Free and Open Source Software for Geospatial (FOSS4G) Conference Proceedings

Volume 15 *Seoul, South Korea* Article 38

2015

Integrating Open Source GIS Software in Undergraduate Curriculum, Research, and Outreach - Recent Experiences at Salisbury University

Arthur J. Lembo, Jr.
Department of Geography and Geosciences, Salisbury University

AT5-04

Integrating Open Source GIS Software in Undergraduate Curriculum, Research, and Outreach - Recent Experiences at Salisbury University

Arthur J. Lembo, Jr.[1]

[1]Department of Geography and Geosciences, Salisbury University
1101 Camden Ave, Salisbury, MD, United States
Email: ajlembo@salisbury.edu

ABSTRACT

The Department of Geography and Geosciences at Salisbury University has a long tradition of teaching geographic information science. Until recently, most of the courses and research activities focused on commercial software offerings. However, the Department recently integrated Free and Open Source Software for GIS (FOSSG) into its curriculum, research, and outreach. Curriculum changes included introducing students to FOSSG in traditional GIS courses using QGIS, and creating two entirely new courses in Enterprise GIS and GIS Programming using PostGIS, GDAL, and SpatialLite. Through a competitive National Science Foundation (NSF) Research Experience for Undergraduates grant (REU), students participated in cutting edge research projects in parallel processing with Hadoop and spatialHadoop for cluster computing, and CUDA for GPGPU calculation on embarrassingly parallel processes for raster data. Finally, undergraduate interns working in the Department's Eastern Shore Regional GIS Cooperative (ESRGC) developed geodashboards using node.js, PostGIS, and Leaflet, while a special topics course developed a GIS based iphone and Android application used by 4,000 participants in the annual Sea Gull Century bike ride using GeoJSON, Leaflet, and javascript. In addition to highlighting the successes of these activities, this paper will discuss the process we used to make the necessary changes in our curriculum, secure the necessary funding for external projects, and the training approach we used to get our computer science students proficient in programming with FOSSG tools.

1. INTRODUCTION

The Department of Geography and Geosciences (DoGG) at Salisbury University has a long tradition of teaching geographic information science. As an academic institution, the DoGG focuses on three primary areas: teaching, research, and outreach. Until recently, most of the courses and research activities focused on commercial software offerings. However, the Department has recently integrated Free and Open Source Software for GIS (FOSSG) into its teaching curriculum, research, and outreach.

While many of our students have obtained employment using commercially available GIS software, a growing number of alumni have migrated their work into the open source realm. Additionally, more job postings include a preference for open source software. In response to the increased use of open source GIS software, the DoGG Faculty made a decision to include more offerings of FOSS4G into the student experience. This included introducing FOSS4G in our teaching, research projects, and outreach experiences.

2. TEACHING

In addition to concentrations in Atmospheric Science, Planning, Human Geography and Earth Science, the Department offers numerous courses in GIScience. These courses include *Map and Air Photo Interpretation, Introduction to GIS, Advanced GIS, Remote Sensing, GIS Modelling, GIS Programming, Enterprise GIS,* and *Cartographic Visualization*. Over the last two years, four of our courses were adapted to include FOSSG software including *Advanced GIS, GIS Programming, Enterprise GIS,* and *Cartographic Visualization*.

2.1 Advanced GIS

The Advanced GIS course is a study of current theories and development in GIScience, with an emphasis on using many of the specialized tools in ArcGIS including Network, Spatial, 3D, and Geostatistical Analyst. While our students were gaining a significant amount of advanced skills in GIScience, they mostly understood the world through the ArcGIS lens. Because the job market is dominated by ESRI software, it was impractical in our undergraduate institution to ignore the importance of the ESRI software. However, to supplement the students' ArcGIS skills, we introduced modules using QGIS in order to replicate some of the laboratory exercises they had previously completed using the ESRI suite of products. The main goal was to show the students that open source options exist, and that these options were comparable to the commercial offerings. Because QGIS was yet another technology stack introduced within a semester-long course, we made our exercises more *cookie-cutter* in approach, just to give the students the opportunity experience the overall feel of QGIS, and experiment with the various plug-ins, comparable to the ArcGIS *Analyst* extensions.

