



DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

WorldMap – A Geospatial Framework for Collaborative Research

The Harvard community has made this article openly available.
[Please share](#) how this access benefits you. Your story matters.

Citation	Guan, Weihe Wendy, Peter K. Bol, Benjamin G. Lewis, Matthew Bertrand, Merrick Lex Berman, and Jeffrey C. Blossom. 2012. WorldMap – a geospatial framework for collaborative research. <i>Annals of GIS</i> 18, no. 2: 121-134.
Published Version	doi:10.1080/19475683.2012.668559
Accessed	February 19, 2015 1:10:08 PM EST
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:11130564
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP

(Article begins on next page)

WorldMap – A Geospatial Framework for Collaborative Research

Weihe Wendy Guan, Peter K. Bol, Benjamin G. Lewis, Matthew Bertrand, Merrick Lex Berman, and Jeffrey C. Blossom

Center for Geographic Analysis, Harvard University
1737 Cambridge Street, Suite 350, Cambridge, MA 02138 USA

wguan@cga.harvard.edu

pkbol@fas.harvard.edu

blewis@cga.harvard.edu

mbertrand@cga.harvard.edu

mberman@fas.harvard.edu

jblossom@cga.harvard.edu

Abstract— WorldMap is a web-based, map centric data exploration system built on open source geospatial technology at Harvard University. It is designed to serve collaborative research and teaching, but is also accessible to the general public. This paper explains WorldMap’s basic functions through several historical research projects, demonstrating its flexible scale (from neighbourhood to continent) and diverse research themes (social, political, economical, cultural, infrastructural, etc.). Also shared are our experiences in handling technical and institutional challenges during system development, such as synchronization of software components being developed by multiple organizations; juggling competing priorities for serving individual requests and developing a system that will enable users to support themselves; balancing promotion of the system usage with constraints on infrastructure investment; harnessing volunteered geographic information while managing data quality; as well as protecting copyrights, preserving permanent links and citations, and providing long term archiving.

Keywords— WorldMap, ChinaMap, AfricaMap, web mapping, GeoNode, geospatial

I. INTRODUCTION

Harvard’s mission is to advance knowledge through teaching and research. One of the primary ways the faculties of its several schools (Arts and Sciences, Engineering, Design, Medicine, Public Health, Business, Education, Divinity, Law, and Government, etc.) to collaborate is through its research centers. It has a long history and expanding set of area research centers: the Fairbank Center (China), Reischauer Institute (Japan), Korea Institute, Davis Center (Russia and Eurasia), Rockefeller (Latin America), Center for Middle Eastern Studies, South Asia Initiative, etc., each with their historical as well as contemporary research agenda.

Seeing the world through a geographic interface is proving to be an intuitive and accessible way of sifting through and visualizing the rapidly growing body of data about all areas of the world, allowing us to see the world as a whole without losing sight of its parts. WorldMap, developed by the Center for Geographic Analysis at Harvard (CGA) together with scholars from some of these area centers, is the University’s means of creating an information environment that serves research and teaching about the world; makes research available to a world public; and facilitates collaboration across institutions.

WorldMap is a web-based and map centric data exploration system built on open source geospatial technology. It is designed to be publicly accessible, simple to use for non-technical scholars, fast in search and mapping, and rich in geographic content. Its first implementation for the continent of Africa, with a beta release in November 2008, proved to be an effective solution for integrating and making accessible dispersed humanities collections and a wealth of mapping materials, allowing for cross-disciplinary

inquiries. AfricaMap has since been cloned to create portals focused on other parts of the world including Boston, Vermont, Paris and East Asia, collectively called WorldMap.

WorldMap is designed to handle an unlimited amount of data content, not only by the scalability of its servers, but more importantly through its service-oriented-architecture which allows the system to connect to remote systems and display materials from them following Open Geospatial Consortium (OGC) and various other open specifications. In return, materials hosted on WorldMap servers are available for other systems' live consumption as well. The system handles data in any format as long as it has a geographic location associated with it. Examples include shapefiles, GeoTIFF images, YouTube videos, Picasa photos, museum collections, library archives, wired news, selected blogs, etc. The user may select several data layers to overlay together, and can control transparency of overlaid layers, allowing for visual investigation of spatial patterns and relationships between data sets. The power of WorldMap is its ability to integrate an ever expanding data collection stored locally and remotely, its flexibility to focus on any particular theme (defined by the filtered data layers), on any geographic region (from the world to a city neighbourhood) and its interactive tools which allow the user to discover, combine, and visualize materials from different sources, in different formats, in different languages, all organized through a map.

The usefulness of the WorldMap system has generated much enthusiasm among a broad user community inside and outside of Harvard University, as well as demand for its expansion and enhancement. Based on experience with a proof-of-concept implementation in Africa and elsewhere, a 2nd generation of the WorldMap system was developed and released as in Beta in July 2011. Version 1.0 is scheduled for release in January of 2012. The new system includes many functional improvements to facilitate greater scholarly collaboration.

WorldMap is made available to any person in any part of the world with an access to the Internet. From any of the common web browsers (IE, Firefox or Chrome), people without prior GIS training can visit the WorldMap website (<http://worldmap.harvard.edu>), browse among the hundreds of maps composed by other users, or freely register for a user account with a valid email address to start composing their own maps using any of the data layers stored in the WorldMap database, or brought in from remote map servers, or uploaded from their own local drives. They can share their maps and uploaded data sets with anyone they desire to share with, from named individuals, to the general public. WorldMap has enabled non-GIS academics and researchers to discover, visualize, investigate and communicate their research materials in a spatial framework, greatly enhancing their ability to conduct academic research or instructional activities.

From July to November of 2011 WorldMap had 11583 visitors (of which 77% were newcomers to the system); 1170 registered users (and many more unregistered ones); 2206 user-uploaded data layers; and 382 custom map collections.

This paper will examine the development of WorldMap from both technical and institutional perspectives, and demonstrate several case projects where WorldMap has been used in support of historical research. The objective is to share our experience in implementing as well as enhancing cutting edge geospatial technologies, in serving academic research as well as harnessing volunteered geographic information for sharing, and in promoting cross-disciplinary collaboration as well as advancing geographic information science. The paper will introduce ChinaMap as a case of WorldMap applied to China studies, and a few other instances of the system to illustrate its current use at Harvard in support of historical geographic research and teaching.

II. THE NICHE FOR AN OPEN MAPPING SYSTEM

The launch of MapQuest in 1996 for the first time brought interactive street maps to the public on an Internet web browser^[1]. The release of ArcView Internet Map Server (ArcViewIMS) in 1998 allowed licensed users to compose their own set of maps and publish them for dynamic viewing by their determined web clients^[2]. Google Map and Google Earth launched in 2005 greatly enhanced both content and performance of publicly accessible Internet maps^[3]. However, it was not until recently that Internet

users could start to compose and publish their own maps through a web browser. The few systems that do support such usage are either licensed (GeoCommons), or limited to a small number of spatial features (Google MyMap).

