GIS for Collaborative Research

Piyawan Kasemsuppakorn*

Abstract

This article presents the importance of collaborative research and the growing trend of using Geographical Information Systems (GIS) as a mediator of spatial knowledge, social, and intellectual practice. Collaborative research projects generally involve researchers from multiple disciplines either in the same institution or between institutions toperform research. It is driven by complex problems requiring multidisciplinary solutions and the increased visibility provided by cross disciplinary research. With the rapid advancement of GIS, there is a substantial increase in the use of dynamic maps and advanced spatial functionsfor multidisciplinary research, such as environmental and urban planning. The distributed GIServices is considered as an appropriate infrastructure for developing a pervasive tool that supports data integration, data sharing, and data visualization from anywhere and anytime. The enabling technologies for distributed GIServices are discussed and three existing platforms related to web-based GIS that supports collaborative research: WorldMap, MapStory, and WISERD are illustrated.

Keyword: Collaborative Research, GIS, Data Integration, Web Mapping Services.

1. INTRODUCTION

Collaborative research refers to a research effort done by research teams from multiple disciplines (interdisciplinary) and the collaboration can take various forms ranging from offering general advice and insights to active participation in a specific piece of research[1].Collaborative research is a growing trend for several reasons: 1) the real-world problems associated with nature and society are complex to which one discipline on its own cannot provide a solution[2]; 2) collaborative efforts between researchers and non-academic stakeholders promises to increase legitimacy, ownership, and accountability for the problem, as well as for the solution[3]; 3) the desire for researchers to increase their scientific popularity, visibility, and recognition[4]; 4) improving the links between science and technology; and 5) the need to gain experience or to train new researchers in the most effective way. With these reasons, there are currently an increasing number, size, and diversity of collaborative research projects between universities and industries and also between countries. For example, environmental epidemiology research requires collaborative actions for better exposure assessments between scientists and for larger studying contrasts in exposure and health profiles between various geographical areas[5].

In order to produce a successful collaborative research project, not only are the intellectual and curiosity, conceptual ability, competency, and creativity characteristics important, good communication among discipline-specific individuals is also crucial. Therefore, a mediator platform is needed. This platform would need to enhance communication as well as enable data integration from various data sources and formats. Additionally, there is a growing volume of researchthat includes geospatial referencing data such as geographic coordinates, addresses, and postal codes. Integrating and presenting this data presents a challenge. Fortunately, there is a solution provided by modern Geographic Information Systems (GIS).Dynamic maps, images, and advanced spatial functions in GIS have all been used as mediators of collaboration in multidisciplinary research in areas such as environmental and urban planning, resource management, health care services, social science, business planningand education [6].

GIS is a computer system capable of assembling, storing,

School of Science and Technology University of the Thai Chamber of Commerce

manipulating, analyzing, and displaying geographic data and information [7]. GIS is considered as a powerful tool and has been widely applied in a variety of purposes including integrating information from multiple sources (e.g., see Resch et al., [8]), solving operational, management, and planning problems (e.g., see Chandrasekar et al., [9]), extracting useful information from such data (e.g., see Shi [10]), and describing and characterizing a place of location in the earth (e.g., see Crnovrsanin et al., [11]). Nowadays, GIS technology has not only become a part of everyday life, but it also has become an extremely important platform for undertaking a range of research areas and enabling new ways to do research. In a GIS context, the attempt to utilize GIS technology for research teams in the production of GIS data, the creation of knowledge, and spatial decision making is called "Participatory GIS" (PGIS).

For example, the study on the ecology, distribution and conversation status of species and ecosystems requires the aid of GIS and remote sensing to predict habitat distribution [12]. This study employed GIS technology as a mediator to collaborate between the environmental specialists and GIS professionals. The environmental specialists studied the environmental risks for terrestrial systems of forest ecosystems and animals. The results from the study were stored in a shared database and available through the web for communication with other departments. Biggs et al [13] integrated different disciplines including environmental science, Information science, and GIS to analyze and present the effects of deforestation, intensive agriculture, regional development on stream chemistry, and urbanization. Several data sources, such as census, soils, Landsat, and stream network were integrated using GIS technologies.

This article highlights the GIS technologies, the importance of PGIS, and a suitable framework for supporting collaborative research. The structure of the paper is as follows. Section 2 provides the fundamental of GIS and discusses its evolution and Section 3 provides details of GIS technologies for enabling collaborative research. Section 4 presents exist-ing web-based GIS platforms for collaborative research. Conclusions are discussed in Section 5.