The benefit of introducing FOSSG software into the Advanced GIS course not only provided the students with experience using another tool, but to also expanded their intellectual capacity to evaluate the benefit of using these tools within a GIS context. We wanted our students to be able to articulate the pros and cons of both commercial and open source GIS software. The results were remarkable. Not only did our students learn and understand FOSSG software, but during many job interviews, our undergraduates found themselves teaching their prospective employers about the availability of such software, and were also able to discuss the pros and cons of using the software within the employer's existing structure.

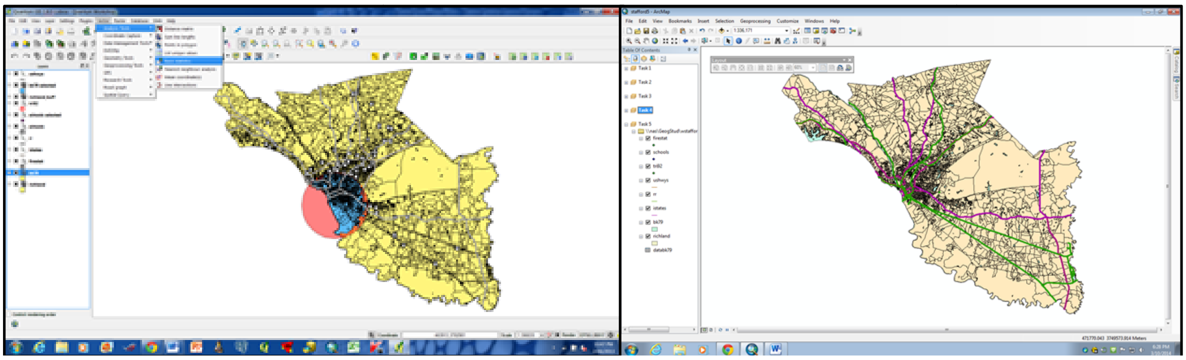

Figure 1 Richland Chemical Lab, performed by students using shown QGIS (left) and ArcGIS (right)

2.2 GIS Programming

Our GIS Programming course focuses on developing custom computer programs that address classical problems in geography and spatial analysis not ordinarily solved using out-of-the-box commercial GIS software. Here again, the dominant industry application is ArcGIS, so we focus 50% of the course on using Arcpy. The benefit is significant as few undergraduate geographers have experience programming in Arcpy. However, similar to our

Advanced GIS course, we recognized the need to expose our students to FOSSG software with respect to the programming environment. However, we did not want to simply replicate the Arcpy experience as a one-to-one match (like attempting to teach Qpy), but rather wanted to expose our students to technologies that the current ESRI offerings could not replicate.

Therefore, the remaining 50% of the course focused on the use of spatial SQL using PostGRES/PostGIS for solving spatial problems. For SQL, the focus was on traditional SQL statements, SQL data types, aggregate functions, and the spatial functions offered by PostGIS. Students worked through the book *How Do I Do That in PostGIS*, to learn how to perform the classic GIS tasks outlined in the book *How Do I Do That in ArcGIS/Manifold*, but within an SQL paradigm. Oftentimes, the laboratory exercises recreated the ArcGIS exercises in a SQL paradigm, allowing the students to see alternate, and oftentimes more efficient ways of completing a task.

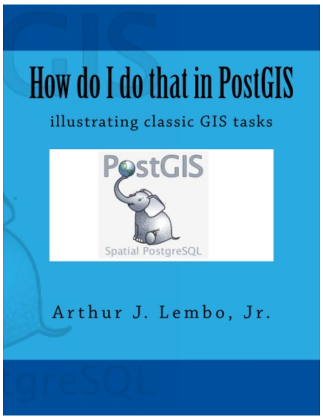

Figure 2 Textbook used to teach students spatial SQL in PostGIS

Similar to the Advanced GIS course, the results were remarkable. In using SQL to solve complex GIS problems, our students discovered a confidence they hadn't experienced before. Also, during the independent final course project, most students opted to utilize SQL to complete the project as it was faster, more elegant in its approach, and easier to write.

