MERGEOS: A PILOT STUDY FOR COLLABORATIVE MAPPING AND DYNAMIC GIS WEB APPLICATION

A THESIS PROPOSAL PRESENTED TO
THE DEPARTMENT OF GEOLOGY AND GEOGRAPHY
IN CANDIDACY FOR THE DEGREE OF
MASTER OF SCIENCE

By
Chau Nguyen

NORTHWEST MISSOURI STATE UNIVERSITY
MARYVILLE, MISSOURI
July 10, 2011
Mergeos: A Pilot Study for Collaborative Mapping and Dynamic GIS Web Application

Chau Nguyen

Northwest Missouri State University

THESIS APPROVED

Thesis Advisor: Dr. Ming-Chih Hung

Dr. Yi-Hwa Wu

Dr. Ming-Hsiang Tsou

Dean of Graduate School
Abstract

The concepts of Neogeography and Volunteered Geographic Information (VGI) allow collaborative data managed by volunteers to be shared in an open environment. One way to explore these concepts and the knowledge about them is to record the creation of a data sharing community with an experimental web application called Mergeos. This web application would have the ability to add or “merge” data together to build a more complete data set of points, polygons, and lines from willing participants. Using tools and practices for the functions of mass data management, collaborative GIS, and social network data collection require us to contemplate and execute GIS solutions in different ways. The main purpose of the Mergeos experiment is to research, describe, and confirm these new concepts by developing a community site at a low cost.

The objectives of this research can be described in two parts. The first is to analyze and describe the process of creating a community site whose purpose reflects on neogeographic concepts. This involves comparing the differences between designing, planning, and developing a project with the traditional methods used for common GIS mapping applications. The second is to create a website to demonstrate a solution that implements these concepts by using the techniques and analyses from the first part of the research. Mergeos is intended for users to begin a data set, setting
the standards by defining what data they want to store and how others will add to this data set, while maintaining it.

Mergeos was developed successfully and has demonstrated that a community site is possible with the accessibility of affordable technical resources. There are various limitations that can be addressed in future research. Overall, providing documentation on the creation of a collaborative community site helps us discover concepts of shared GIS and data management. Mergeos displays the many prospects that neogeographic and VGI concepts have available.
# Table of Contents

Abstract .......................................................................................................................... iii

List of Figures ................................................................................................................ vii

List of Tables ................................................................................................................ ix

Acknowledgements ........................................................................................................ x

List of Abbreviations...................................................................................................... xi

Chapter 1: Introduction .................................................................................................... 1
  1.1 Research Objectives ................................................................................................. 3

Chapter 2: Literature Review ........................................................................................... 4
  2.1 Neogeography & VGI Background ............................................................................ 4
  2.2 Existing VGI Websites .............................................................................................. 8
      2.2.1 OpenStreetMap ................................................................................................... 8
      2.2.2 GeoCommons ..................................................................................................... 10
      2.2.3 ArcGIS Explorer Online ..................................................................................... 10
  2.3 Criticisms and Concerns ........................................................................................... 13
      2.3.1 Crowdsourcing ................................................................................................... 13
      2.3.2 Vagueness ......................................................................................................... 14
      2.3.3 Credibility ......................................................................................................... 15

Chapter 3: Methodology .................................................................................................. 17
  3.1 Study Area ................................................................................................................ 17
  3.2 Data Sources ............................................................................................................ 18
  3.3 Site Design ............................................................................................................... 20
  3.4 User Experience ....................................................................................................... 25
  3.5 Technology & Architecture ..................................................................................... 30
      3.5.1 Existing Tools ................................................................................................... 30
      3.5.2 Web Architecture ............................................................................................. 31
      3.5.3 Database Design .............................................................................................. 35
  3.6 Security & Roles ....................................................................................................... 38
  3.7 GIS Integrations ........................................................................................................ 41
      3.7.1 Server – Side Utilities ....................................................................................... 41
      3.7.2 Geographic Data File Format ......................................................................... 43
3.7.3 Mapping API ............................................................................. 45

Chapter 4: Analysis Results ........................................................................ 48
  4.1 Site Design Results ............................................................................ 48
    4.1.1 Search Data Sets ......................................................................... 49
    4.1.2 Create a Data Set .......................................................................... 52
    4.1.3 Manage your Data ......................................................................... 57
  4.2 VGI Sites Comparisons ........................................................................ 63
    4.2.1 OpenStreetMap ............................................................................ 63
    4.2.2 GeoCommons ............................................................................... 67
    4.2.3 ArcGIS Explorer Online ................................................................. 70
  4.3 Google Analytics Results .................................................................... 73
  4.4 Future Promotions ............................................................................... 74

Chapter 5: Conclusion ............................................................................... 76

References .......................................................................................... 81

Appendices .......................................................................................... 83

  Appendix A – Term Glossary ................................................................. 83

  Appendix B - Logo References .............................................................. 86

  Appendix C – Term Glossary References ................................................ 88
List of Figures

Figure 1: Japan’s earthquake and tsunami travel time ........................................6

Figure 2: OpenStreetMap website .........................................................................12

Figure 3: GeoCommons website ...........................................................................12

Figure 4: ArcGIS Explorer Online website ..............................................................12

Figure 5: Study Area ..............................................................................................18

Figure 6: Mergeos procedure hierarchy .................................................................20

Figure 7: Website draft ..........................................................................................23

Figure 8: LinkedIn website ...................................................................................23

Figure 9: Open Source Web Design website ..........................................................23

Figure 10: Alert Messages upon submitting invalid information ..............................24

Figure 11: Website Architecture workflow ...........................................................33

Figure 12: Entity Relationship Diagram ................................................................39

Figure 13: Security roles process ..........................................................................40

Figure 14: GIS Integrations process .......................................................................44

Figure 15: Mergeos homepage .............................................................................48

Figure 16: Search Data Set process .......................................................................51

Figure 17: Data Bundle review page ....................................................................51

Figure 18: Create Data Set process .......................................................................53

Figure 19: Data Bundle review notification to owner ............................................53
Figure 20: Approve e-mail notification ...............................................................53
Figure 21: Data Bundle descriptions ...................................................................55
Figure 22: Data Bundle viewer ...........................................................................55
Figure 23: Data Bundle approval ........................................................................56
Figure 24: Data Bundle review confirmation .....................................................56
Figure 25: About Mergeos page ..........................................................................57
Figure 26: Login page ..........................................................................................58
Figure 27: Forgot Password page .........................................................................58
Figure 28: Create Account page ..........................................................................59
Figure 29: My Account page .................................................................................60
Figure 30: My Data Bundles / Data Sets page ......................................................60
Figure 31: Edit process and page .........................................................................61
Figure 32: User’s roles process and page ..............................................................62
Figure 33: OpenStreetMap Export tab .................................................................65
Figure 34: OpenStreetMap Edit tab ....................................................................65
Figure 35: GeoCommons Upload data .................................................................68
Figure 36: GeoCommons Edit and Make a Map ..................................................68
Figure 37: ArcGIS Explorer Online Presentation tab .........................................71
Figure 38: Google Analytics Results in January 2011 .........................................73
List of Tables

Table 1: Data sources ........................................................................................................19
Table 2: User experience features .....................................................................................27
Table 3: UCD Five Stage Framework for Mergeos ..............................................................29
Table 4: Web architecture tools ..........................................................................................32
Table 5: Development framework / standards .................................................................35
Table 6: Database server ...................................................................................................37
Table 7: Security roles process ..........................................................................................40
Table 8: GIS integration additions ....................................................................................43
Table 9: Map Application Programming Interface options ..............................................46
Table 10: Comparison list between OpenStreetMap and Mergeos .................................66
Table 11: Comparison list between GeoCommons and Mergeos .................................69
Table 12: Comparison list between ArcGIS Explorer Online and Mergeos .........72
Table 13: Challenges and limitations ...............................................................................78
Acknowledgements

First off, I would like to mention that I feel privileged to be a part of the Master’s Program of Geographic Information Sciences at Northwest Missouri State University.

I would like to acknowledge the help of all those who have assisted me in my thesis work. The members of my committee, Dr. Ming Hung, Dr. Yi-Hwa Wu, and Dr. Ming-Hsiang Tsou (San Diego State University), have all given me excellent guidance and advice throughout the project. I’m grateful to them for having an open mind and providing valuable input to my thesis.

I would also like to extend my gratitude to everyone who participated in testing my project. I truly appreciate all your help and time spent providing valuable feedback. I thank my friends and proofreaders for providing feedback on my thesis paper. I would like to thank all of my co-workers at EdomGroup, Inc and Getfused, Inc. for their continued support and interest in my research and education. They also allow me the opportunity to apply GIS solutions for our clients in the future. I would also like to give special thanks to my husband, Mike, for his endless support of my research and academic endeavors.
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>API</td>
<td>Application Program Interface</td>
</tr>
<tr>
<td>CF</td>
<td>ColdFusion</td>
</tr>
<tr>
<td>CSS</td>
<td>Cascading Style Sheets</td>
</tr>
<tr>
<td>ESRI</td>
<td>Environmental Systems Research Institute</td>
</tr>
<tr>
<td>GDAL</td>
<td>Geospatial Data Abstraction Library</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic Information System</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technology</td>
</tr>
<tr>
<td>KML</td>
<td>Keyhole Markup Language</td>
</tr>
<tr>
<td>OGR</td>
<td>Opensource Geospatial Reference</td>
</tr>
<tr>
<td>OS</td>
<td>Operating System</td>
</tr>
<tr>
<td>SQL</td>
<td>Structured Query Language</td>
</tr>
<tr>
<td>UCD</td>
<td>User-Centered Design</td>
</tr>
<tr>
<td>UI</td>
<td>User Interface</td>
</tr>
<tr>
<td>XHTML</td>
<td>Extensible Hypertext Markup Language</td>
</tr>
<tr>
<td>XML</td>
<td>Extensible Markup Language</td>
</tr>
</tbody>
</table>

See Appendix A for Term Glossary.
Chapter 1: Introduction

The idea of connecting the problem of data accessibility and new concepts of web-based data “sharing” strategies has led to new possibilities for Geographic Information Systems (GIS) professionals and students. These new possibilities could enable users to collaborate on both geospatial data collections and their management, leading to the establishment of more comprehensive data sets over time. These data sets will be managed by a community of authors, users, and contributors. A major challenge of collaborative web-based GIS is the creation of a place where one can reliably find and access data that others might willingly share with one another, such as data sets and maps.

The need for data sharing amongst volunteered participants was clear after one of the deadliest U.S hurricanes, Hurricane Katrina, hit New Orleans, Louisiana on August 23, 2005. This unfortunate natural disaster led people to want to provide information on the hurricane and re-population patterns to support the community in New Orleans. They wanted an online solution that was easy to use, affordable, and recognizable to users to prevent them from having to re-learn anything regarding site appearance and tools. The Greater New Orleans Community Data Center (GNOCDC.org) was able to accomplish this task by providing a site where assorted
maps and data were uploaded by volunteers. Examples of some of these maps are post-Katrina commuter patterns, housing damages, and population loss.