2. GIS AND ITS EVOLUTION

GIS, a special type of information system dealing with geographically referenced data, comprises of fourmain components: hardware, software, data, and users[7]. Hardware is the computer on which a GIS operates. GIS Software incorporated with database management system (DBMS) provides functions and tools for managing geographic information, performing spatial analysis, and displaying geographic information in the form of maps or reports. Anexample of professional GIS software is ArcGIS that supports relatively strong data editing and analytical functions as well as providing connections to many remote data sources compliance with standard geospatial data sharing protocols. However, it is expensive to implement and requires substantial training to customize and maintain the system. With the increasing popularity of the Internet, public web-based GIS systems, such as Google Map and Bing Map are open to public and provide a fixed set of base maps and few analytical functions (e.g., geocoding, and traffic routing). Although it is simple and easy to use, it cannot handle large volume of data and requires programming skills in order to connect to other data services and combining data layers on the maps. Therefore, there is a need for developing an internet mapping systems that combine the strengths of professional GIS software and public web-based GIS systems.

Conventionally, geographic data containingspatial features and their corresponding attribute information are collected by cartographers and geographers by using advanced specialized technologies and are disseminated by government agencies (e.g., Ministry of Transport, Thailand), non-profit organizations (e.g., Department of Survey Engineering, Chulalongkorn University), and commercial mapping companies (e.g., TeleAtlas). These geographic data are generally subject to strict copyright laws. With the lack of publicly available geographic data and the rapid development

of mobile positioning technology,spatial data are collected and contributed for free by volunteers, called Volunteered Geographic Information (VGI)[14]. OpenStreetMap (OSM) and WikiMapia are extensive and effective projects that currently facilitate access to collaboratively collected map data for the whole world.

The development of GIS has been highly influenced by the development of computer technologies. It evolved from centralized GISystems to distributed GIServices[15]. The major difference between them is that distributed GIServices works through both wired and wireless networks while the centralized GISystems only works on wired networks. Obviously, the distributed GIServices has more accessibility than the centralized GISystems, as shown in Figure 1.
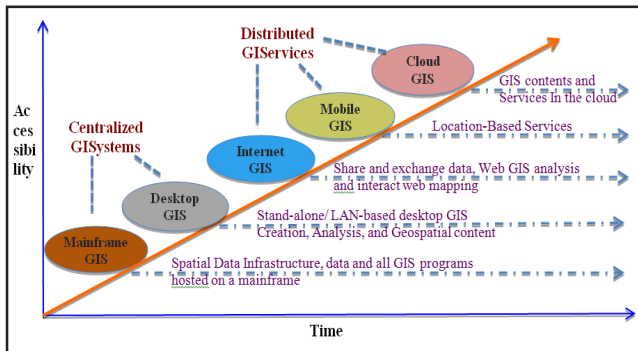

Figure 1. GIS evolution

Mainframe GIS and desktop GIS are categorized as the centralized GISystems that could be a stand-alone or a client/server system. The stand-alone GISystems contains all the GIS programs, interfaces, and spatial data in one computer. A client/server GISystems is based on generic client/server architecture in which the server side stores the spatial database and GIS programs and distributed clients can access a server remotely through the desktop computers. Thus the GISystems is generally platform dependent, which means that each client can access only one specified server at one time[15].

Advancements of wireless communications and the increasing popularity of the Internet have changed the way of conducting GIS processing systems. Rather than relying on desktop GIS programs, users can access the GIS analysis functions and spatial databases from anywhere and anytime with the Internet access and wireless network communications, called distributed GIServices. There are three categories of distributed GIServices: Internet GIS, mobile GIS, and cloud GIS. Internet GIS employs Internet technologies to access remote geographic information and geoprocessing tools, and to distribute maps, images, data sets, and reports. The clients for Internet GIS are mostly desktop computers without GIS software. The launch of MapQuest in 1996for the first time brought interactive street maps to the public on an Internet web browser [16]. Mobile GIS refers to the access and use of GIS data and functions through smart phones and wireless devices, such as tablets. With the improvement of Global Positioning Systems (GPS) and wireless communication technologies, mobile GIS plays an important role in real-time access to spatial data and location-based services (LBS). The clients for mobile GIS are mostly smartphones, GPS-enabled phones, and tablets. Cloud computing, an emerging trend, defined by Buyya et al., [17] as "a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements". With cloud technology, rather than running GIS programs on clients' systems, all GIS programs and data reside on cloud infrastructures maintained by a third party, such as Amazon Web Services , Microsoft Azure , and IBM Cloud and are available for users through web technologies.

