**A proposed open source architecture for collecting and analyzing performance indicator data from the United States Agency for International Development**

Erin C. Goodnough

Pennsylvania State University

Department of Geography/World Campus

College, PA 16802

**ABSTRACT**: Aid donors fund numerous projects, using contractors to implement their programming objectives globally. These large project portfolios create significant challenges in monitoring progress, evaluating program efficacy, and making evidence-based decisions in developing countries. Current performance analysis in the aid community remains dependent upon aggregate data, leaving planners and evaluators with a shortage of sub-national data for evaluative and planning purposes.

This paper describes an open-source architecture and proof of concept for collecting, storing, visualizing and analyzing high-fidelity (i.e. better accuracy, precision, and resolution) performance indicator data in a geographic information system. The architecture uses an existing multi-donor dataset from Malawi, Africa, combined with standardized performance indicator pseudo data that meet the structure of the U.S. Foreign Assistance Framework. While the data structure can be ported over to proprietary or commercial systems, open source tools are explored because their licensing costs are attractive to non-profit organizations and meet the challenges of basic data entry processes involving a perfect storm of low bandwidth, low capacity and low budget. The architecture and prototype is also scalable- it works for an organization with few activities or a large organization such as USAID, which must coordinate tens of thousands of activities all over the world. Activity performance pseudo data are then compared to traditional reporting methods.

***Keywords:*** *monitoring and evaluation, performance, USAID, data management, open source, quality control, project management*

Purpose and goals

The goal of this project is to provide an architecture that uses location, an integral piece of information needed for project performance monitoring and management decisions. In order to achieve this goal, the following requirements were developed for the scope of this project:

1. Keep it financially accessible: while some implementing partners have started using commercial software to manage their project GISes, many implementing partners are unable to invest in licenses for servers, mobile devices, and desktop solutions. This architecture’s point of entry from a hardware perspective is a computer with 4GB of RAM. Every piece of software used for this project is license-free and open source, and downloadable from the internet.

2. Keep implementation accessible: the aid community does not have a large budget for systems development. This program, as completed, can be used by someone without programming experience, but to create the system from scratch requires some minimal programming capabilities. Mature open source technologies were chosen for this project, so there is a large user and developer community to draw support from in case implementing partners experience challenges during setup.

3. Provide a common foundation for data management: to successfully address the goal of collecting and analyzing sub-national performance data, an agency-wide, efficient data structure is required. This paper used the Foreign Assistance Framework (FAF) F-indicators spreadsheet published by the State Department and analyzed each indicator to identify the data entry points, the geographic precision expected, and whether the indicator is categorized as an output or outcome under the FAF.

4. Test feasibility and usefulness: the prototype was tested to ensure a complete data collection to data analysis walk-through could be completed successfully. A summary of this walkthrough shows the functionality the system provides to end users, including project managers, monitoring officers, and strategic planners. Secondly, a storyline was developed to explain a performance issue within a geographic locale. Pseudo data was created to align with the storyline, and a comparative analysis between current indicator reporting requirements and proposed indicator reporting requirements was performed to test whether the storyline performance depression was identified with better precision and in a more timely fashion.

Background

The United States Agency for International Development (USAID) has missions in 92 countries around the globe (USAID n.d.). Each mission operates numerous projects at a given time, utilizing contractors and grantees to implement its programming. Due to a lack of standardized data collection procedures, uniform high-fidelity data are hard to collect and analyze, and as a result, quantitative historical context is lost.

Lack of standardization at a minimum geographic threshold prevents analysts from finding correlations to other, non-USAID collected data. Consequently, the question of whether a project or Mission strategy had a lasting positive impact is largely subjective in nature. This paper proposes that if precise lat/lon data is collected in lieu of administrative boundary aggregates, development practitioners can analyze this data for clusters, patterns, and correlations to other data sets, allowing program planners to discern best practices and lessons learned for better program planning at a sub-national level that is not yet available world-wide or across multiple sectors.

History and Structure of Foreign Assistance (F-) Indicators

The State Department released the U.S. Foreign Assistance Framework (FAF) in 2006 and required USAID to follow the framework during its planning sessions (U.S. Department of State, 2006). Within this framework, development hypotheses are formed under standard global categories based on the both the type of goals the country has, and where along the development continuum the country is located. Within each global category, there are several standard development indicators that will track and measure the outputs and outcomes of the contract. F-indicators are chosen from a standardized list (United States Department of State 2012) in order to help align projects with the hypothesis of the FAF. Through tracking these indicators, implementing partners and government officials hope to substantiate the development hypothesis.