One could argue that market forces are guiding geospatial technology to meet prevailing demands from two ends - and rightly so: one for the professionals with advanced analytical needs and institutional financial abilities, and the other for the general public who only need a simple and free platform for sharing light-weight personal data. But the authors argue that there is a middle ground which has been largely underserved. In this middle ground are academic scholars and small organizations who are not GIS professionals, and do not have access to expensive commercial systems, but are doing serious geospatial investigations requiring access to a variety of spatial data, and in the meantime are creating valuable spatial data in their own right. This middle group is the driving force for volunteered geographic information (VGI) creation^[4]. They need a system that comes to them in a web browser, does not require extensive training to use, and connects them to the vast and ever increasing online geospatial materials that are scattered across many systems. They need this system to allow them to upload their own local datasets, compose their own dynamic maps with a combination of data layers from multiple sources, save them online, and share them with whomever they want to share them with, let them exchange comments on these maps, and publish the maps in their blogs or other publication venues. Such a system would provide a permanent place to archive created data online, and (optionally) make their newly created data available to others as public goods.

WorldMap is the authors' attempt to provide such a system. It is designed to be publicly accessible, simple to use for non-technical scholars, fast in search and mapping speed, and rich in geographic content. Figure 1 illustrates our view of WorldMap in relation to the other Internet mapping systems such as ArcGIS Server for institutional professionals and Google Map for the general public.

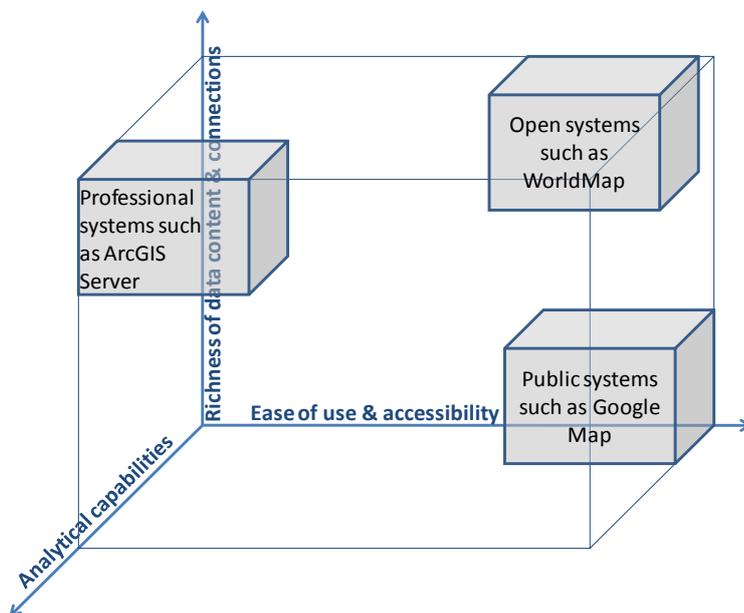


Figure 1 The unique niche for WorldMap in relation to other types of Internet mapping systems

In summary, professional systems such as ArcGIS Server can provide rich data content, supporting connections to many remote data sources that are in compliance with standard geospatial data sharing protocols, and support relatively strong data editing and analytical functions. But such systems are expensive to implement, and require substantial training to customize and maintain^[5]. On the other end of the spectrum, public systems such as Google Map are open to the public, simple and easy to use, and fast

in response if the user does not load much local data to them. But these systems provide a limited and fixed set of base maps and few analytical functions (such as address geocoding and traffic routing). The public systems are also not effective in handling large volumes of local data the users create or upload. Connection to other data services and combining data layers on the map usually requires programming skills.

WorldMap is an open source system that combines the strength of the professional systems' rich data content and ability to connect to remote data sources together with the public systems' ease of use and open accessibility. Its limitation currently is its weak data editing and analytical capabilities compared with the professional systems, and the relative complexity of its user interface which requires some minimum training for end users compared with the public systems.

It is worth noting that in recent years, there have been many web-based mapping systems made public, and more are appearing each day. Most of them are read-only to the public, with theme-based contents. The following are some prominent ones we have seen for historical studies:

- The US National Historical Geographic Information System (NHGIS) <http://www.nhgis.org/>
- Great Britain Historical Geographical Information System (GBHGIS) <http://www.port.ac.uk/research/gbhgis/>
- Belgian historical GIS http://www.hisgis.be/start_en.htm
- HGIS Germany <http://www.hgis-germany.de/>
- Regnum Francorum Online <http://www.francia.ahlfeldt.se/>
- Pleiades <http://pleiades.stoa.org/>
- The Digital Himalaya <http://www.digitalhimalaya.com>
- Chinese Historical GIS (CHGIS) <http://www.fas.harvard.edu/~chgis>
- The Digital Atlas of Roman and Medieval Civilization (DARMC) <http://darmc.harvard.edu>

More are available on the Historical GIS Clearinghouse and Forum

http://www.aag.org/cs/projects_and_programs/historical_gis_clearinghouse

Furthermore, Google Map API (and other similar mapping APIs such as that from ESRI and Microsoft Bing Maps) allow programmers to develop web based applications that can take user submitted data and map them against Google base maps. One successful example is “map Places in Time” (Mappit) <http://mappit.mpiwg-berlin.mpg.de/mappit>.

WorldMap is not “yet another Google Map API” system. Although it takes advantage of the best publically available base maps including Google base images and maps, as well as Open Street Map, it does not rely on any one of them. Any of WorldMap's base maps can be replaced at any time, and the system is not at the mercy of any commercial provider.

There is no public system we know of that allows users to upload large raster or vector GIS files then symbolize, save, and control access to them online. Commercial systems such as GeoCommons and ArcGIS Online offer a limited, vector oriented online version for free, with a premium one for sale.

We call WorldMap an open mapping system because it is open for the public to use, its source code is licensed as an open source product, its architecture relies on open geospatial standards for interoperability with other data sources, and its content is open for sharing at the individual data providers' discretion.

III. THE PROMISE OF OPEN GEOSPATIAL STANDARDS, OPEN SOURCE GIS AND VOLUNTEERED GEOGRAPHIC INFORMATION

Open geospatial standards, open source GIS and volunteered geographic information form a three-legged stool, a foundation for building an interoperable geospatial system on the web. The technology of the web and of open source GIS has been evolving rapidly in recent years and has at last become mature enough to support an implementation like WorldMap^[6]. It is only because of such a foundation that the

development of WorldMap would be possible at a small center in a university, with less than one full time programmer and a part time project manager.