In sum, distributed GIServices enhances the accessibility and reusability of GIS functions and spatial data from different data providers. The Internet, wireless communication, and cloud technologies enable dissemination of spatial information to a much wider range of clients than do centralized GISystems. GIS users and non-GIS users can now directly explore and request spatial information from their web browsers at home and everywhere on their mobile devices. Also, Internet and wireless communication technologies alter the hardware and software requirements for GIS access.

3. GIS TECHNOLOGIES FOR ENABLING COLLABORATIVE RESEARCH

As stated previously, collaborative research involves researchers and professionals from two or more disciplines either in the same institution or between institutions. To support collaborative research, a mediator is required in order to ensure a continuous and interactive communication among research team members throughout the researching process. It has been proved that GIS is a competent and effective tool to align research approaches and facilitate collaboration. Three key unique features of GIS are data integration, data sharing, and data visualization. Data integration refers to the process of combining heterogeneous data residing in different sources, data models, and scales. It requires the standards for interoperability among GIS and geographic information technologies themselves, and between GIS and other information technologies. Data sharing feature allows multiple users to simultaneously view and manipulate data using a central spatial data storage. Data visualization or geovisualization allows research team members to explore complex geographic data, perform geospatial data analysis, and view the analysis results through the use of interactive medium.

Among several types of GIS infrastructures, distributed GIServices is a suitable infrastructure for developing a pervasive mediator supporting collaborative research. This is because of the potential to bring GIS to the communities. Distributed GIServices is cross-platform and interoperable meaning that it is not limited to any kind of machine or operating system. With the rapid development of wireless communication technologies, it enables users to manipulate GIS data and maps interactively over the Internet or wireless networks and the users does not need to have geographic data and GIS software installed in their own machines.

To be able to access and share remote GIS data and functions in the heterogeneous environment, it requires high interoperability and standards for distributed GIServices. The Open Geospatial Consortium (OGC), Inc. [18], an interna-tional worldwide organization, set industry standards for the full integration of geospatial data and geoprocessing resources into mainstream computing and the widespread use of interoperable geoprocessing software. OGC's Simple Feature Specification is a standard for storing, retrieving, and updating simple geospatial features. Many Database Management Systems (DBMSs) that employ and follow OGC's standards include Oracle Spatial, Microsoft's SQL Server, ESRI's ArcGIS Geodatabase, and the Postgres extension PostGIS. Also, today's DBMSs enable efficient management of geographic data by supporting spatial data attribute types, spatial operations in query language, and spatial indexing methods

Clearly, World Wide Web (WWW) is currently used as a main platform for most applications including collaborative research projects to access remote data, process components, and display information. However, the traditional web-based technology using HTML does not support geographic data and spatial functions. Therefore, OGC defined a specific standard language for expressing geographical features, called Geography Markup Language (GML). GML is "an XML encoding for the transport and storage of geographic information, including both spatial and non-spatial properties of geographic features"[19]. Using GML, users can design their own markup elements to describe spatial and non-spatial features, geometries, and spatial reference systems, which enable efficiently data sharing, and easily interchange over the web. GML is based on XML; thus other languages (e.g., Java, Javascript, C++. VML, or SVG) that work with XML will also work with GML. A variety of distributed GIServices can be developed with specific functions based on GML, such as routing, tracking, spatial analysis, and spatial information discovery. However, to execute spatial operations, other programming languages, such as C++ and Visual basic are needed for data processing. To present geographic feature, the map styling process is also required to interpret the GML contents [15].

4. EXISTING WEB-BASED GIS PLATFORMS FOR COLLABORATIVE RESEARCH

This section provides a brief overview of existing web-based GIS platforms for online spatial data sharing and web mapping services. Currently, there has been significant improvement in both commercial and public online platforms. The commercial platforms (e.g., ESRI's ArcGIS online and GeoCommons) provide powerful and complex spatial analysis tools. However, they generally only offer a limited version for free and require substantial training to customize and maintain. On the other hand, there are web-mapping developer communities launched open-source web-based GIS platforms that combine the strength of the commercial platforms with the public systems' ease of use and open accessibility. At the time of writing, several opensource web-mapping platforms for collaborative research include WorldMap, MapStory, and WISERD.

