IEEE International Symposium on Technology and Society, Illinois: IEEE pp. 1–11.

González-Bailón, S. (2013). Big data and the fabric of human geography. *Dialogues in Human Geography*, 3(3), 292–296. https://doi.org/10.1017/S0262325113515379

Goodchild, M. F. (1999). Future directions in geographic information science. *Geographical Information Sciences*, 5(1), 1–8. https://doi.org/10.1080/1082409990480507

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, 69(4), 211–221. https://doi.org/10.1007/s10708-007-9111-y

Goodchild, M. F. (2008). The use cases of digital earth. *International Journal of Digital Earth*, 1(1), 31–42. https://doi.org/10.1080/17538940701782528

Goodchild, M. F. (2009). Geographic information systems and science: Today and tomorrow. *Annals of GIS*, 15(1), 3–9. https://doi.org/10.1080/19475680903250715

Goodchild, M. F. (2010). Twenty years of progress: GIScience in 2010. *Journal of Spatial Information Science*, 1(2010), 3–20. https://doi.org/10.5311/JOSIS.2010.1.2

Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial Statistics*, 1(2012), 110–120. https://doi.org/10.1016/j.spasta.2012.03.002

Goodchild, M. F., Yuan, M., & Cova, T. J. (2007). Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science*, 21(3), 239–260. https://doi.org/10.1080/1365881060965271

Goodchild, M.F. (2019). GIS in the era of big data. Cybergeo: European Journal of Geography. http://journal.openedition.org/cybergeo/27647

Gordon, E., Adriana de Souza e, S. (2011). *Net locality: why location matters in a networked world*. Wiley-Blackwell.

Gore, A. (1998). The digital earth: Understanding our planet in the 21st century. *Australian Surveyor*, 43(2), 89–91. https://doi.org/10.1080/00050348.1998.10558728

Gould, P. (1999). Becoming a geographer. Syracuse University Press.

Graham, M., & Shelton, T. (2013). Geography and the future of big data, big data and the future of geography. *Dialogues in Human Geography*, 3(3), 255–261. https://doi.org/10.1007/2043820613513121

Hey, T., Tansley, S., & Tolle, K. (Eds.). (2009) *The fourth paradigm - data-intensive scientific discovery*. Microsoft Research, Redmond, Washington.

Jason Vincent, A.C. (2014). Net locality: why location matters in a networked world, by Gordon, E., & Adriana de Souza e, S. *Popular Communication: The International Journal of Media and Culture*, 12(1), 69–71. https://doi.org/10.1080/15405702.2013.838468

Jiang, B. (2015). Big Data is not just a new type, but a new paradigm. *Computers, Environment and Urban Systems*, 53, 1–122.

Jokar Arsanjani, J., Zipf, A., Mooney, P., & Helbich, M. (2015). An introduction to OpenStreetMap in Geographic Information Science: experiences, research, and applications. In J. Jokar Arsanjani, A. Zipf, P. Mooney, & M. Helbich (Eds.). *OpenStreetMap in GIScience. Lecture notes in Geoinformation and Cartography* (pp. 1–18). Springer. https://doi.org/10.1007/978-3-319-14280-7_1

Kaplan, A. D., Cruise, J., Endsley, M., Beers, S. M., Sawyer, B. D., & Hancock, P. A. (2020). The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis. *Human Factors*. https://doi.org/10.1177/0018720820904229

Kaplan, R.D. (2009). The revenge of geography. *Foreign Policy*, no 172 (May/June 2009). Washingtonpost, Newsweek Interactive, LLC, 96–105. Available at http://www.jstor.org/stable/20684874