WorldMap stands on the shoulders of many open source components. At its core is GeoNode, a Django-based framework which ties many of the WorldMap pieces together and provides the foundation for feature symbology, map composition, metadata management, and user profile administration. The PostgreSQL relational database with the PostGIS libraries is used for storing vector spatial features. GeoServer is used for spatial feature rendering. Large raster files are served from a tile cache server through a JavaScript-based map client. Tile generation occurs based on WMS (Open Geospatial Consortium Web Map Service) requests. Data can be stored in any source projection and re-projected on-the-fly to Spherical Mercator for overlay with commercial map services such as Google Maps. WorldMap uses the OpenLayers JavaScript library, which contains connectors to many proprietary mapping services such as Google Maps and Microsoft Virtual Earth, as well as non-commercial map services such as Open Street Map. Figure 2 and Table 1 show the major open source components in the WorldMap system architecture. Figure 3 is the WorldMap system architecture diagram.

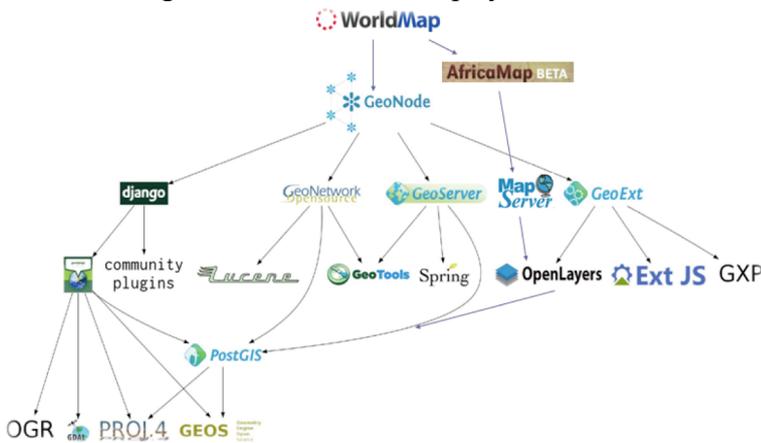


Figure 2 Major open source components in the WorldMap system architecture

TABLE 1 KEY ELEMENTS OF WORLDMAP ARCHITECTURE

Component	Primary Open Source Project	Language	Interoperability standards
Map viewer client which is used within a browser such as IE or Firefox	OpenLayers	JavaScript	HTTP, WMS, WFS, GeoRSS
Map rendering and symbolization which allows the user to control the appearance of a map.	GeoServer	Java	WMS, WFS, SLD
Data export to provide data in various useful formats	GeoServer	Java	GML, KML, GeoJSON, CSV, PDF, JPEG, PNG
Authentication to allow users to control access to materials they contribute	Django	Python	
Metadata catalog to store information about data layers to support search and documentation	GeoNetwork	Java	CSW
Spatial RDBMS to support a wide variety of spatial data storage, manipulation, and query operations	PostGIS	C++	Spatial SQL
Map caching to allow the GeoServer rendering engine to handle much greater loads by eliminating need to render any map twice	GeoWebCache	Java	WMS-C

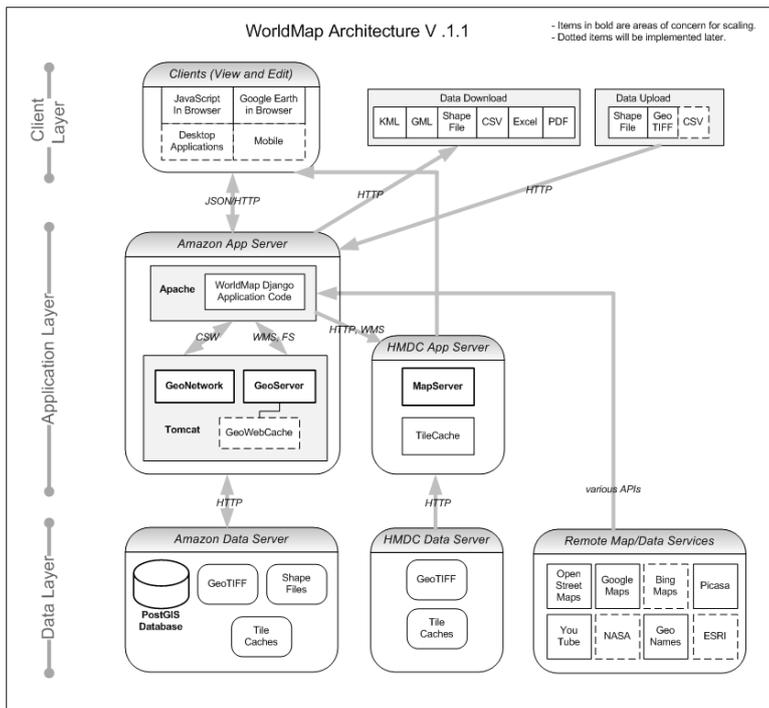


Figure 3 WorldMap system architecture diagram

WorldMap uses open standards, open web service protocols, and open data formats to allow the system to consume services hosted by other organizations as well as for providing map services to other systems. Data content on the WorldMap system accumulates 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. Figure 4 illustrates open data access, exchange, and collaboration enabled through WorldMap.

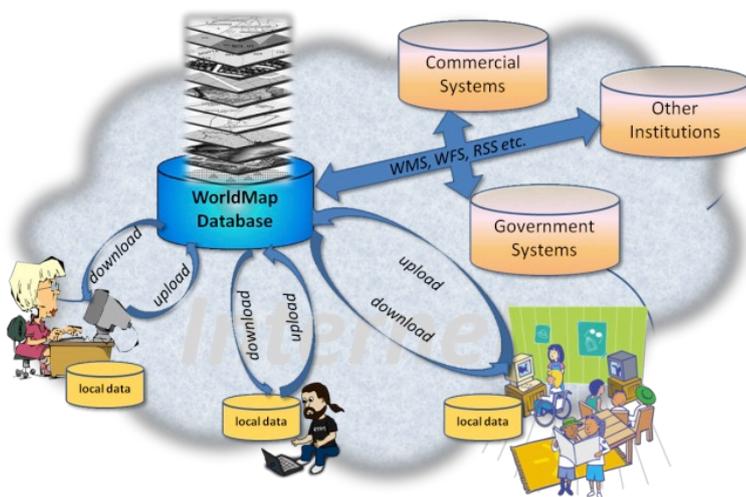


Figure 4 Open data access, exchange and collaboration through the WorldMap system

IV. THE INSTITUTIONAL DILEMMA – USER DEMAND AND FUNDING CONSTRAINTS

The development of WorldMap has been a challenging and rewarding experience for the authors. The CGA is dedicated to supporting research and teaching which makes use of geospatial technology. Through our interaction with Harvard scholars across many fields over the years, we noticed that when scholars are

not provided with adequate technical help, they will continue their research in the traditional way. Once they are shown how technology can help, their demand for technical support surges. To address these demands case-by-case requires little initial investment, but the surging demand can quickly overwhelm the service capacity. To build a system like WorldMap to enable scholars to help themselves is the ultimate solution, but that requires initial funding and staff time, which competes for resources against case-by-case services immediately.