2.3 Enterprise GIS

Our Enterprise GIS course is a special topics course offered every few years. The special topics course is usually only open to those students who have been introduced to SQL in the GIS Programming class. In this course, the students implement an enterprise GIS using PostGRES/GIS as the spatial data server, with ArcGIS, Manifold GIS, and Quantum GIS as the client applications. The students focus on database design and management of user roles, data views, triggers, and the development of end user applications. Typically, the design includes users (creatively named Mo, Larry, and Curly) with different editing capabilities. The context of the project was to create an Enterprise GIS in PostGRES for the Town of Denton, MD, as it was the same town the students created an ESRI file geodatabase for.Also, the students then perform various operations to understand the multi-user nature of spatial databases, such as multi-user editing, or perform large spatial selections using different underlying databases.

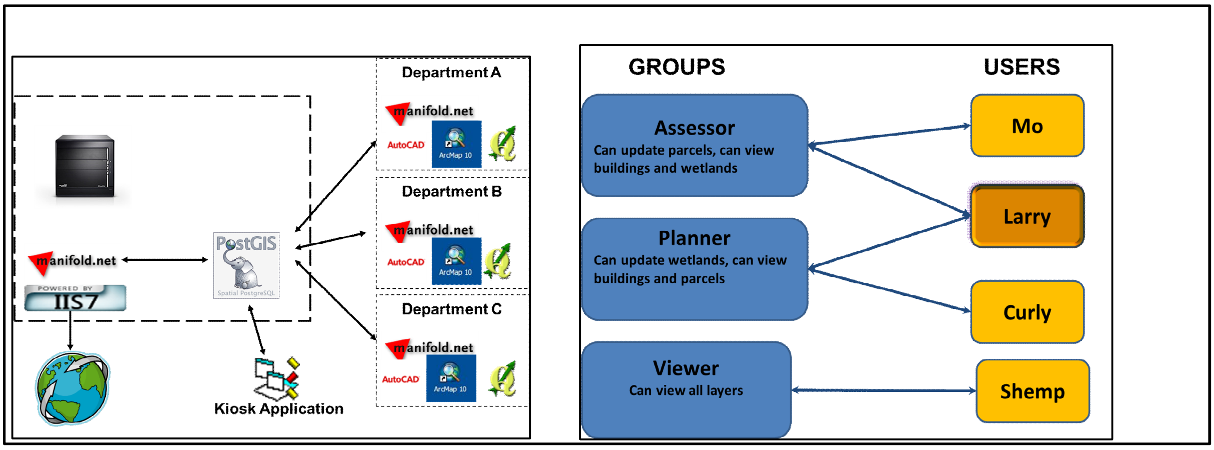

Figure 3 Enterprise GIS class designed and implemented a multi-user, multi-vendor client/server GIS

Again, the results of the course were remarkable, with our undergraduate students learning how to build an enterprise GIS with multiple users and multiple client applications. In addition, our students presented their work on implementing a multi-user, multi-vendor enterprise GIS at two regional GIS conferences. Finally, having previously built a File Geodatabase in the Advanced GIS class, students could articulate the strengths and weaknesses of both models.

2.4 Cartographic Visualization

The Cartographic Visualization class introduces students to the theory and application of cartographic principles and practices in cartographic design. Commercial products like Adobe Illustrator are used with ArcGIS for half of the course, as the students are already familiar with ArcGIS. However, the other half of the course introduces students to MapBox and TileMill so that students become familiar with open source cartography principles. Finally, student are required to make interactive and animated maps for final projects TileMill and MapBox.