WorldMap is a web-based and map-centric data exploration system developed by the Center for Geographic Analysis (CGA) at Harvard University. It is designed to provide a means of creating an information environment for collaborative research, teaching, and general use[20]. WorldMap provides interactive tools for building map portals with customcontents overlaid on abase map from a variety of sources including Google, MapQuest, and Bing. The interactive tools include features of data integration, data sharing, and data visualization. Data contents on the WorldMap system accumulate as both institutional and individual users upload their data for their own project use, and at the same time contribute them to the WorldMap geodatabase for others to share.The WorldMap system is built on open-source geospatial technologies, such as GeoNode, OpenLayers, PostGIS, and GeoServer, in order to create an interoperable geospatial system. Its platform is hosted on Amazon's EC2 cloud infrastructure. GeoNode is a Django-based framework providing the foundation for feature symbology, map composition, metadata management, and user profile administration. The PostGIS libraries are used for storing spatial features and the OpenLayersJavascript

library is employed to connect to other proprietary mapping services such as Google maps and Bing maps. The WorldMap system was released as a beta in July 2011 and WorldMap version 1.4 was released in January 2014 available for non-GIS academics and researchers to discover, visualize, investigate, and communicate their research materials in a spatial framework. As of August,2014, WorldMap had 11,600 registered users; 13,665 data layers added by users, containing 111,486 data fields; 3,771 map collections created; and 780,000 unique visitors coming from every country in the world. It had about 1,000 unique visitors per day, spending 1-2 minutes each on average [21].
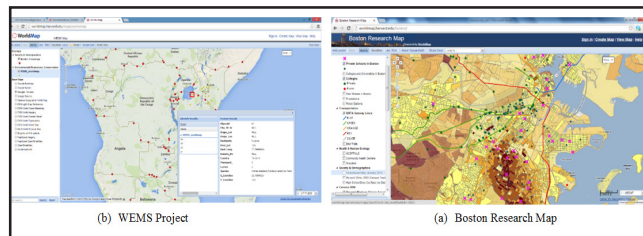

(b) WEMS Project   (a) Boston Research Map

Figure 2. *Examples of Collaborative Projects using the WorldMapplatform.*

Several collaborative research projects using the WorldMapplatformare related to historical, social, economic, cultural, and environmental. The Widelife Enforcement Monitoring System (WEMS) is a collaborative project carried out by United Nations University, private industries, and research institutions. The objective of WEMS is to assist monitoring trafficking and illegal widelife crime. Through WorldMap, the system allows users to map and track widelife crime from WEMS database and to understand how it is related to socioeconomic from the WorldMap data.Another project is Boston Research Map. It is one of historical collaborative research projects, created by Rappaport Institute for Greater Boston at the Harvard Kennedy School and the Radcliffe Institute for Advanced Study. The Boston Research Map portal contains several data layers of both contemporary and historical data supplied by researchers and make is accessible online. Examples of data layers are society and demographics, census 2010, transportation, health and human
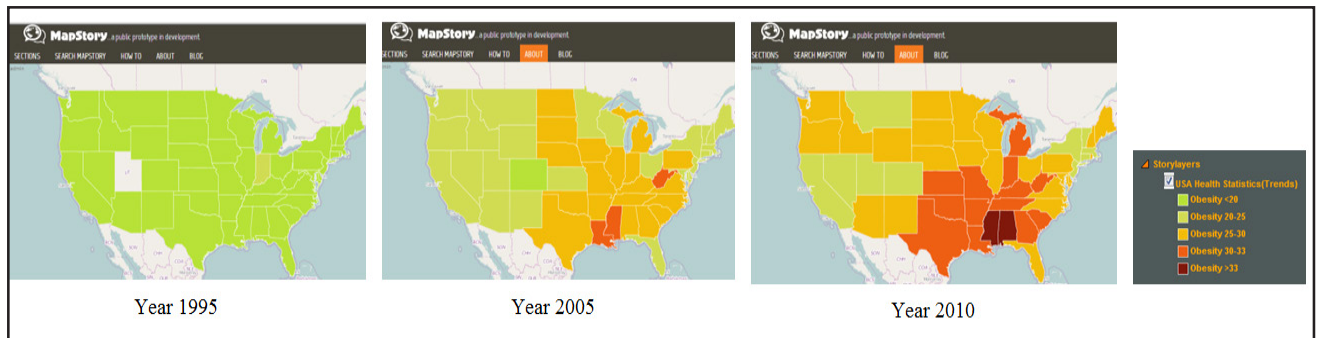
Figure 3. *The study of the obesity in the states using the MapStory platform*