Thus we started system development in small steps. The first implementation, for the continent of Africa, released in November 2008, proved to be an effective solution for integrating and making accessible a very wide range of mapping data on Africa that was not available online previously. AfricaMap was then cloned to create portals focused on other parts of the world, including Boston, Vermont, Paris and East Asia. Because the system didn't have the ability to allow users to create their own customized mapping sites and upload their own data, the demand for "a clone of AfricaMap for my region with my data" quickly became overwhelming to the CGA staff. A budget crisis prevented us from starting the second phase development of the system until two years later.

We started building the new WorldMap system in August of 2010. By January of 2011, several research groups had already started using the alpha version of the new system for project work. The early adopters provided us with valuable feedback and helped test the system, but their projects were built on an immature version of the platform and required some extra work to migrate to the later releases. It became clear that the new WorldMap system was more appealing than the previous AfricaMap system because of key new functions that were added, including the ability for users to create custom maps, to upload data, control access to their data, and define symbology for the map layers. The authors have had to strike a delicate balance between introducing the system to potential users in order to gather feedback, and not having too many users on the system before it has been fully tested.

In July of 2011, WorldMap Beta was released for public use. By November of 2011, the system has had over a thousand registered users; more than two thousand data layers containing tens of thousands of attribute fields; hundreds of maps which are unique compositions of the data layers; and about fifty thousand visits by unique IP addresses. Six Harvard courses used WorldMap in the classrooms in fall of 2011, and 22 courses are planning to use it in the spring semester.

Feedback from the early adopters helped the authors build and constantly update a system enhancement task list. However, other than bug fixing, all enhancement tasks are prioritized by funding, not by simple popularity of requests. In short, WorldMap's funding model is "first ask, first pay". When a user group needs a certain additional function and can fund its development, we will build it with that funding, and the product becomes available to all. Cost sharing occurs whenever two groups want the same feature at the same time.

Major improvements from Alpha to Beta include improvements on caching images and better handling of local service requests, thus significantly improved system performance; implementation of the "Jump tool" to bring social media such as Picasa and YouTube into the mapping platform; enhanced layer search in the system; and the ability to identify and report on attributes by map clicking across layers.

WorldMap version 1.0 is scheduled to be released in January 2012, which will include ability to consume Esri REST services; integrated Google Street view capability; injection of ArcGIS style creator for 32 and 64 bit Windows OS clients; an online image georeferencing tool; and direct layer creation and editing capabilities on WorldMap browsers.

More enhancements scheduled for 2012 are time animation of layers; commenting and ranking of maps and layers; map annotation creation and sharing; mobile clients; and improved layer search. Most of them are still pending for funding approval.

WorldMap also benefits from the global Open Source developer community which cumulatively contributes millions of dollars each year to improve the platform by building on core WorldMap elements such as Django, GeoNode, GeoServer, OpenLayers, ExtJS, GeoNetwork, PostGIS, and GDAL.

V. APPLYING CUTTING EDGE TECHNOLOGY – A DOUBLE EDGED SWORD

Because WorldMap is open source, powerful new functionality can often be added rapidly at relatively low cost. If another group in the world has built a tool we need then we integrate it. Sometimes we need something that doesn't exist so we build it ourselves. In other cases we build on a tool that has been started by someone else, contributing the resulting code back to the commons where it can be further improved. This approach has allowed the system to be examined, improved on, and extended by researchers and developers around the world. The approach also allows the system to remain vendor neutral and capable of working with virtually any mapping system through the use of open standards, formats, and web protocols.

Building a system with open source components enabled us to quickly prove the concept and deploy much needed functions in a short time. The excited user community welcomed it, and quickly started to demand more functions, better user interfaces, faster performance, broader data contents, etc. In short, WorldMap exposed, and partially fulfilled an unmet need in the academic community.

The initial success of the system quickly brought web traffic beyond our in-house server's capacity. Anticipating a load increase, the authors deployed WorldMap on Amazon EC2, a commercial virtual computing environment. This allowed us to quickly expand the system capacity in response to load changes. However, scalability also requires more investment in the infrastructure, but end users do not see this directly, thus are not interested in sharing the cost. Technical solutions alone will eventually be insufficient for resolving the scalability issue. When financial decisions are made, the inevitable question is how much an educational institution is willing and able to subsidize an open system for the general public to use.

At some point in the future, we may need to persuade some heavy institutional users to set up their own WorldMap instances, thus separating out their web traffic and storage load from the Harvard system. WorldMap architecture allows for "cloning" of the system with minimum developer effort and at no software cost. The ultimate solution for long term sustainability of the system is an endowment, which is yet to be explored.

Rapid customization of open source components gave WorldMap a jump start. This approach allowed the authors to bring in the newest technology for much needed functions. However, developing at the cutting edge can lead to bleeding too. Some of the technology components are not yet mature. For example, GeoNode version 1.0 was only released in December of 2010. It offers excellent capabilities at the proof-of-concept level, but some of these capabilities are not ready for large scale deployment yet. Much of WorldMap system development work has been on customizing these components, expanding their capabilities, and sometimes re-engineering them to handle much larger data sets and more complex use cases. Often times it requires extra effort to synchronize our customized new features with new enhancements rolling out in the software components' subsequent releases.

VI. ISSUES WITH GEOSPATIAL INFORMATION SHARING

Besides system development, another challenge is content management. Intellectual property rights are protected by permission settings set by the data contributor, with the assumption that people contributing data to WorldMap are responsible for obtaining and safeguarding the proper rights to the data. The system offers a platform for them to conveniently share the data with whomever they wish to grant access to. They may change the permission settings for data they own at any time. Such a design enables WorldMap to support the complete research life cycle, in effect creating a "permeable membrane" for spatial data sharing (Figure 5).