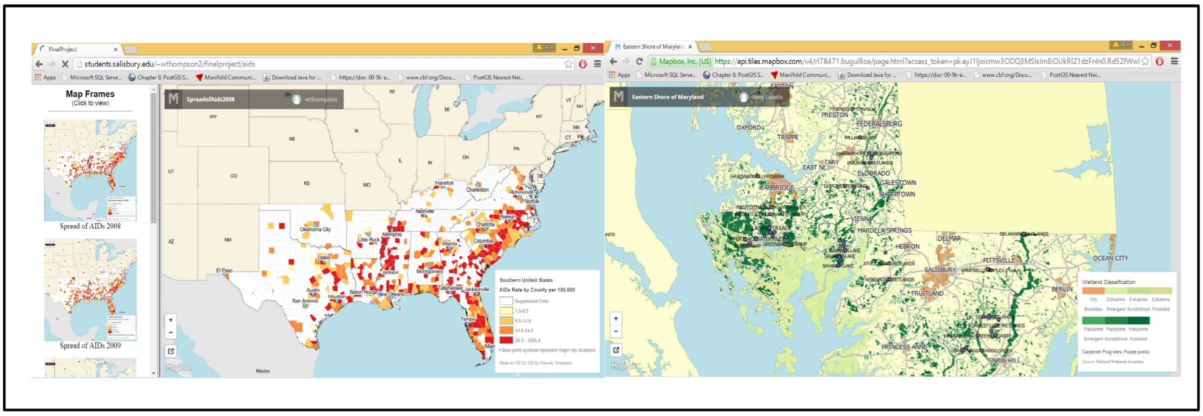

Figure 4 Examples of student final projects include interactive maps produced using MapBox and TileMill

3. RESEARCH

As an undergraduate institution, Salisbury University emphasizes research experiences for undergraduates. Research take the form of traditional sponsored research, and directed research projects by undergraduates with faculty mentors. The use of open source tools were used in sponsored research on parallel processing for the National Science Foundation (NSF) and food security mapping for the United States Department of Agriculture (USDA). Directed research was conducted as part of the Government and Non Profit Application Workshop (GNAppWorks) and use mapping APIs for Android devices.

3.1 The National Science Foundation

As part of a competitive grant, Salisbury University was awarded a 3 year NSF REU in parallel processing. The REU included 8 undergraduates each summer, working with 4 faculty mentors. Generally, the REU projects focused on classical parallel processing projects in the areas of cryptography, tomography, and graph theory for social networks. However, each summer, a parallel processing project in GIS was undertaken.

3.1.1 Parallel Processing with General Purpose Graphical Processing Units

An NSF REU gave undergraduates an opportunity to write raster based terrain algorithms utilizing GPUs on video cards. The students wrote the algorithms in C++, and utilized NVIDIA APIs. Throughout the process, the students not only improved the speed of the analysis, but also developed work-arounds for I/O bottlenecks to achieve greater speed improvements. The students developed algorithms for embarrassingly parallel applications including slope, aspect, and terrain ruggedness. For the kernel based terrain functions, students saw an improvement of 10x when using GPGU processing. However, when the students increased the kernel size (and thus the number of computations), the speed improvements approached 60x.

Figure 3 shows the general process used for performing the GPU analysis, along with a breakdown of times for input, output, and processing under a parallel and serial implementation. Most notable is the reduction in processing time of the GPU implementation compared to the serial implementation.

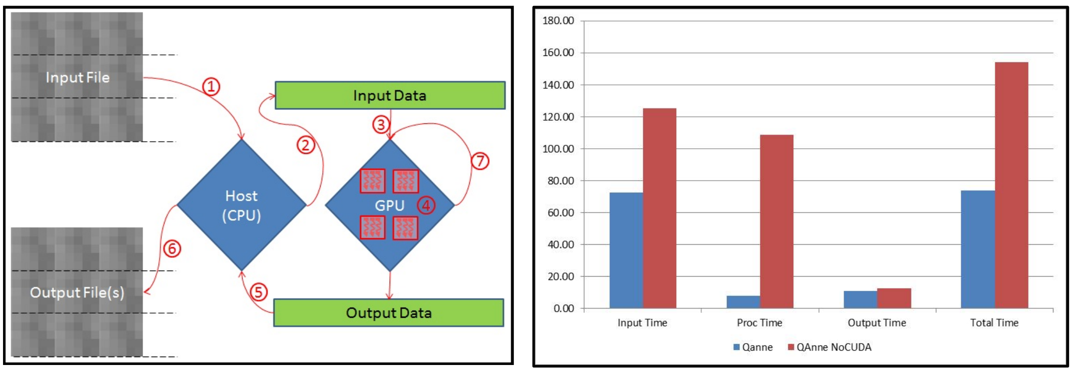

Figure 5 NSF research performed raster GIS analysis using video card showed dramatic speed improvements when utilizing GPUs on video cards