Concepts that play a part in these new solutions include Neogeography and Volunteered Geographic Information (VGI). Neogeography is a new term that includes new tools and techniques that would be considered outside of traditional GIS (Turner 2006). VGI is a specialized information collection method which allows collaborative data creation and management by volunteers and citizens in an open and distributed environment. One way to explore these concepts and knowledge is to document the creation of a data sharing community with an experimental web application. This experimental application would have the ability to add or “merge” data together to build a more complete data set of points, polygons, and lines from willing participants. Going forward, this experiment will be referred to as “Mergeos” in this thesis study.

Using tools and practices for the purposes of mass data management, collaborative GIS (Geographic Information System), and social network data collection require us to think and implement GIS solutions in atypical ways. The main intentions of the Mergeos experiment are to research, describe, and confirm these new concepts by developing a community site at a low cost.
1.1 Research Objectives

The objectives of this research can be described in two parts. The first is to analyze and describe the process of creating a community site whose purpose depends on neogeographic concepts. This involves comparing the differences between designing, planning, and developing a project with the traditional methods used for common GIS mapping applications.

The second objective is to create a website to demonstrate a solution that implements these concepts by using the techniques and analyses from the first part of the research. Mergeos is intended for users to begin a data set, setting the standards for other contributors, by defining what data they want to store and how others will add to this data set, while maintaining some control over its accuracy and usability. This usage will be tracked using Google Analytics and encouraging users to contact the site manager if they have any problems. These data sets are then made accessible to other GIS professionals to use in their research.
Chapter 2: Literature Review

Currently, there are many different sites that utilize Neogeography and VGI concepts. This chapter will discuss some of these sites along with the background and concerns of Neogeography and VGI. This chapter will show how powerful these concepts are, the capabilities of such a collaborative site, and how it could help different individuals and communities expand their data sets.

2.1 Neogeography & VGI Background

Neogeography brings the technology of online mapping, map mashups, and the lives of individuals to a new level of interaction, regardless of a users’ expertise. This trend has brought new uses for data in relation to maps, despite implementation complexity and traditional standards. Developments in Neogeography have resulted in interactive maps that allow novice users to comprehend and visualize spatial data without programming, web, or GIS backgrounds. However, the creation of these maps would require some programming and GIS knowledge depending on their complexity. With all of the interest in Neogeography over the years, it is not until recently that we have actually begun to realize the real tools for quick, immediate data collection and visualization that we can use today.

In the last three years, a trend and tool in Neogeography called VGI has emerged onto the scene, although the process has been around for about two decades.
VGI is rapidly contributing to a collection of user content and data volunteered by individuals along with government agencies or private companies. There have already been sites utilizing VGI for this purpose, such as OpenStreetMap (www.openstreetmap.org) and MapTube (www.maptube.org). Even after the creation of these tools, there are still opportunities for using VGI to access a wide range of data beyond the common sources like census and terrains. Compiling a collaborative database from various GIS professionals and students could lead to a wider variety of low cost alternatives. Having a collective GIS database could also encourage collaboration to make more data sets available that the originators could not complete on their own.

Governments and private companies may be providing valuable data, but users often need access to more specialized data, like recent natural disasters. An example of the latest natural disaster is the 8.9-magnitude earthquake that hit the northeastern coast of Japan on March 11, 2011 and caused a large tsunami. These two events caused massive damage to cities and nuclear power plants. See Figure 1 of Japan’s earthquake and tsunami travel times map from NOAA (National Oceanic and Atmospheric Administration).

In an effort to aid Japan in rebuilding and to provide additional details on the two events, volunteer experts and non-experts uploaded their map data to community sites to share with others. One of these community sites is the Japan Sendai Earthquake Data Portal (http://cegrp.cga.harvard.edu/japan/), which is the Harvard
Center for Geographic Analysis’s response to the earthquake in Japan. This site encourages people to share their data regarding this natural disaster and resulted in compiling various types of maps for academic use only.

By combining smaller bundles of GIS data together, users could create larger data sets that would not exist otherwise (like for Hurricane Katrina or the Japanese earthquake), as an alternative to commercial data sources. These data sets could fill needs that were previously ignored and inspire new research. This could also motivate individuals to volunteer their data for cost-efficient alternatives (Goodchild 2007).

Figure 1: Japan’s earthquake and tsunami travel times
The implementations of VGI will continue to emerge as data is contributed by volunteers and shared with the public adding to the long-term success of VGI (Goodchild 2008). Wikimapia (wikimapia.org) and Google MyMaps (http://maps.google.com/support/bin/answer.py?hl=en&answer=62843) are examples of some early implementations that utilized volunteered GIS data to the extent of adding to street maps and viewing certain government or agency data. Most provide users with basic functions, such as making minor changes, exporting data to different formats, and creating their own data set, however, there are demands to allow users to merge their own data with existing data provided by others.

The result of increasing the use of VGI is that citizens become sources for data and information, which signifies a change from top-down and expert-driven GIS (Elwood 2008). VGI is an important source for geographers because it may take a lot of time, but is a low cost alternative and it often focuses on locations or data that may otherwise go unnoticed.

VGI has also opened new doors in environmental monitoring by including the input of citizens. Using a combination of open source tools, sensing devices, and back-end infrastructure, environmental modeling products can be designed based on the interests of citizens, who become active participants in building the required data set (Gouveia and Fonseca 2008). Gouveia and Fonseca discussed the Senes@Watch project that endorses VGI by finding problems and determining data issues in relation to user’s complaints. Liu and Palen (2010) presented neogeographic practice analysis
by creating a genesis of crisis map mashups, evolution features, and design decisions.

In conclusion, the authors recommended that the increased accessibility and interoperability of geotechnology’s professional and participatory forms aid in shaping Neogeography.

2.2 Existing VGI Websites

In this section, some of the most common community sites that users have been introduced to are discussed, including:

- OpenStreetMap
- GeoCommons
- ArcGIS Explorer Online

2.2.1 OpenStreetMap

OpenStreetMap (OSM) is a community site with over 130,000 users worldwide where users can share data, enter constructive spatial information, and be authors of their own maps (Lyons 2009). Its purpose is to display thorough street-level maps from individuals willing to share and publish their data. In addition, users can organize workshops for other potential users in their area to expose them to the OSM experience of collecting data to the project. The use of intuitive icons and the ability to track users’ changes make this website visually appealing. OSM uses the Creative
Commons Attribution Share Alike 2.0 licensing to make vector graphics freely available for download and sharing. As a result, users are allowed to copy, circulate, and transfer maps. Users can reuse the map data as long as the terms of share alike and attribution are followed.

Share alike requires users to share their new creations with the data they reuse in the same way that it was provided to them, while attribution requires that the original users’ work is recognized in the manner specified by the author or licensor.

OSM software is free, but it can be expensive to purchase GPS devices and computer hardware. The server software is from various projects with their relative licensing, meaning that OSM uses various other projects and users are subject to each project’s license for that chunk of functionality. This means that they might use a map technology that is their own and completely free-open source. However, other elements of the project may be free but not open source, or open source but not for commercial use. Out of all the emerging user-generated mapping projects, OSM is probably the most effective project as of 2008 (Haklay and Weber 2008). The usage of OSM is increasing and mapping events are being established in social and technical environments. These editable maps can be accessed at www.openstreetmap.org (see Figure 2).


2.2.2 GeoCommons

GeoCommons, created by FortiusOne (founded in 2003), allows users to upload, download, and search for data in addition to combining data sets (Schutzberg 2007). Although professionals may appreciate this type of site, the main concern is encouraging users to upload their data. From an academic perspective, researchers can publish their spatial data and promote their work through the site. Users can browse through various categories (agriculture, demographics, education, etc.) and either view, edit, or create their own maps. Mashups are easily created and there are data storage and other tools for that purpose. GeoCommons was built on the FortiusOne GeoIQ platform, using a Creative Commons with Attribution licensing. The Google Maps API (Application Program Interface) is used as the backdrop with Flex to draw map features upon it. GeoCommons provides tutorial demos to assist users with different features around the site. GeoCommons can be accessed at http://geocommons.com (see Figure 3).

2.2.3 ArcGIS Explorer Online

The ESRI (Environmental Systems Research Institute) recently launched ArcGIS Explorer Online. This is an online site where users can share maps amongst different groups and communities (even the iPhone community). ArcGIS Explorer Online provides access to ArcGIS.com for a collection of various basemaps to choose from and add mapping features to (e.g.: measurements, identification, drawing shapes,
etc.). ArcGIS.com is a place where users can share maps with particular groups or communities and invite others to share their maps (ESRI 2010). JavaScript Web Mapping application, ESRI online services, and Silverlight are used to quickly build the maps. ArcGIS Explorer Online can be accessed at http://explorer.arcgis.com/ (see Figure 4).

Besides online community sites, other factors have emerged over the years that have contributed and added to the growth of VGI and Neogeography concepts. This includes the wikification of GIS from those differing levels of GIS experience and mass media. The wikification of GIS is a large collaboration of willing participants who depend on each other to work together and solve problems using Web 2.0 technology. According to Sui (2008), the growth of wikification of GIS could be a result of earlier contributions from open-source developments like Linux, Wikipedia, and eBay as well as successes from other community sites, such as Facebook, YouTube, and MySpace. An interesting example of this development is a basic map created by amateur cartographers based on the tattoo found on Angelina Jolie’s left arm. The tattoo lists the latitude and longitude of the birthplaces of her four adopted kids. To visualize the locations, a map was produced by volunteers using Google Maps with mashup capabilities, Google Earth, and KML (Keyhole Markup Language) for street-level details that were exposed to 100 million users from Google Earth (Sui 2008). This example shows how a community defined by common interests
Figure 2: OpenStreetMap website

Figure 3: GeoCommons website

Figure 4: ArcGIS Explorer Online website
worked together to solve a problem with accessible tools beyond the traditional solutions of GIS. The release of Google Earth, Microsoft’s Virtual Earth, and other forms of mass media has affected the growth in VGI and Neogeography by involving citizens in sharing information. Overall, media and wikification of GIS has created an emergence of a new geography without geographers (Sui 2008).

2.3 Criticisms and Concerns

Despite all of the benefits, there have been some criticism and concerns due to the idea of users and communities volunteering data.