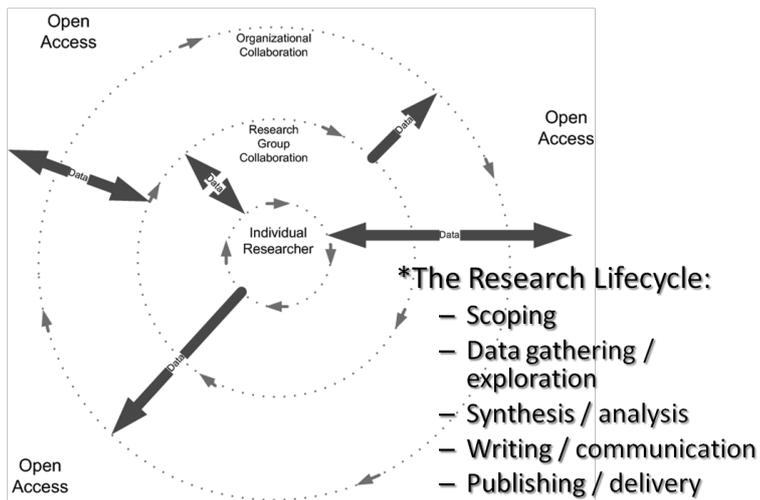


Figure 5 Permeable Membranes for Spatial Data Sharing

An individual researcher can start her project with private data which are only available to herself on WorldMap (inner circle in fig. 5). As the project develops, she can open up permission to a few of her collaborators for review and comments, by simply changing her data layer and map site permissions in WorldMap (middle circle in fig. 5). When her work is mature enough to be used in classrooms or publications, she can again change permissions to allow more people, even her entire institution, to access her data and/or maps (outer circle in fig. 5). When the project is completed and published, she may choose to open it up for public access (outside of all circles in fig. 5). In this entire process, there is no need to duplicate any data, or transfer any data to any other cyber locations. Work progresses will be reflected in the same version that is viewable in real time by all with the proper permission.

Most maps on WorldMap are the result of collaboration. It begins by assembling data layers from multiple sources, contributed to the system by different users. It continues when multiple map editors, often from different organizations in physical locations, login to WorldMap from their Internet browser and edit the same map at different times, symbolizing and organizing different data layers in the map. Soon WorldMap will allow users to make comments and rank maps and data layers online, and mark up maps with annotations and share them with other users. WorldMap serves as a virtual “mapping room”, where people from any part of the World can come together through the Internet, assemble data layers from many different sources into various maps, each serving a unique research purpose. These maps present different scholars’ understanding of that part of the world, facilitate discussions on their commonly interested topics, and publish their findings to the designated audience, often to the general public.

As WorldMap starts to invite users to contribute data content, new challenges emerge. The system needs to implement a set of mechanisms to differentiate institutionally vetted content from individual casual uploads. It also needs to enable crowd-sourced monitoring for crowd-sourced contents – allow users to report improper content, vote on data quality, rank data by its popularity, and share comments on data content. Most of these system features are to be developed in the next few months.

Other data management issues include permanent links, citation for publication, and long term data archiving. WorldMap has started to solve these problems but more work remains to be done. WorldMap creates a snapshot of a map any time it is saved, making it possible to go back to earlier versions of a map. WorldMap also supports live map embedding of any map view into any HTML page.

The authors’ approach to long term data archiving is through system integration with more mature digital archives. The service oriented architecture of WorldMap allows it to consume data services from any external systems adhering to the same open standards. One possibility is to archive data in systems such as the Harvard Geospatial Library (<http://hgl.harvard.edu>) and the Dataverse Network

(<http://thedata.org/>), and consuming such archived data through web services from these systems. Detailed implementation is yet to be worked out between these project teams.

VII. USING WORLDMAP FOR HISTORICAL GIS RESEARCH

Several of the early adopters of WorldMap happen to be projects with a historical GIS component. We will introduce a few of them briefly, to showcase the system's general functionality as well as its capability in support of historical GIS research.

Case 1 - ChinaMap

The objective of ChinaMap is to provide a historical and contemporary geographic window on China, which has a powerful searching tool for information and a way of visualizing and analyzing data on China. It brings together information from multiple sources, and various formats (maps, photos, video), and organizes them in space and in time to supports research and teaching on China. Figure 6 shows the opening view of ChinaMap on November 25, 2011.



Figure 6 ChinaMap powered by WorldMap

Data in ChinaMap includes political, infrastructural, demographic, economic, topographic, environmental, and historical topics, covering a variety of scales and extents, from the entire country, to one province on the Yangtze delta. Users may compare the social and economical development of the country through the past decades, or observe differences in these parameters among different geographic regions.

ChinaMap loosely connects to other systems such as Social Explorer for specialized data investigation (Figure 7), and accepts data contributed by users, such as the railroad tracks and stations in Figure 8. Overlaying 2011 high speed railways on the Qing Dynasty (1644 to 1912) courier routes revealed striking patterns of similarity (Figure 9), while comparison of the spatial-temporal distribution of the civil service examination in different dynasties shows both stark inequalities in regional representation and important shifts over time (Figure 10).

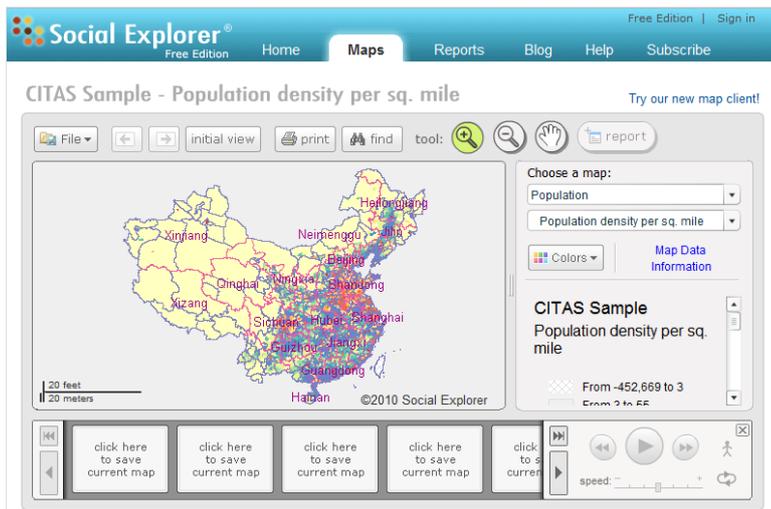


Figure 7 Social Explorer with China Data, accessed from the “Jump Tool” in WorldMap