ecology, and historical of the Boston city area. Figure 2a shows the border crossings between African countries and the widelife species in each location. Figure 2b presents the percentage of African American in Each district area, where the dark red presents the highest percentage and the light yellow presents the lowest percentage.

MapStory is an online social media platform that enables a community of experts to organize knowledge about the world spatially and temporarily.Its platform is built entirely on open source geospatial softwares (e.g., GeoNode); therefore it is completely open and free to use for educators, researchers, and practitioners. With MapStory, researchers are able to conduct interdisciplinary research in a variety of studies. For example, a research study of obesity trends in the United States used the MapStory platform to communicate between the researchers from different disciplines (e.g., public health and information sciences)[22].Figure 3shows the obesity of state populations over time from 1995 to 2010. Based on this MapStory, we can see the increasing trend of obesity in all states, especially people from the southern states.

The Wales institute of social and economic research, data and methods (WISERD) initiates a project involving five higher education institutions in Wales (e.g., Aberystwyth, Bangor, Cardiff, Glamorgan, and Sawnsea Universities), aiming to enhance collaborative socio-economic research through the Wiserd Geo-portals (WGP)[23]. Several expertise in criminology, economics and finance, geography and GIS, public health, urban studies, political science, sociology and social policy join the research and provide both quantitative and qualitative socio-economic research data. Examples of stored quantitative data are household, British election, labor force, and British crime survey. Examples of qualitative archive are interview transcript and field notes. WGP allows users to perform spatial, temporal and thematic queries to search for metadata using both dynamic web-based mapping interfaces and text-based interfaces.

WGP was designed and developed based on open-source geospatial software (GeoFoss), providing a strong and sustainable framework for integrating, managing and disseminating quantitative and qualitative socio-economic data.The used GeoFoss components include PostGIS libraries, GeoServer, and OpenLayers. As of August, 2011, there are 63 survey data sets in the WISERD archives, such as labor force survey 2010, living in Wales (2004-2008), Welsh health survey (2003-2011), and working in Britain (1999-2002); 5122 tables were stored in the metadata; and 6700 questions from researchers[24].

5. CONCLUSIONS

Collaborative research refers to a research effort done by research teams from multiple disciplines (interdisciplinary) within the same institution and between institutions. The collaboration can take various forms ranging from offering general advice and insights to active participation in a specific piece of research. The collaborative research is a growing trend because of several benefits of collaboration:

(1) sharing of knowledge, skills and techniques; (2) creating a source of creativity; (3) providing intellectual companionship; (4) extending the professional network in the scientific community; and (5) enhancing the potential visibility of the research. However, collaboration requires a mediator platform enhancing collaborative knowledge among researchers and enabling data integration from various data sources. GIS has been extensively employed in a variety of research fields that involve geographical data, such as environmental, health, and transportation and it has been proved that GIS is an effective tool to facilitate collaboration. GIS infrastructure evolved from centralized GISystems to distributed GIServices based on the development of computer technologies. Distributed GIServices is considered as an appropriate infrastructure for developing a pervasive tool that supports data integration, data sharing, and data visualization from anywhere and anytime. The open-source geospatial technology and the software are deemed necessary for the development of distributed GIServices. WorldMap, MapStory, and WISERD are existing open source web-based GIS platforms that provide a means of creating an information environment for collaborative research and teaching. Users of WorldMapand MapStoryare from all over the world, while WISERD has been widely used in the universities in Wales.

6. REFERENCES

[1]   D. Beaver. "Reflections on scientific collaboration (and its study): past, present and future." *Scientometrics,* Vol. 52, pp. 365-377, 2001.

[2]   A. Bruce, C. Lyall, J. Tait and R. Williams. " Interdisciplinary integration in Europe: The case of the fifth framework programme." *Futures,* Vol. 36, pp. 457-470, 2004.