2.3.1 Crowdsourcing

Some projects have asked the public for their help (which is also known as outsourced or crowdsourcing) on certain open source projects or tasks. Most times, this results in affordable costs but is a tradeoff for uncertainty of quality since it is from users of various levels of experiences. Crowdsourcing is a concern because the labor costs a lot less than paying traditional employees and could have a negative influence on other professionals’ future business (Howe 2006). This is because cost barriers were established to distinguish the professionals from the amateurs. In the recent years, cost barriers have changed in order to make buying more affordable for more users. Howe (2006) discussed four situations where crowdsourcing impacted
negatively on professionals / freelancers and businesses. For instance, iStockPhotos (www.istockphotos.com) is a site where user-generated stock photos and videos are made available, royalty-free for prices in $1.00 US increments. Having photos at such a low price point has made the competition difficult for free-lance photographers like Mark Harmel. Howe (2006) described that Mark lost a sale to iStockPhoto from Claudia Menashe, who he quoted $600 for photos of sick people even at a discount. Menashe chose the affordability of getting her licensed photos at a much lower cost and the people who took the photos benefitted as well. Unfortunately, crowdsourcing situations do not have a clear line of who is wrong since both parties have legitimate reason to justify their actions. Therefore, there is not a clear solution to resolve this ongoing issue since everyone’s goal is different.

2.3.2 Vagueness

Vagueness is another concern of VGI, especially in altering and affecting peoples’ perceptions due to uncontrollable factors. For example, climate change can modify views from stakeholders which can result in undependable data due to the environment. However, solutions are being developed like the web based VGI system (eVGI prototype) described by De Longueville et al. (2009). The eVGI prototype collects and stores vague geographic information to analyze perceptions of stakeholders’ testimonials from sectors in the Barents region. This allows VGI and
metadata to be created to portray geographical vagueness while OSM ignores this.

### 2.3.3 Credibility

Other concerns are the credibility and reliability of user-generated contents from willing participants. With the ongoing debates about data quality, a possible solution is to employ peer-to-peer credibility assessment to increase quality control (Flanagin and Metzger 2008). Credibility is an important field for future works that could be based on rating and commenting on functionalities (De Longueville et al. 2009). Data gathering, management, and representation are other concerns that cannot be controlled directly due to the disjointed nature of many different parties and their volunteered data. There can only be standards and rules put in place, such as having the owner of a data set approve changes, which may help, but cannot provide a perfect solution to an ongoing issue. Willing participants’ values, priorities, demographics, and different experiences may provide various concerns and results (Sieber 2004). However, having a variety of volunteered data from various backgrounds may increase understanding and aid in analysis.

Some of these concerns cannot be fixed due to numerous factors that could only favor one side or the other, but proper procedures can be put in place to help the situation. Regardless of the concerns, “in essence, everyone is a NeoGeographer and NeoGeographical information is for everyone” (Rana and Joliveau 2009, p. 80). The literature shows that interest in using community sites with VGI and Neogeography is
growing. This can benefit everyone in different ways, even as concerns and issues are identified and addressed. This will add to the endless opportunities for collaborative websites utilizing VGI.
Chapter 3: Methodology

It is very important to develop a methodology for achieving the goals set forth with what is accessible in a given situation. In this case, affordability and open source alternatives must be considered due to limited funds and student status. The proposed methodology consists of five parts: site design, user experience, technology & architecture, security & roles, and GIS integrations.

Eventually, users will be able to find the data set they need and evaluate its quality for their purpose using the maps and tools that Mergeos provides. Users would also have the ability to download and control data sets for use in their own maps and analysis. In addition, other users will be able to find an existing data set that is similar to what they have collected. To make it a more complete set, users can add their data to the existing data set. The site makes the responsibility of quality control up to the users because the site manager is not controlling their data.

3.1 Study Area

A particular study area is not clearly defined since data can come from various sources and be distributed by different participants. To set study area limitations for this project, the focus will be on vector data within the United States as seen in Figure 5. Figure 5 can be seen at http://xoax.net/childrens/USCapitals_support/Image1.png. The target audiences are GIS students, educators, professional, and the general public.
3.2 Data Sources

As a constraint on the endless possibilities of different data types that can be uploaded; the thesis only focused on vector data consisting of lines, points, or polygons. Data may come from various counties within the United States. Data from various sources may be used in the testing phase to stabilize and test the capabilities of the site. Refer to Table 1 for the original sources of testing data.

Figure 5: Study Area
Table 1: Data sources

<table>
<thead>
<tr>
<th>Data Name</th>
<th>Source and URL</th>
<th>Data Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lakes</td>
<td>Tarrant County, TX [<a href="http://www.tarrantcounty.com/egov/site/default.asp">http://www.tarrantcounty.com/egov/site/default.asp</a>]</td>
<td>Polygons</td>
</tr>
<tr>
<td>School Districts</td>
<td>Crook County, OH [<a href="http://gis.co.crook.or.us">http://gis.co.crook.or.us</a>]</td>
<td>Polygons</td>
</tr>
<tr>
<td>School Districts</td>
<td>Union County, OH [<a href="http://www3.co.union.oh.us">http://www3.co.union.oh.us</a>]</td>
<td>Polygons</td>
</tr>
<tr>
<td>Zapata County, TX Boundary</td>
<td>Texas Natural Resources Information System [<a href="http://www.tnris.state.tx.us/datadownload/download.jsp">http://www.tnris.state.tx.us/datadownload/download.jsp</a>]</td>
<td>Polygons</td>
</tr>
<tr>
<td>Kinney County, TX Boundary</td>
<td>Texas Natural Resources Information System [<a href="http://www.tnris.state.tx.us/datadownload/download.jsp">http://www.tnris.state.tx.us/datadownload/download.jsp</a>]</td>
<td>Polygons</td>
</tr>
<tr>
<td>LA Sex Offenders</td>
<td>Louisiana’s Courts [<a href="http://www.la.gov/offender">http://www.la.gov/offender</a>]</td>
<td>Polygons</td>
</tr>
<tr>
<td>U.S. State Parks</td>
<td>Texas State Parks [<a href="http://www.tpwd.state.tx.us/landwater/land/maps/gis/">http://www.tpwd.state.tx.us/landwater/land/maps/gis/</a>]</td>
<td>Polygons</td>
</tr>
<tr>
<td>U.S. State Parks</td>
<td>Tennessee State Parks [<a href="http://www.tn.gov/environment/parks/gis/data/">http://www.tn.gov/environment/parks/gis/data/</a>]</td>
<td>Polygons</td>
</tr>
</tbody>
</table>
3.3 Site Design

Figure 6 demonstrates the procedures and pages that each user encounters on the site. To briefly describe the procedure hierarchy, users will first see the welcome page (the site’s homepage), with four buttons to choose from, “Search Data Sets”, “Create A Data Set”, “About Mergeos”, and “Log In.” “Search Data Sets” lets users search for data sets, view the results, add a data bundle, review and approve data bundles if they are the owner, and download a data set in KML format. Creating a data set requires an account, which consists of providing account and data details. An

Figure 6: Mergeos procedure hierarchy
account is needed if users want to add or edit a data set. Once an account is created, data sets and data bundles can be uploaded. Mergeos uses the term “data bundles” to describe a group of data that is contributed to a data set by an individual user. With the appropriate access, owners can edit data sets. Owners and managers can set user’s roles and everyone else can view the list of contributors to a data set. If an account already exists, users can log into their account and view contributors and their data bundles along with their security roles. If a user forgets his or her password, a “Can’t remember your password?” link is provided that emails their password. The option to have the site remember users’ email is provided on the login page. The “About Mergeos” page briefly discusses the purpose and functions of Mergeos.

Figure 7 displays an example website draft page of the “Data Bundle and Data Set Review” pages, where the data will be displayed in a data viewer at the bottom of the screen. The menu will consist of “Search Data Sets”, “Create A Data Set”, “About Mergeos”, and “Log In” that will be consistent throughout the site. The design processes is intended to improve the definition and usability of each section of the website, using the current draft as a foundation. Descriptions of the users’ experiences navigating the site are found under the 3.4 User Experience section. The final website designs and pages are shown as screenshots that can be viewed in Chapter 4 under the “Site Design Results” section. The details of procedures are also discussed in Chapter 4.
An application's appearance is the first impression, so it is important to create a pleasing look by supplying a simple and inviting design. The objective is to engage users to the site and set the stage for an intuitive user interface (UI) that efficiently aligns functionality with the task the user wishes to complete. Having a slim, task oriented design with a contrast of colors, such as the LinkedIn site (www.linkedin.com) is helpful. An example of a task-oriented design is a white background with accent colors and rich simple icons that clearly identify points of action (see Figure 8).

By learning from applications like LinkedIn, it is easier to develop a website template draft or make better decisions when choosing a premade template. There are community sites, such as Open Source Web Design (www.oswd.org), where a collection of web designs are created and posted by the community that the public can use for free. To save time and cost, the Open Source Web Design site was used for its variety of web design selections. Figure 9 shows a screen shot of the Open Source Web Design homepage, where users begin searching for the template they need. An appropriate design was chosen from this website and compared to the template draft that was developed as well as the website references.
Figure 7: Website draft

Figure 8: LinkedIn website

Figure 9: Open Source Web Design website
Mergeos has the following site design elements:

- Template easily to modify regardless of technical experience
- Big buttons for users to effortlessly locate when looking to complete an action
- Color contrast to give the site a clean appearance and well defined areas of content that the user can repeatedly find as they browse the site
- Centered layout to put the application in front of the user regardless of their screen size
- Easily identifiable alert and information boxes when an action is performed (see Figure 10)
- Site design works in various internet browsers (Chrome, Internet Explorer, Safari, and Firefox)
- Readable fonts and wider line spacing allows users to comprehend and read content easily

Figure 10: Alert Messages upon submitting invalid information
3.4 User Experience

A user friendly application encourages its users to continue coming back for the functionality they need. For Mergeos, it serves as the real story of how a user will find and use a data set as well as share one. This section describes guidelines that should be considered when designing to ensure a good user experience from two references. The following is a list of guidelines from Isaacs and Walendowski (2002) to use in UI design for websites:

- **Visual Cues** (example: bold text, font colors, text size, and borders)

- **Simple Icons** (example: question mark symbol that indicates the meaning of “help” or presenting a “question”)

- **Tables and Charts** (example: a chart representing a business sales and profits)

- **Simple Menus** (example: menus for common navigation requirements like going to the home page, content page, contact page, etc.)

- **Appropriate Colors** (example: green can represent completed tasks or complementary colors can be used to make content more readable)

- **Footers** (example: a footer can be used when a user encounters an error by displaying the message to the user in the footer)
- Notifications of User Expectations (example: notify users of required fields)
- Buttons labeled clearly with actions (example: use action labels like “Add” and “Search”)

One of the guidelines of user design is using visual cues or simple icons that can be interpreted with ease and without needing conscious thought to convey their meaning. Labels can be used if users need to interpret the cue or icon’s meaning quickly. Information is easier to understand in forms of visual grid structure like in tables and charts. This helps to organize data and focus user attention on significant information. Another guideline is using simple menus with the minimum number of items possible to avoid complexity and confusion. Using appropriate colors to emphasize content, represent redundant cues, or to provide an appealing website appearance can also enhance a user’s experience with the site. It was also stated that it is good practice to notify users if something is expected from them, such as required fields, instead of surprising them later when they do not perform the action. Having such surprises can lead to negative user experiences and frustration. Another guideline is to label buttons with actions instead of generic words like “Yes”, “OK”, or anything that would pressure users to read text carefully. When writing content, everyday language should be used versus jargon of specific industry. If jargon were used, it would place an assumption that users know the terminology, which is not always the case.
Based on these guidelines, a list of features on user experience was created for Mergeos, as shown in Table 2. These features will be provided in the website in an attempt to improve the user’s perceptions and response while using it.