3.1.2 Parallel Processing with Multiple CPUs and Hadoop

In addition to GPU processing, students also experimented with distributed computing, utilizing the open source spatial Hadoop on very large spatial databases. The results demonstrated that the classic point-in-polygon test on tens of millions of points that required five days of processing with commercial GIS software could be completed in 200 seconds using Hadoop and special modifications to the underlying data structures. Within this project, students built their own cluster in the lab, in addition to renting time on an Amazon EC2 server. The ability to affordably rent 20 computers with SSDs on Amazon reduced the processing time to seconds, while our own cluster of 4 computers completed in 24 minutes.

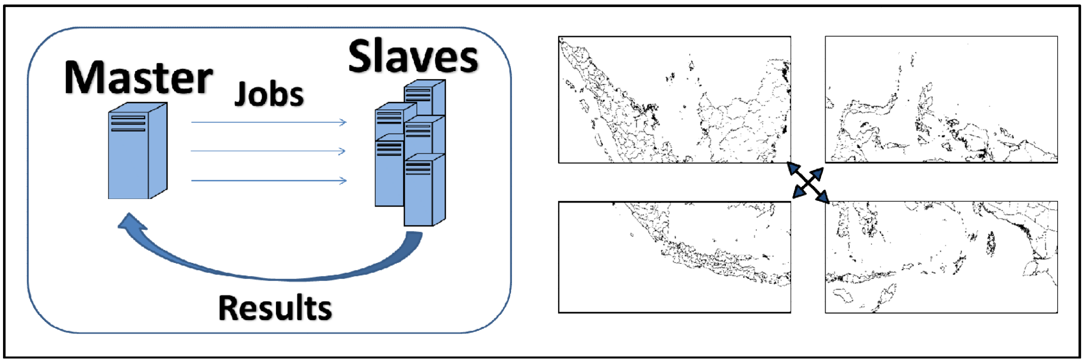

Figure 6. The use of spatial Hadoop splits up the data to work over multiple machines simultaneously to complete the task much faster than traditional GIS software on bigger data.

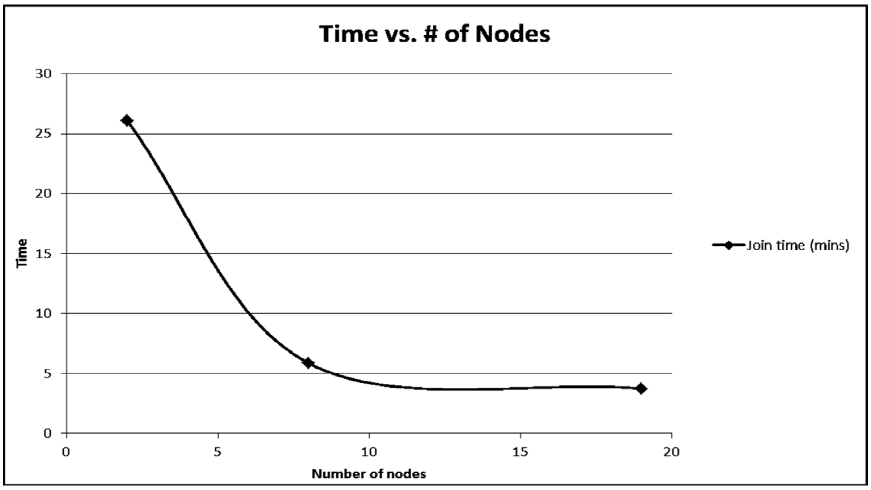

Figure 7 Results of the Hadoop project showed the runtime of the spatial join on a Hadoop cluster of n computers working together. The X values represented 2, 8, and 19 CPUs.