[3]   J. Spangenberg. "Sustainability science: a review, an analysis and some empirical lessons." *Environmental Conservation,* Vol. 38, pp. 275-287, 2011.

[4]   D. Beaver and R. Rosen. "Studies in Scientific Collaboration: Part II - Scientific Coauthorship, Research Productivity and Visibility in the French Scientific Elite, 1799-1830." *Scientometrics,* Vol. 1, pp. 133-149, 1979.

[5]   M. Kogevinas, A. Nybo Andersen and J. Olsen. "Collaboration is needed to co-ordinate European birth cohort studies." *International Journal of Epidemiology,* Vol. 33, pp. 1172-1173, 2004.

[6]   D. Sui. "GIS, cartography, and the "Third Culture": Geographic Imagination in the computer age." *Professional Geographer,* Vol. 56, pp. 62-72, 2004.

[7]   C. P. Lo and A. K. W. Yeung. Concepts and Techniques of Geographic Information Systems. New York: Pearson Prentice Hall, 2006.

[8]   B. Resch, M. Mittlboeck, F. Girardin, R. Britter and C. Ratti. "Real-Time Geo-awareness-Sensor Data Integration for Environmental Monitoring in the City." *presented at the International Conference on Advanced Geographic Information Systems & Web Services (GEOWS '09),* Cancun, 2009.

[9]   K. Chandrasekar, M. V. R. S. Sai, P. S. Roy, V. Jayaraman and R. R. Krishnamurthy. "Identification of Agricultural Drought Vulnerable Areas of Tamil Nadu, India-Using GIS Based Multi Criteria Analysis." *Asian Journal of Environment and Disaster Management,* Vol. 1, pp. 43-64, 2009.

[10]   G. Shi. "Thematic data extraction from Web for GIS and applications." *presented at the 2011 IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services (ICSDM),* Univ. of Calgary, Canada 2011.

[11]   T. Crnovrsanin, C. Muelder, C. Correa and K.-L. Ma. "Proximity-based visualization of movement trace data." *presented at the IEEE Symposium on Visual Analytics Science and Technology, 2009 (VAST 2009),* Atlantic City, NJ, 2009.

[12]   D. S. Boyd and G. M. Foody. "An overview of recent

remote sensing and GIS based research in ecological informatics." *Ecological Informatics,* Vol. 6, pp. 25-36, 2011.

[13] T. W. Biggs, T. Dunne, D. A. Roberts and E. Matricardi. "The rate and extent of deforestation in watersheds of the southwestern amazon basin." *Ecological Applications,* Vol. 18, pp. 31-48, 2008.

[14] M. F. Goodchild. "Citizens as sensors: the world of volunteered geography." *GeoJournal,* Vol. 69, pp. 211-221, 2007.

[15] Z.-R. Peng and M.-H. Tsou. Internet GIS: Distributed geographic information services for the Internet and wireless networks: Wiley, 2003.

[16] M. P. Peterson. *Maps and the internet,* Vol. 1, Amsterdam: Elsevier, 2003.

[17] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg and I. Brandic. "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility." *Future Generation Computer Systems,* Vol. 25, pp. 599-616, 2009.

[18] OGC. OpenGIS Reference Model. Available online at http://portal.opengeospatial.org/files/?artifact_id=3836 December 2003.

[19] OGC. "GML 2.0, Recommendation Paper." presented at the Open GIS Consortium, Wayland, Massachusetts, 2001.

[20] W. W. Guan, P. K. Bol, B. G. Lewis, M. Bertrand, M. L. Berman and J. C. Blossom. "WorldMap-a geospatial framework for collaborative research." *Annals of GIS,* Vol. 18, 2012.

[21] CGA. WorldMap Usage Statistics. Available online at http://about.worldmap.harvard.edu/icb/icb.do?keyword=k28501&pageid=icb.page480821. September 2014.

[22] MapStory. US Obesity Trends. Available online at http://mapstory.org/maps/180/#. September 2014.

[23] R. Berry, R. Fry and G. Higgs. "Building a geo-portal for enhancing collaborative socio-economic research in Wales using open-source technology." *Journal of Applied Research in Higher Education,* Vol. 2, pp. 78-92, 2010.

[24] R. Berry, R. Fry and S. Orford. Introducing the WISERD Geoportal. Available online at http://www.data-archive.ac.uk/media/338403/inspire7oct2011so.pdf. December 2011.