<table>
<thead>
<tr>
<th>Features</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Font and Sizes of Buttons, Links, and Text</td>
<td>Are clear and legible with no unrelated content placed near or around action items.</td>
</tr>
<tr>
<td>Location of Buttons, Links, and Fields</td>
<td>Are purposely placed for user’s convenience. For example: the download link is next to “Load Data Set” button above the data viewer for accessibility.</td>
</tr>
<tr>
<td>Light Contrast of Background Colors</td>
<td>Are used to make the content easier to read.</td>
</tr>
<tr>
<td>Forms</td>
<td>Are consistently laid out throughout the site so the user can easily navigate from each field to the next one.</td>
</tr>
<tr>
<td>Highlighted Fields</td>
<td>Assist users by indicating clearly where the user’s focus is on the form.</td>
</tr>
<tr>
<td>Instructional Context and Descriptions</td>
<td>Are available on some of the pages to instruct users on how to accomplish a particular task with everyday language.</td>
</tr>
<tr>
<td>Error Messages</td>
<td>Are displayed in footer when an error occurs.</td>
</tr>
<tr>
<td>Menu Buttons</td>
<td>Are displayed on the top if the page to assist users in navigating through the site. There are four menu buttons for simplicity.</td>
</tr>
<tr>
<td>Action Buttons</td>
<td>Label buttons with actions such as “Search”, “Go”, and “Login.”</td>
</tr>
<tr>
<td>Required Fields</td>
<td>Have text that notifies users that required fields are indicated by a red asterisk.</td>
</tr>
<tr>
<td>Forgot Password Link</td>
<td>Is available for users who do not remember their password.</td>
</tr>
<tr>
<td>Email Remembered Option</td>
<td>Have the option to have the site remember a user’s email for the next time he or she logs in.</td>
</tr>
</tbody>
</table>
Tsou’s and Curran’s (2008) hydrological case study on User-Centered Design (UCD) approach is another reference to consider as a guideline for user friendly site design. UCD is a design process that is focused on the users’ needs, behaviors, and wants in using a site. This research suggests a five stage framework when designing a website that is discussed in Jesse James Garrett’s book (2002): Strategy, Scope, Structure, Skeleton, and Surface. Tsou and Curran (2008, p. 315-316) states the following about each stage as procedures to web mapping applications:

- **Strategy Stage** – Determine what users want and need out of the site and mapping services.
- **Scope Stage** – Based on the web and mapping service scope, determine the required features to support the site’s needs and metadata.
- **Structure Stage** – Plan how everything on the site will be composed and focus on the mapping functionality and map data contents. Each function should be intuitive to satisfy the functional requirements indicated in the Scope Stage.
- **Skeleton Stage** – Make concrete and tangible structure with necessary components and ensure that data objects are arranged into a significant category while mapping function groups with a conceptual hierarchy. The data layers are in the hydrologic layers (the main theme) or base layers (supporting information) category.
**Surface Stage** – This stage is the most important because it results in the actual finished site. The map contents, graphical designs, and styles are completed and utilized by users at this stage.

Table 3 shows the five stage framework based on the UCD approach for Mergeos.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Web Map User Interface Design (functions)</th>
<th>Web Map Display Contents (data layers)</th>
</tr>
</thead>
</table>
| Surface  | • Window and data viewer arrangements  
          • Use of graphic icons, buttons, windows, and links for the user interface                              | • Search for existing data sets  
          • Allow users to upload their own data sets and data bundles  
          • Allow users to download datasets                                                                 |
| Skeleton | • Map Display Window  
          ° Zoom-in  
          ° Zoom-out  
          ° Pan  
          ° Rotate  
          • Information Dialog Box Window  
          ° Data information supplied in metadata (feature name, shape area, etc…)                               | • Base layers defaulted from Google Earth:  
          ° Satellite Photos  
          • Main layers posted by users                                                                 |
| Structure | • Map display functions (Data Viewer)  
           • Search and comment functions  
           • My account and log in functions  
           • Security role and contributors functions  
           • Create data set and upload/submit functions  
           • Data download in .KML and Google Earth view functions  
           • Load Data Set function                                                                                     | Refer to Table 1: Data sources for list of data available                                                      |
| Scope    | • Interactive map manipulation  
           (Zoom-in, Zoom-out, tilt, and rotate to interested features)  
           • Upload GIS data in .KML or .SHP zip  
           • Download GIS collaborative data                                                                                        | • Display and map volunteered data  
           • Provide Descriptions, Title, Created Date, Updated Date, and zip file of these maps                       |
| Strategy | **User Need**: Web mapping community site is a possible solution to accessing uncommon and unique data. Users can access site online with reliable internet with no additional software requirement.  
           **Map Service Objective**: This mapping service allows users to share and create collaborative data sets and merge data bundles for other users. |                                                                                                                                                                      |
3.5 Technology & Architecture

Another aspect to consider is what technologies best fit the desired functionality of Mergeos while being compatible with each other. With this in mind, CF (ColdFusion) was chosen for server-side integration of data and mapping features. CF fits the needs of the project due to its thorough integration features with Java, Flash / Flex, .NET and many other technologies. There are commercial and free / open-source implementations of the CF server available. The current state of open-source GIS is that some projects use .NET, others use Java, and still others use web service APIs, etc. CF can use all of these with relative ease of development. The Google Maps API (which uses Flex or JavaScript) or GeoServer (uses Java) may provide much of the mapping functionality and are freely available. However, it may be necessary to expand the set of tools due to the limits and ranges of features in the growing open-source GIS projects.

3.5.1 Existing Tools

It is important to find existing code libraries, projects, and applications to make use of and learn from in the course of creating an application like Mergeos. While availability may vary depending on the creator’s programming language or other technology choices, many tools support integration with other technologies. In addition, existing resources can still act as a reference to both the developer / creator even if they are not in the same programming language or capable of being integrated.
Particularly for common site functions like content management, user management, etc., this assists in cutting costs, avoiding “reinventing the wheel”, and saving time.

The tools used are discussed later under the section “Server-Side Utilities”.

The following are other benefits of using existing resources:

- Existing resources are often more stable because they have already passed the scrutiny of multiple projects before.
- Features not previously thought of are often found because others have shared their ideas in the form of an existing project.
- In some projects, the interface has the benefit of a great amount of time spent on it in comparison with the new project being created.

### 3.5.2 Web Architecture

When creating an application, attention must be paid on choosing the patterns and tools that the application will be built with to best handle all of its functions, purpose and growth for future needs. Web application architecture is comprised of selecting tools, such as programming languages, software, frameworks, and standard patterns to implement the application.

For free, user-driven applications like Mergeos, there are quite a few concerns when making architecture and technology decisions. The biggest concerns are cost and scalability, as these are often the main challenges to scale an application to larger
audiences even if the initial goals of the application are attained. While it is not the most ideal situation, it happens regularly that one must choose a tool that is less expensive even if the more costly alternative is better suited. Other concerns, emphasized consistently across all applications, include the creator’s familiarity with the technology, frameworks, and software they intend to implement as they relate to the purpose. Table 4 explains the tools that were used.

<table>
<thead>
<tr>
<th>Web Tools</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apache Webserver</td>
<td>Is a standard web server that is free and readily available. This webserver is also very scalable as the application grows and gains more users. See: <a href="http://httpd.apache.org">http://httpd.apache.org</a></td>
</tr>
<tr>
<td>Linux OS</td>
<td>Is a powerful and free OS (Operating System) that runs many web servers. This is also very scalable with no licensing issues when creating many web servers. See: <a href="http://www.linux.org">http://www.linux.org</a></td>
</tr>
<tr>
<td>ColdFusion Application Server and Markup Language</td>
<td>Has two free open-source servers available and can integrate with a large range of technologies. Allows access to useful codebases to help build basic website functions. See: <a href="http://www.adobe.com/products/coldfusion">http://www.adobe.com/products/coldfusion</a></td>
</tr>
</tbody>
</table>
Figure 11 shows the website architecture, which explains how the different tools connect together to create Mergeos and the user’s direct interaction with the webserver. The user’s requests are sent to the web server, which runs the Linux Operating System and Apache web server. Apache processes many of the requests for images or other supporting files and sends any dynamic requests to ColdFusion pages to the ColdFusion server application. CF then interacts with the database server, implemented with MySQL, to process data queries and interacts with GDAL (Geospatial Data Abstraction Library) to process geographic data. CF sends the response back to the Apache webserver, which finally returns it to the user’s computer. Figure 11 has the following logos: Google Chrome [Google 2008], Firefox [Mozilla Cooperation 2006], Opera [Opera Software ASA 1996], Safari [Apple, Inc. 2003], Internet Explorer [Microsoft 1995], Google Earth Plugin [Google 2005], Linux [Ewing 1996], Apache [Apache Software Foundation 1999], ColdFusion [Adobe 2005], GDAL [Geospatial Data Abstraction Library 2009], and MySQL [Oracle 2010].
In addition to choosing the appropriate technologies for implementing Mergeos, several development standards were chosen. Some of these standards speed up development time, while others provide consistency across all functions of Mergeos. These choices are invaluable for the long term management of an application and certainly affect its beginning greatly. Table 5 briefly describes the development standards.

ColdBox is very fast and even sets a standard for organizing code into folders and clearly contained segments. ColdSpring is used to organize the creation and dependencies of all the different objects necessary for an application. In addition, it makes more efficient use of server resources by controlling which objects may be reused for many requests and which must be re-created. Cascading Style Sheets (CSS) separates application elements from the actual display Hypertext Markup Language (HTML) making it both more readable and controllable for the developer. jQuery can then be used to implement basic animation or effects to improve user experience.