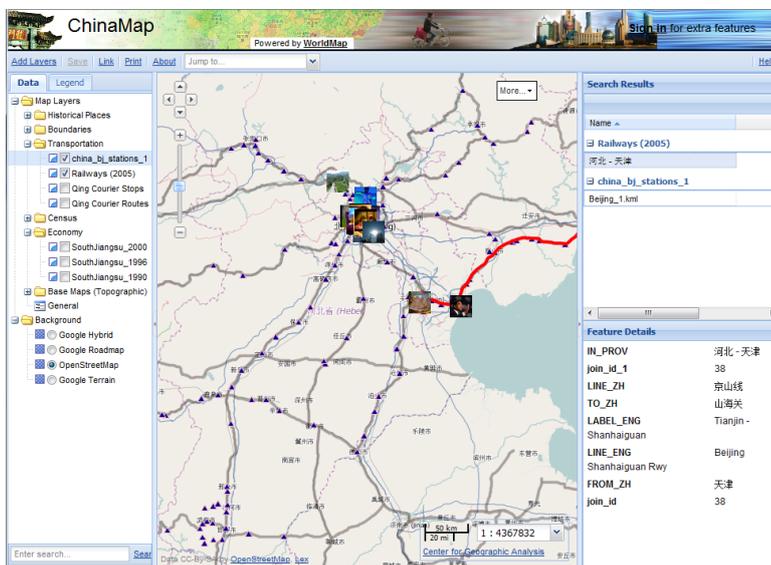


Figure 8 China’s railroad system

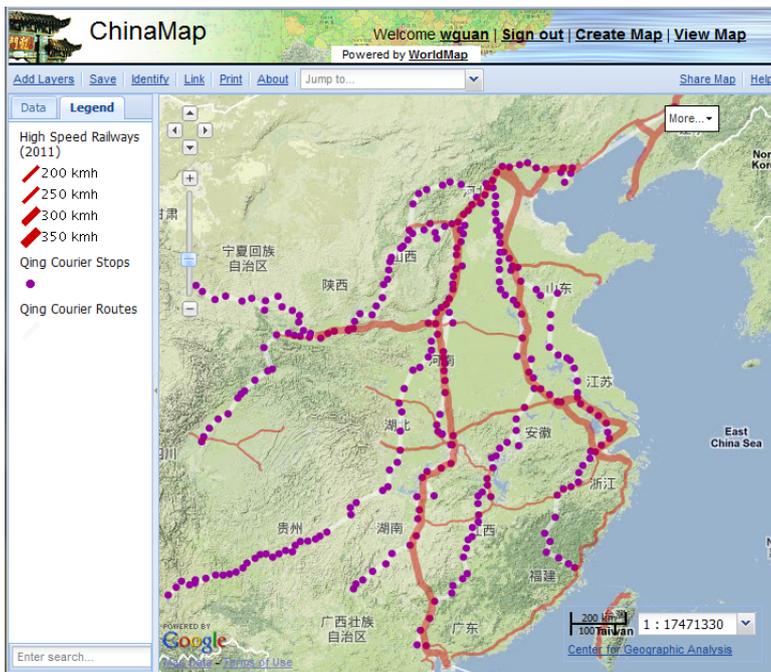


Figure 9 China 2011 high speed railways overlaid on the Qing Dynasty courier routes

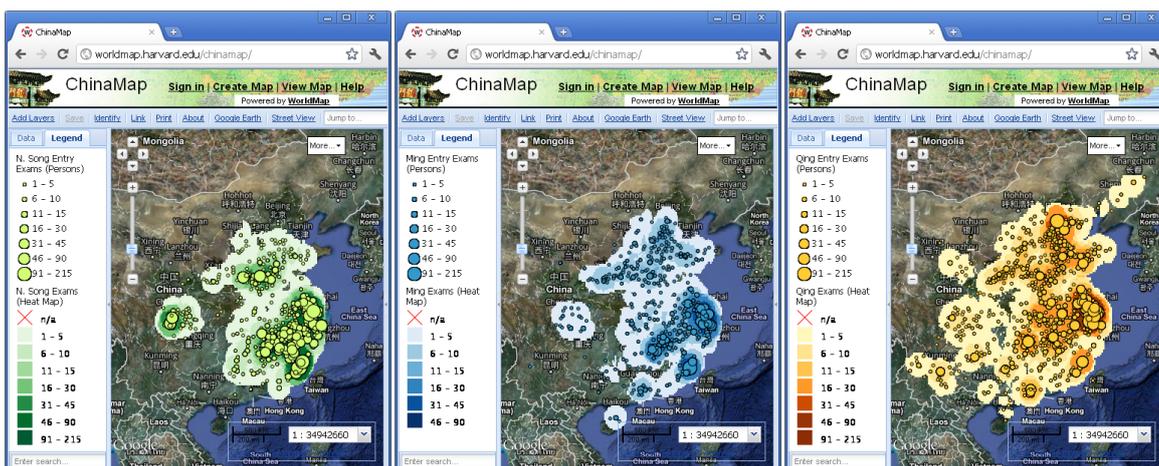


Figure 10 Civil service examination degree holders (left: North Song, middle: Ming, right: Qing)

The Chinese civil service examination became the primary means of recruiting educated men to serve in government during the Song dynasty (960-1279). Candidates began at the county level and worked their way up through prefectures and provinces to a three-day blind test at the capital. Although it was little used by the Mongols during the Yuan dynasty (1260-1367), it was restored during the Ming (1368-1643) and Qing (1644-1911) dynasties. The analysis on ChinaMap was done using extensive data from the China Biographical database (CBDB) [7].

WorldMap does not currently have system level integration with CBDB. Data for fig. 10 was manually processed in ArcGIS and uploaded to WorldMap. However, the authors are planning to develop a new temporal gazetteer API in WorldMap which would receive a place name with time, then look up the name in the system designated temporal gazetteer datasets as well as publicly available gazetteer services (such as GeoNames and Google place names), and return a list of candidate locations with geographic coordinates and temporal durations for that place name. The first implementation of this API will not have

any natural language intelligence built in, but will depend largely on a straight text string match of the place name and its time designation between the input and the system designated temporal gazetteer datasets. The user's final selection among the returned candidate locations will be captured and used in the system to improve future matches. The system will be able to handle any language in UTF8 format.

ChinaMap serves researchers as a system for finding, analyzing, visualizing and publishing geographic data. It serves courses as a platform for exploring places virtually through various base maps such as satellite images, terrain models, or thematic maps. It supports demonstrating change over space and time, and provides a platform for organizing and comparing student created data sets and facilitating classroom discussions. ChinaMap also serves the general public as a window to understand a place through map layers and media linked to locations, such as Picasa photos or YouTube videos (Figure 8). ChinaMap, like many other implementations of the WorldMap system, is a living system that is growing rapidly in content and functionality.

Case 2 – AfricaMap

AfricaMap was the first instance of the WorldMap platform, and has evolved into a mapping system that presents a diverse collection of mapping data, many of which are uniquely created by Harvard researchers, and all available for public use. The topics include archaeology, economy, environment, ethnicity and languages, health, demography, and political administration. It also includes public media sources such as Wikipedia, in addition to YouTube and Picasa. Its Trans Atlantic Slave Trade layer depicts the Trans-Atlantic Slave Trade Database in geographic form and includes almost 35000 slaving voyages^[8] (Figure 11).