3.2 The United States Department of Agriculture – Foodshed Mapping

The Mapping Local Food Systems project, funded by the U.S. Department of Agriculture, examined New York's ability to supply its own food needs. The project was a joint effort between Salisbury University, Cornell University, and Tufts University. The Salisbury University portion was to create an Internet Mapping Service (IMS) to visualize the optimization algorithms developed as part of the research.

The IMS solution included the query and display of foodshed footprints for individual cities, and was implemented with GeoJSON as a datasource, and Leaflet.js as the client GUI.

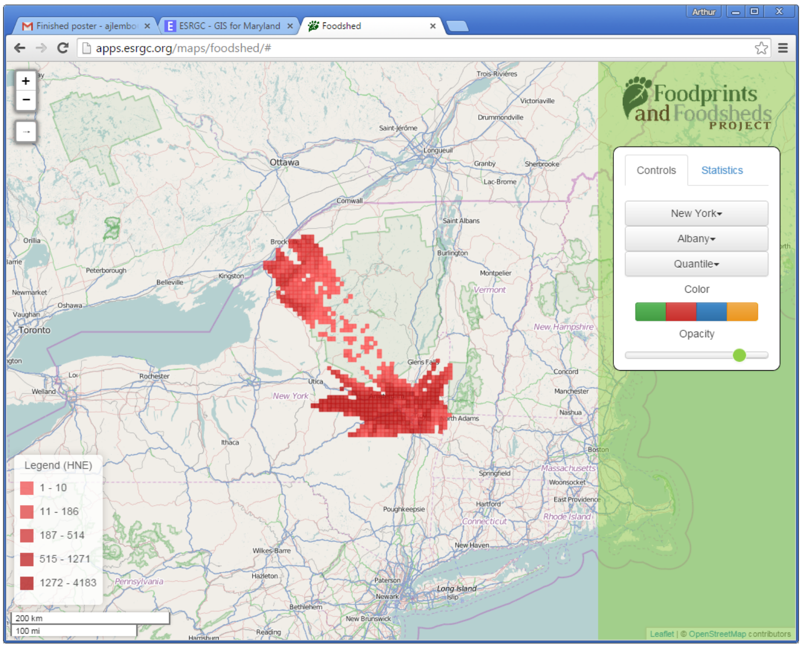

Figure 8 The USDA Foodshed project utilized GeoJson data files, Leaflet, and OpenStreeetMap to display foodsheds for selected cities in New York State

3.1.3 GNAppWorks – The Sea Gull Century

In order to integrate our teaching and outreach programs, the Geography Department, the Henson School of Science, and the Eastern Shore Regional GIS Cooperative created the Government and Non-profit Application Workshop (GNAppWorks). GNAppWorks is a mobile application development initiative focused on providing applications to the government and nonprofit sector.Nonprofits with increasingly smaller budgets, citizens in lower socio-economic classes, and small government entities have needs for mobile applications, but lack the means to develop or fund development of high quality products. With the backing of Salisbury University, GNAppWorks provided computer science majors with an opportunity to work on real-world smartphone applications under the direction of a Faculty GIS mentor. Some of the clients have included The Nature Conservancy and the City of Havre de Grace.

As part of a research class, Computer science, Geography, and Business students worked collaboratively to develop Android (and iPhone) applications to guide cyclists participating in the Sea Gull Century bike ride.The application was used by thousands of riders, and showed specific rider locations along the three Lower Shore routes, displaying their speed, nearby vendors and rest stops. Geography undergraduates digitized the bike routes, the computer science students programmed the application, and the Business students worked with ride sponsors to participate in the marketing of the ride. They also worked with the University legal department to ensure that the application utilized the appropriate logos, colors, and descriptions. The students also gained an understanding about the collaboration required within a small start-up organization.

In addition, the research project, "Creating the Sea Gull Century Mobile Application – A Multi-Discipline Effort," was presented at the University of Maryland Eastern Shore's fifth annual Regional Research Symposium, and earned first place in the undergraduate oral presentation category.