Generally, frameworks like these are available in any programming language the creator may choose. The creator should understand the benefits and deficits of each framework to choose the right one. Upon doing so, they will see similar advantages regardless of the programming language. Web standards, such as CSS, allow the use of simple HTML and Style Sheets to maintain a consistent look and feel across an entire application regardless of the programming languages used.
### Table 5: Development framework / standards

<table>
<thead>
<tr>
<th>Framework / Standard</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ColdBox</td>
<td>A ColdFusion application framework that speeds development by using an easy to “Model, View, Controller” pattern for implementing an application. See: <a href="http://www.coldbox.org/">http://www.coldbox.org/</a></td>
</tr>
<tr>
<td>CSS</td>
<td>Cascading Style Sheets (CSS) is a development standard used by browsers to more uniformly and easily control aesthetic elements of an application such as fonts, borders, layout, etc. See: <a href="http://en.wikipedia.org/wiki/Cascading_Style_Sheets">http://en.wikipedia.org/wiki/Cascading_Style_Sheets</a></td>
</tr>
<tr>
<td>jQuery</td>
<td>A JavaScript development framework and library for common JavaScript functionality interacting with the structure of a web page. See: <a href="http://www.jquery.com">http://www.jquery.com</a></td>
</tr>
</tbody>
</table>

#### 3.5.3 Database Design

To store information about users, data sets, data bundles, and all other elements of the application, Mergeos requires the use of a database. Essential components of database planning include choosing a database server and determining a database design. Taking the time to plan how data is being stored in a database and the challenges in gathering the information for display can save huge amounts of time and frustration. Database design can also help with overall performance and determine how easy or hard the maintenance process can be because many of the
variables are known ahead of time. The database servers researched are Microsoft SQL Server, MySQL, and Oracle Database. Table 6 explains each database server.

Microsoft SQL Server is considered the most well rounded database server. It is very fast and powerful with many geo data capabilities available in the latest version. A free “express” version is available for smaller databases. Unfortunately, licenses can be costly once the free version is outgrown and licensing issues could occur when scaling the application to use multiple database servers.

MySQL is the most popular free database solution and is very scalable by handling databases of any size with no licensing restrictions. This server has a large community and resource availability; however, the tools to manage and work on the database are not as good as many commercial databases. It does have some geo-integration capabilities that improve with each version.

Last is the Oracle Database, which is incredibly powerful and can handle very large databases better than the others. Unfortunately, the management tools are not very easy to use and this server is very expensive to license and implement. For a free application like Mergeos, this removes Oracle as a database server option.

After consideration, MySQL was chosen as the database technology to support the application due to its cost, lack of feature limitations such as database size, and being freely usable in both small projects as well as large clustered environments. As
Table 6: Database server

<table>
<thead>
<tr>
<th>Database Server</th>
<th>Description</th>
<th>See:</th>
</tr>
</thead>
<tbody>
<tr>
<td>MySQL</td>
<td>Is an open source relational database management system based on SQL. Allows dynamic web pages and is popular for database-driven websites. See: <a href="http://www.mysql.com">http://www.mysql.com</a></td>
<td><a href="http://www.mysql.com">http://www.mysql.com</a></td>
</tr>
</tbody>
</table>

the load and functionality grows, choosing MySQL helps ensure that all the technologies used can be scaled to handle far more users and features without too much expenditure. While these are not issues for an application at its inception, they become huge burdens or advantages when users deem the application a success and it is quickly growing.

The database design follows traditional relational database standards requiring separate tables describing each discreet object in the application. These include a separate table for data sets, data bundles, users, comments and many others for the purpose of organizing data more efficiently. These relationships are then defined so the application can efficiently join each data set to its data bundles. This provides a
clear separation of concerns and an easier path to add information to any item later without reworking other areas of the database. The database server does have some geographic capabilities. The decision was made, however, to store this information directly in geographic format files so it could be accessed more efficiently for display, and processed using other GIS utilities for the functionality required.

Figure 12 shows the entity relationships between these tables. The User table has a one-to-many relationships to the Comment, Data Set, Data Bundle, and Data Set Roles tables. The Data Set table has a one-to-many relationship to the Comment table, one-to-many relationship to the Data Bundle table, and many-to-many relationship to Data Set Roles table. The Role table has a many-to-many relationship to the Data Set Roles table.

3.6 Security & Roles

Mergeos’ functionality puts users in management roles of their data, allowing them to get help from others in building a larger and more accurate data set. This introduces problems in managing user access to the data sets and brings questions about who is allowed to perform all of the different actions on the data. These permissions for the data sets were separated by having the user roles on each data set depicted in Table 7.
These roles can be removed and altered as necessary by managers and owners with the supplied actions of promoting to writer or manager, demoting to writer, and removing from the system. When a user’s contributed data bundle is approved, the contributor can be given either role according to the owner or manager. Figure 13 presents the process of the security roles.

Figure 12: Entity Relationship Diagram
Table 7: Security roles process

<table>
<thead>
<tr>
<th>Role</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owner</td>
<td>Has complete control of the data set and can remove or change the role of another user on any data set they own.</td>
</tr>
<tr>
<td>Manager</td>
<td>Can set user’s roles. Managers can approve, add, remove, or update data bundles at any time.</td>
</tr>
<tr>
<td>Writer</td>
<td>Has some control of the data set and cannot give others the right to manage the data set. Writers can add, write, or edit data bundles at any time.</td>
</tr>
<tr>
<td>Contributor (General User)</td>
<td>Has the right to submit a data bundle to any data set for review. Once their data is added and approved, their contributions will be known to the community.</td>
</tr>
</tbody>
</table>
3.7 GIS Integrations

Mergeos is intended to be an easy application for its users to upload and manage data with using the fewest steps possible. However, it has some complex needs in terms of GIS integrations that would apply to any new GIS project that tries to take advantage of newer GIS concepts. For data aggregation, conversion, and display, Mergeos requires powerful tools at a low cost of entry and compatibility with any tools GIS professionals and students choose to use now and in the future. The three key GIS integrations required for Mergeos, and very common to almost any other neo-geographic application, are:

- Server-Side Utilities for data management and conversion
- Geographic data file format for display, export and manipulation
- Mapping API for displaying and communicating the value and contents of data

3.7.1 Server – Side Utilities

Mergeos explicitly requires users to be able to upload geographic data regardless of the software used to create their data. The application must be able to manage and convert that data into a result that can be shared with others and improved by the contributors of a data set. To this end, a lot of searching was done to find utilities that could convert and navigate different types of geographic data.
Research was also done to find utilities that can be executed automatically by the application and perform well enough to give the immediate feedback that users expect from a web application.

For this project, a lot of code and applications from my previous websites were used to complete the user, content, and other core functions. Many other projects were integrated to build the GIS functionality for Mergeos. These are GDAL, OGR2OGR, and OSGeo4W which are listed in Table 8.

Of the many open-source GIS projects, GDAL proved to be the ideal solution for this problem. GDAL is an open-source collection of utilities released by the Open Source Geospatial Foundation. GDAL abstracts, manipulates, and converts geographic data in many different formats via the command-line, perfect for use on a server by a web application. The most useful utility for Mergeos turned out to be a program called OGR2OGR. It directly converts geographic data files between formats and can even manipulate projections. Using OGR2OGR, Mergeos is able to receive and quickly convert ESRI Shapefiles or KML files to many other formats that GIS professionals use. In addition, as Mergeos grows, the functionality could easily be expanded and tested to allow users to upload those same formats (see: http://www.gdal.org/ogr/ogr_formats.html). As it is an abstraction library, OGR2OGR was built in a modular format. This means that many of the formats it supports and operations it performs are supplied by modules and adapters installed with the library.
### Table 8: GIS integration additions

<table>
<thead>
<tr>
<th>GIS Integration</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDAL / OGR2OGR</td>
<td>Provides a powerful command-line utility for converting GIS data between many formats. See: <a href="http://www.gdal.org/ogr/">http://www.gdal.org/ogr/</a></td>
</tr>
<tr>
<td>OSGeo4W</td>
<td>Provides a common installer for many GIS utilities, including for GDAL / OGR2OGR with many extensions for specific formats that do not come in the standard install. See: <a href="http://trac.osgeo.org/osgeo4w/">http://trac.osgeo.org/osgeo4w/</a></td>
</tr>
</tbody>
</table>

OSGeo4W was discovered after research into larger projects. OSGeo4W combines many smaller projects such as GDAL/OGR2OGR with many common adapters and utilities into one easily installable package. This includes installing the adapters for ESRI Shapefiles along with the other utilities needed for the operations being done on the server. Figure 14 describes the GIS integrations process. This figure has the following logos: Google Earth Plugin (Google 2005), ColdFusion (Adobe 2005), and GDAL (Geospatial Data Abstraction Library 2009).

### 3.7.2 Geographic Data File Format

After conversion tools/utilities were selected, the next challenge was choosing a common format to export and manipulate the data in a way that would allow data providers to feel empowered. The chosen format should allow the widest possible range of GIS users to be able to access and use the data in their own projects. The common data format must be capable of:
- Being imported, modified, and exported by as many various GIS applications as possible
- Presenting many different layers and types of data

The ESRI Shapefile format was the first to come to mind, however, it is encumbered by proprietary standards and poor support in many GIS applications due to its proprietary nature. While it has no doubt proven its capabilities storing great amounts of widely varying data, its lack of open specification and wide support removes it from consideration as a common processing format for Mergeos.

Figure 14: GIS Integrations process
KML, however, was found to be an open standard. KML was pushed forth by Google after purchasing a company named Keyhole, and is a widely supported format in almost every GIS application studied. It is an XML (Extensible Markup Language) based file format, making it very flexible for developers and GIS users, while managed by an open specification that defines how data is organized and written for many different purposes. In addition, it solves the problem of standardizing coordinate projections as the KML specification only allows EPSG (European Petroleum Survey Group):4326. This is convenient since GDAL/OGR (Opensource Geospatial Reference) already supports conversions to and from this projection. For these reasons, KML was selected as the common data format to which all data is converted / managed in Mergeos.

### 3.7.3 Mapping API

In order to visually understand and review the data sets published via Mergeos, a mapping API must be used as a canvas to display the data set on. The requirements involved are fast parsing and display of potentially large data sets. The ability for a user to consistently view the entire data set at one time with complete control over their perspective is another requirement. Table 9 describes the free map APIs that were considered.
Note that the only non-server dependent solution found is Google Earth. In some cases, this might be considered a weakness, however, in the case of Mergeos, it is a strength because there are a lot of variables that must be accounted for with the data sets. Managing each item of every data set individually on a server is too resource intensive and prone to error. Putting the process of displaying the data on the user’s computer is more intuitive in this case. This is so because it removes the challenges of processing images of many data sets from the application without placing large demands on the users’ computers since it is displaying one data set at a time. In the

<table>
<thead>
<tr>
<th>APIs</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsoft Bing Maps</td>
<td>Provides an integrated platform for building map applications. Also has a wide range of flexible API choices like AJAX Control 7.0. See: <a href="http://www.bing.com/maps">http://www.bing.com/maps</a></td>
</tr>
<tr>
<td>Yahoo! Maps</td>
<td>Allows embedding of contextual and interactive maps in web applications. Flash, Ajax, and Map Image are platform choices that can be used. See: <a href="http://maps.yahoo.com">http://maps.yahoo.com</a></td>
</tr>
<tr>
<td>Google Maps</td>
<td>Provides a variety of APIs that can embed functionality or overlay personal data in web applications. JavaScript, Flash, and Static Maps are some of the APIs available. See: <a href="http://maps.google.com">http://maps.google.com</a></td>
</tr>
<tr>
<td>Google Earth</td>
<td>Provides a 3D digital globe of the Earth embedded into web applications by using the Google Earth plug-in and JavaScript API. See: <a href="http://www.google.com/earth/index.html">http://www.google.com/earth/index.html</a></td>
</tr>
</tbody>
</table>
process of planning the site, it was discovered that Google Earth is more flexible. It can also handle large sets of data points quicker as compared to the other server driven mapping solutions. Mergeos does not attempt to limit the amount of data in a single data set and it must host many different data sets. Therefore, the key performance element of server-based solutions and layer image caching is far less attractive. In addition, it is widely accessible to users, has KML support, easy to add 3rd party content, and a simple API to implement. Google Earth has a very fast and powerful KML parsing engine that allows Mergeos to pass KML data converted from uploaded data bundles directly to the Google Earth software for display. It also allows the KML to be compressed so the user does not have to wait long to download data and the development time to have this feature is much less. Fitting the needs of Mergeos so clearly, Google Earth was chosen to display the data sets in Mergeos.
Chapter 4: Analysis Results

This chapter describes the results of the site design and how Mergeos measures up to other VGI sites. The Google Analytics results and limitations encountered in the process of creating Mergeos are also addressed in the following chapter.