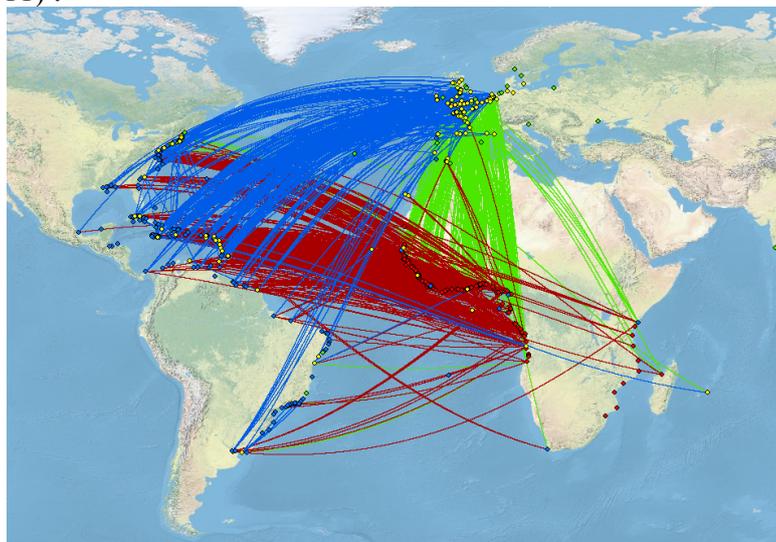


Figure 11 A Portion of the Trans-Atlantic Slave Trade Database Georeferenced. Green is leg one of the voyage, red is leg two, blue is leg three.

The historical maps in AfricaMap span from early 1600s to 1900s, allowing a semi-transparent overlay of any combination, which not only offers historical views of the continent through the eyes of generations of cartographers, but also make evident changes in understanding of different parts of the continent in different eras. The Period maps summarized the administrative regimes' territorial boundaries across Africa from 8000BCE to the present day (Figure 12), and the numerous base maps scale range from 1:1,000,000 to 1: 2,500; and cover different regions, countries and cities with different themes at different times. Of special note are two sets of ethnographic mapping unique to AfricaMap: one by Murdock (http://www.library.pitt.edu/articles/database_info/hraf.html) and one by Felix^[9]. Also unique are the Russian 500k and 200k mapping layers which cover the continent with detailed topographic mapping, dating from the 1960's to the 1990's. Each of these data layers has detailed metadata that is available for

public examination. Users may visit <http://africamap.harvard.edu>, right click on each data layer name, select “Layer Properties”, select the “About” tab in the popup window, and read the data description.

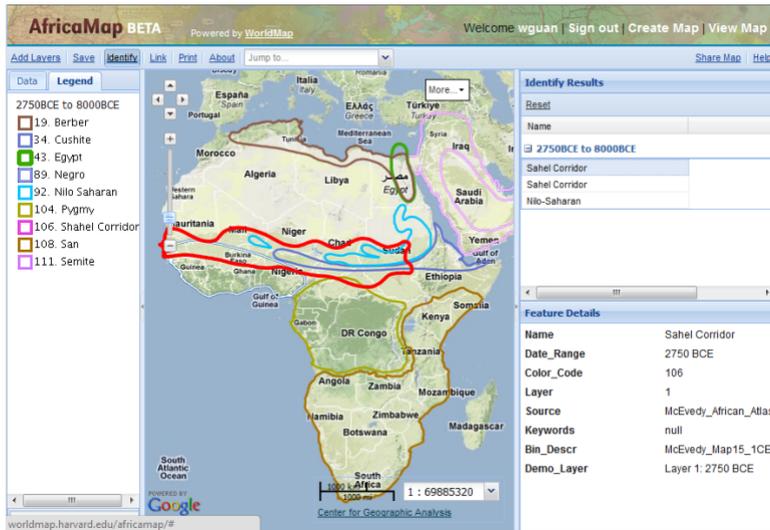


Figure 12 Africa Period Map of 2750BCE to 8000BCE

Case 3 – BostonMap

The creators of BostonMap (Rappaport Institute for Greater Boston at the Harvard Kennedy School and the Radcliffe Institute for Advanced Study) have focused mostly on contemporary data. Much of the data in the current BostonMap is census related. The group also added new tools to the WorldMap platform to support contemporary urban explorations including the “Jump” tool for accessing Social Explorer, Birdeye views, and Yelp. In addition, the “Street View” tool makes Google Street views available within BostonMap as well as within any instance of the WorldMap platform where Google has Street view images available (such as Japan).

A large number of new layers have been created for BostonMap by joining federal census boundaries to data gathered by the city of Boston. These materials are being organized in other maps such as Census 2010 (<http://worldmap.harvard.edu/maps/283>) and will be moved to BostonMap as needed. In addition to contemporary data for Boston, historical map layers help researchers better understand today’s Boston. Figure 13 shows that the location of today’s Logan International Airport was completely under sea water on an 1895 map.

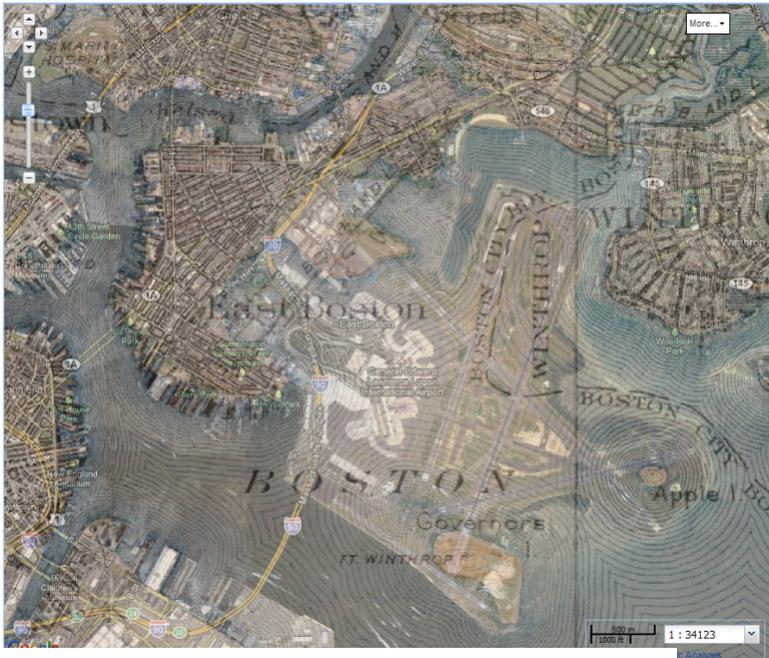


Figure 12 Boston 1895 regional map on Google Hybrid background

Case 4 – GizaMap

GizaMap is just starting to be built, with high resolution satellite images, aerial photos, survey maps, and archaeological sketches. A new set of historic mapping has just been acquired, which will be added to the system in the coming months (Figure 14).