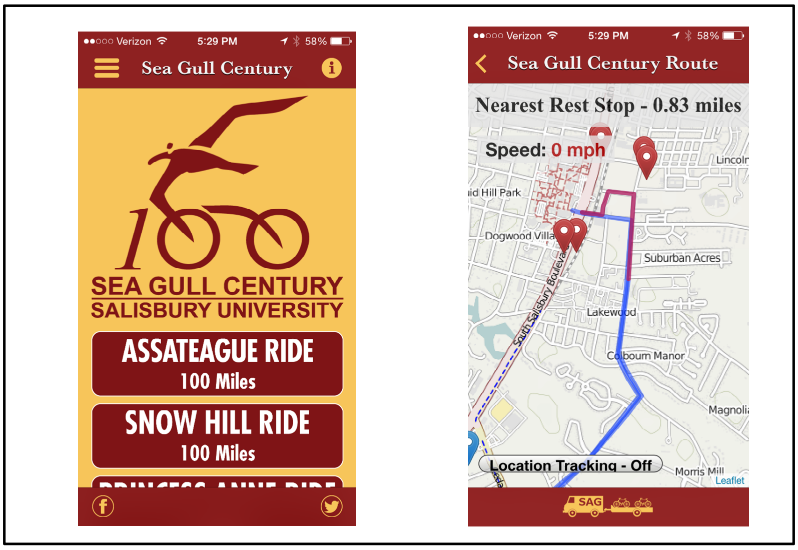

Figure 9 An Android application using Leaflet, GeoJSON, and OpenStreetMap was used to assist riders in the annual Sea Gull Century ride.

4. OUTREACH

4.1 The Eastern Shore Regional GIS Cooperative

The original goal of Salisbury University'sEastern Shore Regional GIS Cooperative (ESRGC)was to provide GIS and other high tech services to under-served markets on the Eastern Shore of Maryland. Over the years, the ESRGC has expanded to include larger work assignments, and now have a staff of 14 full time employees including geographers, computer scientists, and database analysts. As these work assignments grew into more IT related endeavors, our work transitioned from cartography and basic spatial analysis tasks to greater use of database technology, Internet mapping, and spatial analysis. As part of the educational experience, the ESRGC hires undergraduate Computer Science and Geography students to work collaboratively to develop IT based applications.

Early efforts at using commercial based GIS software for Internet delivery proved frustrating, as the commercial products inhibited the creativity the staff desired. So, while commercial GIS software did exist to support many of our projects, a decision was made to migrate almost all of the development work into a FOSS4G paradigm, as the open source solutions allowed us to provide greater flexibility and responsiveness to the application. In addition, the FOSS4G solutions allowed our clients to enter into Internet based IT solutions for a minimal cost.

Examples of some of the ESRGC activities with FOSS4G include our use of GeoDashboards for the State of Maryland, Oklahoma, and the Federal Government, the GNAPPWorks Initiative, and municipal applications for local and regional organizations.

4.1.1 GeoDashboards

Although each geodashboard we create are different, they follow the same general architecture:

1. Data is stored in PostGRES with PostGIS
2. The web server is node.js
3. Front end web pages are written in leaflet.js

4. Data is sent to the web pages either as GeoJSON directly from PostGIS, or as WMS services from TileMill.

The use of Leaflet allows us to create a very responsive GUI, and our connection with PostGRES/PostGIS allows us to issue virtually any kind of SQL or spatial SQL query.