4.1 Site Design Results

The results of the site design process of Mergeos, which can be accessed at www.mergeos.com, defines the procedures and features as they are implemented in the procedure hierarchy in Figure 6 under 3.3 Site Design section. Screenshots of the finished website are provided with a description of how a user would navigate through each page to achieve their goals. The website results are organized by three sub-sections: search data sets, create a data set, and manage your data. The procedure starts with the homepage (see Figure 15), where the user will have the options to search for data sets, create a data set, read about Mergeos, or log into their account.

Figure 15: Mergeos homepage
4.1.1 Search Data Sets

Submitting a data bundle is the most important feature of Mergeos. A data set is comprised of any number of data bundles. Data bundles can also act as a way to organize the segments of data in a data set, by encouraging contributors to submit bundles for each area or category of data that they have collected. The metadata for the data bundles and datasets describes what data is in them. This includes title, date created, date updated, data description, and a list of contributors.

The user can search for a data set from the main menu. This involves entering search key words on the “Search Data Sets” page to determine if a data set, such as roads or lakes, already exists. By default, the results will show all existing data sets along with its title in alphabetical order, created date, and updated date (see Step 1 in Figure 16). Once the desired data set is found, the user has the opportunity to review it further by clicking on the data set’s link. The resulting Data Bundle page then further describes the data set and lists all of the data that has been added to it. All the contributors to the data set are also listed under the “Contributors” link, making the user more aware of where and what information the data set contains. If the user is the owner of the data set, the “Edit Data Set” link will appear allowing changes to the data set’s title and description. On the Data bundle page, the user can display the data set and view it in Google Earth by clicking on the “Load Data Set” button (see Step 2 in Figure 16). The data viewer has zoom, pan, and rotation controls. If the data set is not what the user needs then he or she has the option to keep searching or create
If the user finds that the data set fits their needs by reviewing all of this information or they wish to review it further in their own tools, they can use the “Download or View in Google Earth” link (see Figure 17). This link saves the complete KML file or opens it in the full version of Google Earth. A “Save As Default View” button is available to set the default view when the data set is loaded by using the map controls. The map controls are located on the left hand side of the data viewer. At the end of the page, users can enter comments and their thoughts about the data.

At this point, the user has made the decision to use the data or work on it further. They could also desire to share their work or add their own improvements to the data. This is done by adding a data bundle to the existing data set using the “Submit Your Data Bundle” link. On the contributors’ page, the contributors’ user names and roles are listed (see Figure 17). Refer to 4.1.3 Manage Your Data for details on setting contributors’ roles. Refer to 4.1.2 Create a Data Set section for details on adding, reviewing, and approving a new data bundle.
Figure 16: Search Data Set process

Figure 17: Data Bundle review page
4.1.2 Create a Data Set

From the main menu, the user can create a data set. To do so, the user clicks on the main menu button labeled “Create A Data set” and enters the title and description of their data set. Users that are not logged in are required to login or create an account before being forwarded to the “Create A Data Set” page (refer to Step 1 in Figure 18).

The next step is to populate the data set with at least one data bundle. To add a data bundle to an existing data set, the user clicks on the “Submit Your Data Bundle” when viewing a data set on the “Data Bundle Review” page. The user then enters the title and description of their data bundle and selects the appropriate KML file or zipped Shapefile (see Step 2 in Figure 18). Once submitted, the changes to the data set are not committed and shown to all users of Mergeos until the data set’s owners approve it. A notification is sent to the owner and displays in his or her “My Account” page upon login. See Figure 19 for the add a Data Bundle review notification to owner page. At this point, the user will wait until the owner approves or declines the data bundle. An email notification is sent to the user on the status of their data set addition. See Figure 20 for the email approval notification page.
Figure 18: Create Data Set process

Welcome Back, Chau!

You Have 1 Data Bundle to Review

Figure 19: Data Bundle review notification to owner

**Data Bundle Approved: Block Group**

Congratulations.
The following data bundle was reviewed by Sara Lee and was approved for addition to the data set.

Block Group
Block Group in MO View the Updated Data Set

Comments from Sara Lee:
approved
Thank you for making the availability of geographic data better for everyone through Mergeos.

Sincerely, The Mergeos Team

Figure 20: Approve e-mail notification
When the owner receives a message about a new data bundle that needs to be reviewed, the owner clicks on the link and is directed to the data bundle review page. On this page, there is instructional text that will assist the owner in reviewing the new data bundle to verify that it is a good addition to the final data set. The owner would use the supplied buttons to load the existing data set (if any) first and then load the new data bundle for comparison. It is then up to the owner to approve or deny the bundle and submit a response to the contributor by clicking the appropriate buttons. See Figure 21 for the data bundle instructions questions on the descriptions. The first two questions let users determine if the data bundle is relevant and useful to add to the existing data set.

At the middle of the Data Bundle review page, there is a data viewer that allows a user to compare existing data with new data bundle by using the “Load Data Set” and “Load New Bundle” button. This functionality permits a user to verify that the new data bundle will fit with the data set. See Figure 22 for data bundle viewer. The last step is to approve or deny the new data set. Only owners are allowed to complete this action (see Figure 23). Figure 24 displays the confirmation page once the new data bundle has been successfully reviewed by the data set owner.
Figure 21: Data Bundle descriptions

Figure 22: Data Bundle viewer
Figure 23: Data Bundle approval

Figure 24: Data Bundle review confirmation
4.1.3 Manage your Data

“There About Mergeos” is one of the main menu options. This page explains the purpose and function of Mergeos (see Figure 25). An email is provided on the website for users’ questions and feedback.

In order to add data to an existing data set, create a data set, or do anything with data bundles and data sets, a user account is required. When a user clicks the “Submit Your Data Bundle” link on the Data Bundles Details page, they are required to login using an account they created previously, or create a new account with Mergeos. Users that have not logged in will be led to the login page to either login with a current account or create a new account (see Figure 26). If users can’t remember their password, they have the option to have an email sent with their passwords (see Figure 27). There is also an option available to have the site remember the email address.

![Figure 25: About Mergeos page](image-url)
Figure 26: Login page

Figure 27:Forgot Password page
If the user decides to create an account from the login page, he or she can do so by clicking the link “I don’t have an account!” The user will fill out their names, alias, email, and password in the form that is displayed (see Figure 28). After the new account is submitted and created, the user is directed back to the login page.

Once logged in, the user will see the “My Account” page that is defaulted to the “Latest Contributions” page that lists the user’s latest contributions by details and status along with any data bundles that need to be reviewed. See Figure 29 for the “My Account” page. From this page, users can manage and view their data sets or data bundles. “My Data Bundles” lists details and status. “My Data Sets” lists details and owners are given the option to edit or delete their own data sets. See Figure 30 for the data bundle/data sets page.

![Create An Account](image)

**Figure 28: Create Account page**
Figure 29: My Account page

Figure 30: My Data Bundles / Data Sets page
Users can edit their data sets upon login from the “Data Bundles” page or after selecting a data set from the search results by clicking on the enabled link. Users can edit the details of only their contributed data bundles or data sets. See Figure 31 for the edit process and page.

As stated in section 3.6 Security and Roles, there are four roles: Owner, Contributor, Writer, and Manager. The owner is the creator of the data set and can edit, delete, review, and add all data in addition to managing the roles of other users. Owners can assign Contributor, Writer, and Manager roles to adjust the level of control and responsibility that each user has over a particular data set. A contributor has no write access to other users’ data, but can edit only their own data bundles. A writer has write and edit access to their own or other users’ data. A manager has write, edit, delete, and review access to the data and the ability to set roles on other users.

![Figure 31: Edit process and page](image)
To set contributors’ roles, the user would need to log in to their account by clicking the “My Account” menu button. Once logged in, the user would click on a contributed data set to be directed to the “Data Bundles” page. From here the “Contributors” link will redirect to a page of contributors with their current role list and available actions, if applicable. If the user is currently a manager or owner, they will see the option to promote any contributor to “Writer” or “Manager” and demote them to any of the three roles or remove them. See Figure 32 for the user’s roles process and page.

Overall, a site like Mergeos can be a step forward towards improvement and quality for GIS applications. In addition, monitoring activities can measure and assess data as well as the ability to control and reduce uncertainty.

Figure 32: User’s roles process and page
4.2 VGI Sites Comparisons

This section will display the analysis results with the comparisons of the community sites discussed in the Literature Review. Existing VGI sites involving OSM, GeoCommons, and ArcGIS Explorer Online are discussed in the following tables below by comparing them to Mergeos.

4.2.1 OpenStreetMap

Compared to OSM, there are several differences described as follows. OSM’s concept of having users as authors of their map is common to both OSM and Mergeos. However, it is done differently. In Mergeos, the author controls who can add to their data sets by accepting contributors’ data and can give certain permission roles to the contributors. This allows and encourages contributors to build a consistent data set for others to access rather than simply to accept any data submitted as is. OSM allows users to edit and save live data without the author’s authorization. Mergeos is also different from OSM in that users are not authoring a map with images and features but the data. OSM allows editing of points and offers a limited number of point types (such as hotel, café, airport, bus, etc.) for users to manipulate. It is similar to a Wiki where a topic is set and users contribute to it. Mergeos allows the owner of the data set to set the topic or a user can create their own. As for exporting, OSM offers four formats (XML, Mapnik Image, Osmarender Image, and Embeddable HTML) that do
not expose the raw data or allow users to upload their own data. On the other hand, Mergeos offers common exports of Keyhole Markup Language (KML) and can upload KML or Shapefiles. Figure 33 displays the Export tab from OSM.

Another difference is the way that both sites handle finding a map to start adding data. In OSM, users must search by location and add their points from a selective list to a collaborative map that already has contributions from others. There is also no way for users to add their own content since only specific points, not polygons or lines, can be added. Figure 34 shows the Edit tab from OSM.