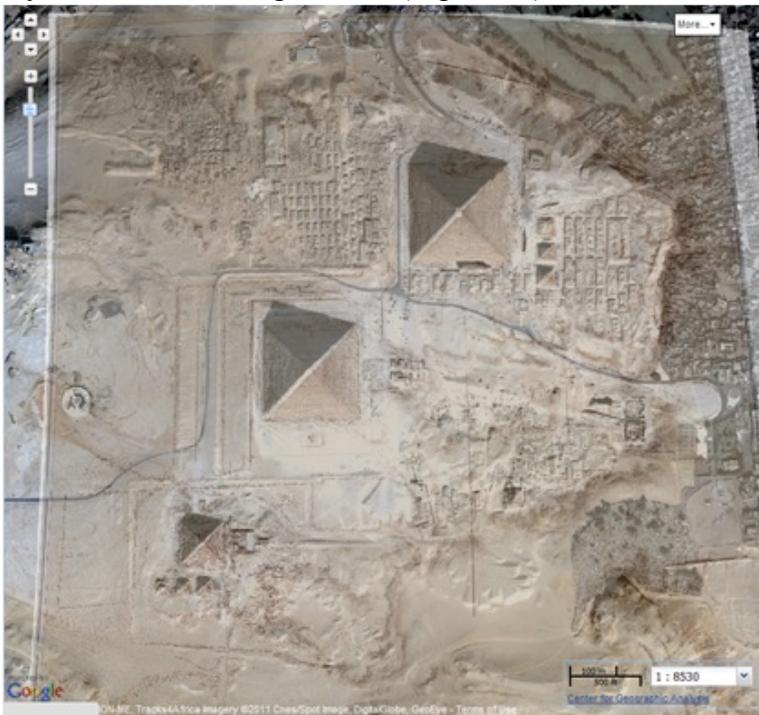


Figure 13 The Prisse d'Availles Plan (1878)
overlaid on satellite images

VIII. SUMMARY AND CONCLUSIONS

WorldMap is a web-based, map centric data exploration system that is built on open source geospatial technology. It is designed to be publicly accessible, simple to use for non-technical scholars, fast for search and visualization, and rich in geographic content. It provides an effective solution for integrating and making accessible dispersed collections, combined with diverse mapping materials to allow for cross-disciplinary inquiries from a spatial perspective. WorldMap allows any scholar to create their own mapping site on a browser, gathering, organizing and presenting mapping materials most relevant to their research, and sharing it with anyone they desire to. By November of 2011, WorldMap includes over 380 mapping portals in various stages of maturity and over a thousand data layers not available anywhere else.

Our experience from the WorldMap project indicated that such an open mapping system fills a niche for the academic user community. WorldMap is developed on open geospatial standards, open source software components, open web service protocols, open data formats and volunteered geographic information. This approach has enabled rapid development of the system with limited resources, but has also brought technical and institutional challenges to the project team, which includes:

- 1) synchronization of many software components under development by multiple organizations;
- 2) juggling competing priorities for serving individual requests and developing a system that will enable users to support themselves;
- 3) balancing the desire to promote public use of the system with constraints on infrastructure investment;
- 4) harnessing volunteered geographic information while effectively manage data quality; as well as
- 5) protecting copyrights, preserving permanent links and citations, and providing long term archiving.

Geospatial technology has enjoyed accelerated development in the past decades, and this momentum seems to be continuing^[10]. New web mapping tools and functions, both for commercial and open source products, are being developed rapidly. Any statement made on what can and cannot be done by a particular product is almost certain to be outdated in a few months. With that in mind, the authors invite you to experience WorldMap as an evolving platform, aimed at supporting collaborative research across multi-disciplines within a cutting edge open geospatial framework.

ACKNOWLEDGMENT

The authors thank Professors Suzanne Blier and Christopher Winship for their vision and guidance in conceptualizing, developing and promoting WorldMap. The WorldMap project team at CGA includes the authors and Guoping Huang, David Siegel, Bo Zhao, Julia Finkelstein, Molly Groome, Giovanni Zambotti, Anna Arzumtsyan and Kirk Goldsberry. For a complete list of sponsors and software development contributors for WorldMap, please visit <http://worldmap.harvard.edu/>.

REFERENCES

- [1] M. P. Peterson, Maps and the internet, Vol. 1, International Cartographic Association, Elsevier, 2003.
- [2] A. A. Alesheikh, H. Helali, and H. A. Behroz, Web GIS: Technologies and Its Applications, Symposium on Geospatial Theory, Processing and Applications, Ottawa, 2002.
- [3] D. Butler, (2006). Virtual globes: The web-wide world, Nature, 2006, vol.439, pp.776–778.
- [4] M. F. Goodchild, Citizens as sensors: the world of volunteered geography, Published online: 20 November 2007, Springer Science+Business Media B.V. 2007.
http://www.ncgia.ucsb.edu/projects/vgi/docs/position/Goodchild_VGI2007.pdf

- [5] K. Leukert, and W. Reinhardt, GIS-Internet Architectures, International Archives of Photogrammetry and Remote Sensing, Vol. XXXIII, Part B4. Amsterdam 2000.
- [6] B. Lewis, and W. Guan, Jump-starting the next level of online geospatial collaboration: Lessons from AfricaMap, Advances in Web-based GIS, Mapping Services and Applications, Li, Dragicevic & Veenendaal (eds), London: Taylor & Francis Group, 2011.
- [7] China Biographical Database, General Editor Peter K. Bol (Cambridge, the Fairbank Center for Chinese Studies at Harvard University; Taipei, the Institute of History and Philology of Academia Sinica; Beijing, the Center for Research on Ancient Chinese History at Peking University: 2005-). <http://isites.harvard.edu/icb/icb.do?keyword=k16229>.
- [8] The Trans-Atlantic Slave Trade Database, Emory University, <http://www.slavevoyages.org>.
- [9] C. Meur and Congo Basin Art History Research Center, Peoples of Africa: ethno-linguistic map, M. L. Felix (eds), Tribal Arts, 2001.
- [10] S. Kumer, Geospatial industry: Here today, world tomorrow, Published online: 25 January 2011, Geospatial World. http://www.geospatialworld.net/index.php?option=com_content&view=article&id=21415:geospatial-industry-here-today-world-tomorrow&catid=366:janeary-2011.