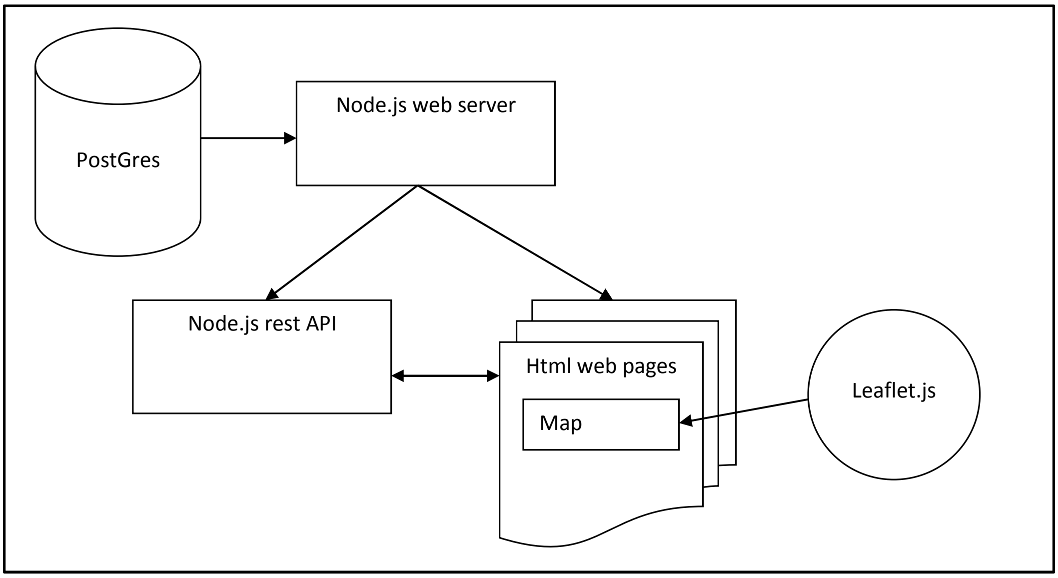

Figure 10 The general architecture of the ESRGC dashboards include the use of node.js as a web server, PostGRES as the back end database, and Leaflet.js for the client-side interface.

Examples

The ESRGC completed geodashboards for Maryland's BayStat program, the Maryland Department of Labor, Licensing, and Regulation (DLLR), the State of Oklahoma Licensing and Regulation, and the National Broadband Mapping Initiative, all using the architecture described in Figure 8.

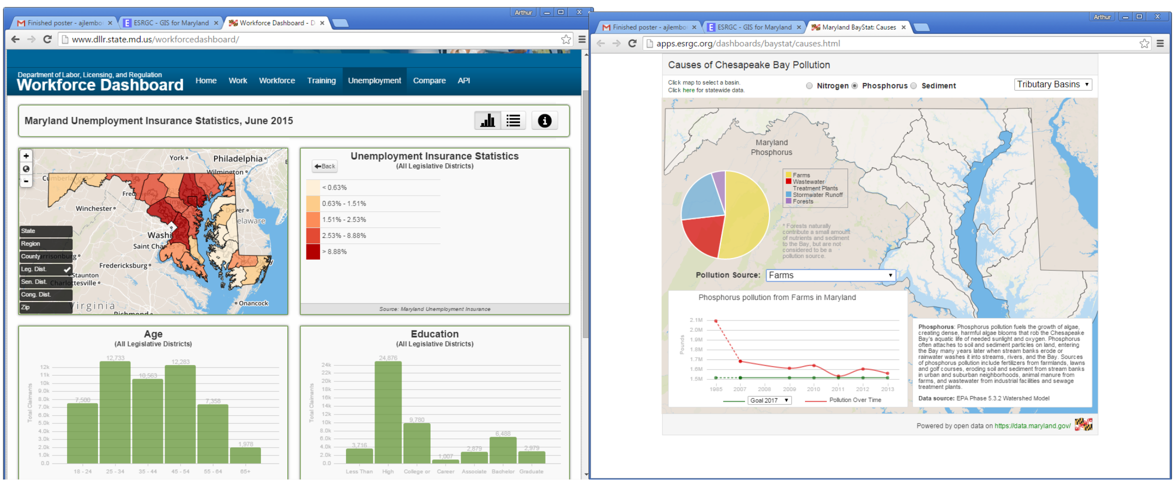

Figure 11 Example of Open Source Dashboards for the Maryland Department of Labor, Licensing, and Regulation, and the Maryland BayStat Program.

5. CONCLUSION

While most undergraduate GIScience programs focus on the use of a single commercial vendor, Salisbury University sought to be broader in our approach, introducing students to different commercial GIS software, in addition to including open source tools in our key GIS courses. In addition, faculty focused on research and outreach projects that included open source tools, giving our students a greater breadth and depth in GIS solutions to problems. Obviously, integrating an open source stack into existing University offerings requires greater dedication and time requirements for the faculty, the results in our course evaluations and Senior exit interviews show that students not only appreciated being exposed to more options, but that the experiences gave them greater opportunities when searching for employment and graduate school admission.