Along with these features, a help section is provided in a blog format of asking questions and users' answers to them. This can be time consuming to search through if one is waiting on an answer. Mergeos does not have a help section (which is a future improvement) or blog section, but does allow users to submit comments about the data set. Overall, OSM is a great tool for all levels of users, regardless of their background, to contribute data and to show where interesting points are worldwide to the public. Table 10 displays a comparison table of OSM and Mergeos.
Figure 33: OpenStreetMap Export tab

Figure 34: OpenStreetMap Edit tab
### Table 10: Comparison list between OpenStreetMap and Mergeos

<table>
<thead>
<tr>
<th>Site Features</th>
<th>OpenStreetMap</th>
<th>Mergeos</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Map</strong></td>
<td>Displays an editable map with mapping tools to edit point data</td>
<td>No editable map, allows data set merging with vector data only</td>
</tr>
<tr>
<td><strong>Login</strong></td>
<td>Required when editing data or using a trace</td>
<td>Required when creating a data set or adding a data bundle</td>
</tr>
<tr>
<td><strong>Export</strong></td>
<td>Export to XML, Embeddable HTML, Osmarender Image, and Mapnik Image with selectable area specified by users which does not include the raw data</td>
<td>Export raw data to KML</td>
</tr>
<tr>
<td><strong>Site Layout</strong></td>
<td>Has six tabs: View, Edit, History, Export, GPS Traces, and User Diaries using OpenMap User Interface with Yahoo Maps as the background</td>
<td>Has four tabs: Search Data Sets, Create Data Set, About Mergeos, and Log In using GoogleEarth as the background</td>
</tr>
<tr>
<td><strong>Viewing Data</strong></td>
<td>Users can use the search to enter the location they would like to view, create a map under Edit, or select an existing edit under History. Able to review data set before downloading it.</td>
<td>Users must search for an existing data set or create their own data set. Able to review data set before downloading it.</td>
</tr>
<tr>
<td><strong>Adding Data</strong></td>
<td>Adding data with mapping tools under Edit tab. Allows live edits on the fly.</td>
<td>Merging and adding data set to someone’s existing data set by adding a data bundle</td>
</tr>
<tr>
<td><strong>Upload Data</strong></td>
<td>Does not allow users to upload their existing data</td>
<td>Allows users to upload their KML or shapefile (zip format)</td>
</tr>
<tr>
<td><strong>Bonus Features</strong></td>
<td><strong>History</strong> – Tracks everyone’s editing changes  <strong>GPS traces</strong> – Upload and view traces  <strong>Help</strong> – Questions can be posted here  <strong>User Diaries</strong> – Blogging entries from users</td>
<td><strong>Permissions</strong> - Allow data set owners to accept or deny new changes  <strong>Comments</strong> – Allow users to comment on others data sets</td>
</tr>
</tbody>
</table>
4.2.2 GeoCommons

Although GeoCommons has a similar concept, the export and upload file type options and manipulation of data differ from how Mergeos is designed. GeoCommons is comprehensive in the amount of features and options it is able to supply for users, but is still fairly easy to use and understand. However, users are not able to review the data before downloading it, after searching and finding maps. Figure 35 shows the Upload Data Page from GeoCommons. Mergeos has no interactive mapping tools because it is designed to merge multiple data bundles together into one larger data set and build upon that without forcing the potential users of that data set into one specific usage of their data. Figure 36 displays the Edit and Map page.

GeoCommons proves to be an excellent tool, but there are some areas of functionality that we hope Mergeos can focus on more successfully. These weaknesses include the lack of ability for users to add their own data that they already created using other tools. In addition, processing map features seems to be very slow due to the usage of Flex to draw on top of the Google Maps API display. Unlike Mergeos, there is no author’s approval that the data being added is acceptable for quality assurance. A comparison of the features in GeoCommons to Mergeos is presented in Table 11.
Figure 35: GeoCommons Upload Data

Figure 36: GeoCommons Edit and Make a Map
### Table 11: Comparison list between GeoCommons and Mergeos

<table>
<thead>
<tr>
<th>Site Features</th>
<th>GeoCommons</th>
<th>Mergeos</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Map</strong></td>
<td>Displays editable featured maps or basemaps with mapping tools to edit vector and raster data which does not merge data together</td>
<td>No editable map, allows data set merging with vector data only</td>
</tr>
<tr>
<td><strong>Login</strong></td>
<td>Required when creating and editing a map or uploading files</td>
<td>Required when creating a data set or adding a data bundle</td>
</tr>
<tr>
<td><strong>Export</strong></td>
<td>Export to KML, CSV, and shapefiles</td>
<td>Export raw data to KML</td>
</tr>
<tr>
<td><strong>Site Layout</strong></td>
<td>Two services with searching ability: Make a Map where users can create interactive maps with various tools and Upload Data which is a basic geocoding service to upload data. Google Maps API and Flex is used.</td>
<td>Has four tabs: Search Data Sets, Create Data Set, About Mergeos, Log In using Google Earth</td>
</tr>
<tr>
<td><strong>Viewing Data</strong></td>
<td>Users can immediately search for maps using the search functionality or make a map and upload their own maps/data. Users are not able to review data before downloading it.</td>
<td>Users must search for an existing data set or create their own data bundle. Able to review data set before downloading it.</td>
</tr>
<tr>
<td><strong>Adding Data</strong></td>
<td>See Upload Data</td>
<td>Merging and adding data set to someone’s existing data set by adding a new bundle</td>
</tr>
<tr>
<td><strong>Upload Data</strong></td>
<td>Batch uploader and geocoder available where data can be uploaded in CSV, Shapefile, KML, and RSS</td>
<td>Allows users to upload their KML or shapefile (zip format)</td>
</tr>
</tbody>
</table>
| **Bonus Features** | **Batch Uploader** – Can upload multiple files  
**GeoJoin** – Can connect names with common boundary data  
**Geocode** – Basic geocoding for addresses  
**Current Events** – Access to maps that are currently in the news (BP Oil Spills, Haiti Earthquakes, etc…)  
**Heat Maps** – Type of thematic map that can be created | **Permissions** - Allow data set owners to accept or deny new changes  
**Comments** – Allow users to comment on others data sets |
4.2.3 ArcGIS Explorer Online

Out of all the existing community sites involving the sharing of GIS data, ArcGIS Explorer Online is the closest one to the concept of Mergeos. Both are similar in the idea of the collaboration of data but different in user experience, as seen in the procedure hierarchy for Mergeos (see Figure 6). ArcGIS Explorer Online provides many great mapping features and tools to allow users to collaborate on maps such as ready-to-use common and topographic basemaps. These maps are then either reposted with added content or shared by sending a link to the maps. Due to Mergeos’ targeted audience, the raw data is far more valuable than a map alone; however, both can export the data into other formats like KML. ArcGIS Explorer Online is able to support rasters and vectors while Mergeos can only support vectors due to the limitations of Mergeos as a prototype. ArcGIS Explorer Online seems to be more comprehensive, with the ability to organize maps in folders, create groups, and create bookmarks, which are all good features to have.

Despite these great features, there is a weakness in that it does not provide an easy way to allow users to add their own data quickly, without using a separate ESRI service. It also does not allow a user to upload their own data. Instead, the user must create a map from scratch to be saved and then added on top of an existing basemap. Figure 37 displays the Presentation tab from ArcGIS Explorer Online. While this product is not very different from Mergeos, it further emphasizes the popularity of
ideas like Mergeos involving neogeographic concepts. It also proves why research into making these tools for data sharing within the GIS community more accessible is important. A comparison of the features in ArcGIS Explorer Online and Mergeos is presented in Table 12.

![Figure 37: ArcGIS Explorer Online Presentation](image-url)
Table 12: Comparison list between ArcGIS Explorer Online and Mergeos

<table>
<thead>
<tr>
<th>Site Features</th>
<th>ArcGIS Explorer Online</th>
<th>Mergeos</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Map</strong></td>
<td>Displays an editable map with mapping and presentation tools to edit vector and raster data. Allows Microsoft Bing maps.</td>
<td>No editable map, allows data set merging with vector data only</td>
</tr>
<tr>
<td><strong>Login</strong></td>
<td>Required when saving data</td>
<td>Required when creating a data set or adding a data bundle</td>
</tr>
<tr>
<td><strong>Export</strong></td>
<td>Export to and e-mail your maps directly from the application.</td>
<td>Export raw data to KML</td>
</tr>
<tr>
<td><strong>Site Layout</strong></td>
<td>Has two tabs: Presentation and Home</td>
<td>Has four tabs: Search Data Sets, Create Data Set, About Mergeos, Log In using Google Earth as the background</td>
</tr>
<tr>
<td></td>
<td><strong>Presentation Tab</strong> – Has text format tools and presentation options</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Home Tab</strong> – Has Search, mapping tools, and slide presentation options</td>
<td></td>
</tr>
<tr>
<td><strong>Viewing Data</strong></td>
<td>Users can immediately view maps using basemaps or the search options which are both fairly easy to use</td>
<td>Users must search for an existing data set or create their own data set. Able to review data set before downloading it.</td>
</tr>
<tr>
<td><strong>Adding Data</strong></td>
<td>Adding existing data with Add Content option under Home tab and can extend capabilities using Add-Ins and the Analysis Gallery options</td>
<td>Merging and adding data set to someone’s existing data set by adding a new data bundle.</td>
</tr>
<tr>
<td><strong>Upload Data</strong></td>
<td>Does not allow users to upload user’s existing data without additional services</td>
<td>Allows users to upload their KML or shapefile (zip format)</td>
</tr>
<tr>
<td><strong>Bonus Features</strong></td>
<td><strong>Presentation mode</strong> – Can setup a slide presentation to communicate data</td>
<td><strong>Permissions</strong> - Allow data set owners to accept or deny new changes</td>
</tr>
<tr>
<td></td>
<td><strong>Capture New Slide</strong> – Can take a screenshot</td>
<td><strong>Comments</strong> – Allow users to comment on others data sets</td>
</tr>
<tr>
<td></td>
<td><strong>Access to Mapping Services</strong> – Access to ArcGIS Server, ArcIMS, GeoRSS feeds, etc…</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>2D and 3D Display Modes</strong> — Switch display between maps and globe modes in any projection</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Ratings</strong> – Users can rate maps</td>
<td></td>
</tr>
</tbody>
</table>
4.3 Google Analytics Results

At this time, there is a small number of users testing the site. Google Analytics is a free service that was used to monitor and produce statistics about visitors’ website activity. In January 2011, there was a 57.14% increase in visits since the site creation in October of 2010. Figure 38 shows the Google Analytics results for the month of January 2011. The graph indicates that the highest activity was in January 21, which makes sense as reminders were sent on January 20. The lowest activity is at the beginning and end of the month. Users spent an average of about 8 minutes on the site at a total of 11 visits for the month of January. Only 1 visitor “bounced,” meaning they left the site rather than continuing to browse. In addition, there were 2 new visitors and an average of 5 pages per visit.

![Google Analytics Results in January 2011](image)

**Figure 38: Google Analytics Results in January 2011**
4.4 Future Promotions

In the future, the site will need some exposure to various people which can be achieved effectively through online social promotions. Some social applications that may increase site traffic to Mergeos are:

- Twitter
- Facebook
- Google AdWords
- Google Earth gallery
- Conference presentations
- Mailing list

Twitter and Facebook are two of the most common means of online communication that are currently being used. Users can post and share updates on Twitter about Mergeos. A Facebook page can be created for the site where users can “Like” the page to share amongst their friends. In addition, a “Like” button can be displayed next to data bundles on the site that would share the link through the user’s Facebook page. Another potential promotion is Google AdWords where the site can display in Google's search results based on certain keywords entered in the search field. The only downside is this service is not free. Posting data bundles in the Google Earth’s gallery is another way to entice more users and make them aware of the site’s existence. Another way to market the site is through conference presentations at
various GIS events. Doing so can promote the site’s purpose and existence to targeted audiences (students and professionals). At a cost, having a vendor booth at conferences can supply networking opportunities to discuss the site. A mailing list can be offered to potential users at these events. Also, a link to the mailing list can be added to the site or various social sites like Facebook to notify users of Mergeos’ news and updates. A banner can be created to share to users who would like to post it on their sites or distribute to their friends. Any of these social mediums will enhance exposure and draw more users to the site.

As various social promotions are used, eventually “word of mouth” will be another way to promote Mergeos, especially in the case of a natural disaster or tragic “event” where people are seeking data. If an event occurs, Mergeos can take the opportunity to advertise in several ways. One is to encourage users to create a data set or work on data sets as a humanitarian cause by using banners or social media. Another way is to "feature" data sets about similar events or bring attention to promising data sets about the events to encourage collaboration. Advertising directly to people and agencies affected by the event that they can share their information using the site is another example. The hope is that eventually word will spread about Mergeos and target more users that can benefit from the site.
Chapter 5: Conclusion

The completion of Mergeos demonstrates a solution in data accessibility to a variety of collaborative and unique data sets for the public. Mergeos has met the objectives of documenting the creation of a website based on neogeographic concepts and allowing volunteers to share merged GIS data. Due to the success of the site’s development, it has proven that a community site is possible even with limitations, such as the availability of affordable and technical resources. Supplying documentation on such a community site helps us explore concepts of collaborative GIS and data management as well as allowing standards to be set and controlled by map owners. Mergeos shows much of the potential that neogeographic and VGI concepts have to offer.

Some ideal scenarios where Mergeos could be used are:

- Students collaborating on uncommon data sets
- Combining and sharing related data quickly like disaster data
- Managing data sets with remote team members that have different software
- Publishing data sets for others to review

Mergeos can benefit the community by supplying uncommon data sets such as cancer data by state. The site can be used to combine unexpected or hard to collect
data like natural disaster information. For example, tornado or hurricane data can be easily collected with no up-front cost and low requirements from contributors to help a city survey the damage and aid in reconstruction. This would depend on the number of volunteers who post their data with a similar type of information in order to be effective. This could be orchestrated very easily with Mergeos.

Another benefit is that users can upload their data from anywhere that has internet access. Therefore, communities can share and publish data from different locations despite the software they are using. However, the GIS data being uploaded must be in either a shapefile or .KML zip file for data display to work properly on the site.

Despite the site’s success, there are some limitations and challenges that were encountered when designing and developing Mergeos to accomplish the powerful functionality these new types of web applications provide. Even the solutions used to solve those challenges were arduous once they were found. Table 13 reviews these challenges with their possible solutions.

Unfortunately, a feasible way has not been found to add the regular map views; therefore, data will be displayed on Google Earth’s satellite virtual globe view. Other issues are that road names do not display very well and not all areas are covered or properly mapped to a good resolution. When it comes to reviewing a data set in the data viewer and depending on the data size, it does not draw the data correctly or at
## Table 13: Challenges and Limitations

<table>
<thead>
<tr>
<th>Challenge Number</th>
<th>Challenge / Limitation</th>
<th>Solution (if any)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Google Earth does not automatically zoom in to data set in the data viewer on the site</td>
<td>Have data set managers set their data set’s default area using the provided controls</td>
</tr>
<tr>
<td>2</td>
<td>No Rasters</td>
<td>There is no solution for this, may be a future improvement</td>
</tr>
<tr>
<td>3</td>
<td>No other layer sets for the data viewer beyond satellite and terrain style views</td>
<td>A potential solution is bringing map layers from different mapping applications into Google Earth, but will require more work as a future improvement</td>
</tr>
<tr>
<td>4</td>
<td>Integration with other online GIS applications and social networks to expand Mergeos’ reach to the wider GIS community</td>
<td>Solutions for incorporating Mergeos better with Google Earth layer and Facebook have been released since Mergeos’ development that could be integrated in the future</td>
</tr>
<tr>
<td>5</td>
<td>Since the exports are in KML format, the default projection is WGS 84 (EPSG 4326) which could cause some data set not to be upload correctly</td>
<td>Sometimes re-uploading the data set or converting shapefiles into KML file to upload would work</td>
</tr>
<tr>
<td>6</td>
<td>The ability to convert KML into shapefiles to be imported in various GIS software</td>
<td>There are scripts on the net and ArcTool box scripts for ArcGIS to apply the conversion</td>
</tr>
<tr>
<td>7</td>
<td>Large data sets may take a while to display</td>
<td>A loading progress icon is displayed to show progress and the large set will appear in time</td>
</tr>
<tr>
<td>8</td>
<td>Not enough testing from various students and instructors due to time</td>
<td>If this project continues, time can be made for a thorough beta testing from academic audiences</td>
</tr>
<tr>
<td>9</td>
<td>Depending on how raw data was created, some data have junk data since the data display as it stored in the database</td>
<td>For future enhancements, create a minimum standard of required data columns to prevent some junk data</td>
</tr>
</tbody>
</table>
all. However, there were no issues with the data set when users downloaded the file by clicking the “Download or View In Google Earth” link. This will display the data set and draw it faster in Google Earth than the data viewer.

Limitations and issues are expected as long as the overall goal is still being achieved. Fortunately, the solutions to the problems of finding unique GIS data will help lay the foundations on how to approach and develop a community site with GIS data. Applying new concepts and determining the success of the design patterns and methodologies will also be beneficial for others to learn from.

There are several improvements that would be beneficial in Mergeos’ outlook, including:

- Adding the ability to handle raster data
- Defaulting to the desired area in the data viewer instead of having to zoom in or zoom out and choose the desired area
- Adding a help system that includes tutorials and video demos
- Adding additional data details such as data type, data source, etc.
- Allowing upload of bulk data sets
- Creating a minimum standard of required columns needed for uploading data
- Providing more time for website beta-testing from more students and instructors
• Creating a survey or feedback page for suggestions and website improvements

The website can be accessed at www.mergeos.com. Feedback and suggestions are always welcome and can be sent to chau.nguyen25@gmail.com.
References


Appendix A – Term Glossary

**Adobe Flash** - is an Adobe Systems multimedia platform that is commonly used for animation and dynamic web pages which was originally called Macromedia Flash [Blue Icon Media 2010].

**Adobe Flex SDK** - is an Adobe Systems software development kit involving cross-platform rich internet applications based on the Adobe Flash platform. By using an Adobe Flex Builder or a free Adobe Flex compiler, Flex applications can be written [Wikipedia – Adobe Flex 2001].

**Apache Webserver** - is a public domain open source web server created by a group of programmers [Webopedia 2011].

**Cascading Style Sheets** – is a style sheet language that is used to display the look and formatting of a markup language document [Wikipedia – CSS 2001].

**ColdBox** – is a ColdFusion toolkit of software and framework for web applications.

**ColdFusion** – this application development and scripting language was built to handle simple HTML pages to a database. It has grown into a feature rich, dynamic language and web application server focused on integration with a wide range of technology.

**Creative Commons Attribution Share Alike 2.0** – Creative Commons is a non-profit association that provides free licenses that they share to the public and license that governs the original work [Wikipedia – Creative Licenses 2001].

**Environmental Systems Research Institute** – a software development company that offers Geographic Information System solutions, applications, and software such as ArcGIS Desktop [Wikipedia – ESRI 2001].

**GDAL** – stands for Geospatial Data Abstraction Library and is a translator library for raster geospatial data formats. This was released under an Open Source license by the Open Source Geospatial Foundation [GDAL 2011].
**Geographic Information System** – a system that manipulates, manages, stores, and analyzes data and displays data visual through computer software [Wikipedia – GIS 2001].

**GeoServer** – open source Java application that provides a way to share and edit spatial data.

**Google Maps API** – the development API used in any Google Maps application, with different map types like satellite, regular map, and terrain.

**Information and Communication Technology** - where concepts, methods and applications involved are continuously evolving [Tutor2u 2007].

**Java** – a language similar to C++ that was created by Sun Microsystems. This object-oriented language built for a wide range of platforms and devices is now one of the most common languages used for computer software and web applications.

**jQuery** – is a quick and concise JavaScript library that simplifies HTML documents, event handling in web development [jQuery 2009].

**KML**– stands for Keyhole Markup Language. It is based on the XML language that expresses geographic visualization from the internet with two-dimensional maps and tree-dimensional Earth browsers (such as Google Earth).

**Linux OS** – is a free operating system based on Unix-type created by Linus Torvalds and other developers around the world [Real Time Operating System 2010].

**MySQL** – is an open source relational database engine offered freely and presently owned by Oracle.

**.NET** – is a software framework that is installed on computers with Microsoft Windows OS.

**Neogeography** – means “new geography” that applies geographic techniques and tools used in community related activities which were officially announced in 2006 [Wikipedia – Neogeography 2001].

**OGR2OGR** - converts simple features data between file formats [Linux-Support.com 2010].
**Silverlight** – a Microsoft product that is an influential development platform for interactive applications [Boardreader 2011].

**SQL** – stands for Structured Query Language which is a database language used to manipulate data in a relation database management system [Wikipedia – SQL 2001].

**SQL Server Express** – is a free version of Microsoft’s SQL Server relational database management system focused on small scale applications [Wikipedia – SQL Express Server 2001].

**User Interface** - is what people use to interact with a machine or use to maneuver around a site or software.

**Volunteered Geographic Information** – are tools to create geographic data willingly provided by individuals [Wikipedia – VGI 2001].

**Web 2.0** – are web applications that interactively share information and allow collaboration among users.

**XHTML** – is an Extensible Hypertext Markup Language that extend versions of HTML (Hypertext Markup Language) that web pages are normally written in [Wikipedia – XHTML 2001].

**XML** – is an Extensible Markup Language used to carry and store data.
Appendix B – Logo References


Appendix C – Term Glossary References