F-indicators include inputs (person days of employment), outputs (kilometers of road repaired), (number of people trained) and outcomes (percentage reduction in childhood disease). Programs are developed to leverage previous activities, and cumulatively aggregate these successes into outcomes, intermediate results, and ultimately, achieved development objectives (USAID n.d.). An example of leveraging is first creating discrete infrastructure activities to dig wells. The planned outputs from this activity are the number of wells dug and the number of people (beneficiaries) who had increased access to potable water. Hand-washing campaigns follow the well construction and use the potable water sources as an input into handwashing demonstrations. The planned outcome from this campaign is a reduction in the occurrence of childhood diarrheal disease.

History of geospatial M&E data usage

Much of the existing Monitoring and Evaluation literature is about outcome or impact indicators such as infant mortality, child literacy, environmental degradation, GDP and rates of disease (Merry 2011) (Campagne 2006) (World Bank n.d.). These types of “outcome” indicators do not identify the contributions of a particular project intervention. More so, the idea of “location” is missing from much of the literature. Using location to track project output data and higher-level outcome data allows for correlative analysis to either add credence to or disprove the development hypothesis under scrutiny. The basis of this paper’s methodology is to collect output indicator data at lat/lon locations so this analysis can take place at a meaningful sub-national level.

Using GIS for project management is not a new concept (Campagne 2006, MEASURE Evaluation 2012, Ott and Swiaczny 2001, Khan, Akhter and Ahmad 2011). More development practitioners are making the case for better M&E processes (Custer 2012) and using geospatial tools to manage their projects. This prototype makes the case to do both. And indicator usage is expanding in the aid community as a function of the ‘decision-making’ shift that accompanies a power shift from political/bureaucratic processes to the democratization of the data. However, as reliance upon data during strategy planning becomes more mainstream, the false perception that data are always truthful must be considered (Merry 2011). This is important in the context of this prototype, as its approach to M&E data collection and analysis is not a cure-all for the ills of missing or inaccurate data, but they are a piece of the puzzle that will help provide a better context for decision-making.

An example of using GIS for decision-making in the aid/development context is combining economic indicators with the physical landscape to select grantees for activities. In Malawi, USAID combined slope, aspect of the sun, sun exposure, roads, and land cover data to identify beekeeping operations that would produce high-quality honey in large quantities for international export through the COMPASS II program (USAID/COMPASS 2007). This paper proposes leveraging performance data so it can be used in single-objective analyses such as this – by identifying high performing areas that can rise to the challenge of more ambitious projects.

USAID has sponsored several program-focused (eg: health) indicator databases (USAID Demographic and Health Surveys 2014). It has also collected sub-national indicator data in high-profile countries such as Afghanistan (USAID n.d.), although it is still tied to administrative boundaries such as provinces or districts. The agency does compel data collection at the country administrative boundaries from all of its implementing partners. But it has not collected M&E data across all program areas with lat/lon coordinates. Below are several examples of approaches to monitoring and evaluation data that USAID has undertaken:

USAID’s GeoBase, or its successor, Afghan Info, was designed to coordinate efforts across agencies within Afghanistan and is an example of a cross-program database in a high-profile country. Its success in collecting subnational results data was limited in that it was collected at an administrative boundary level which mitigated, but did not correct, the modified area unit problem[[1]](#endnote-1), and according to the GAO, data continued to be decentralized as organizations continued to use their own project tracking tools (Government Accountability Office 2011).

The MEASURE DHS Database (USAID Demographic and Health Surveys 2014) is an example of a program-specific database designed to improve transparency and coordination between USAID and other donors. It does collect some standard and many customized outcome indicator data, some of which are high resolution and available to researchers in statistical software format, but it does not provide worldwide coverage at project site lat/lon precision. Catholic Relief Services has developed, using ESRI products, a mobile data collection system for performance indicators, including customized indicators. This system is proprietary to CRS, and not available to other implementing partners (Bothwell 2014).

AidData.org is funded in part by USAID and tracks project activity and dollars at high, mostly country-level, aggregates. They have some subnational datasets available to the public. To create these datasets, Aiddata.org georeferences project locations reported in quarterly and final project reports, newspaper articles, and other sources. Their data fidelity runs from coarse administrative unit centroids to lat/lon points. Attributes for these project locations are not M&E related, but instead track money flows, what projects are going where, and general project titles (AidData 2014).

World-wide, USAID uses F-indicators at a high administrative boundary level, typically in per-country, aggregates. The agency requires reporting from implementing partners, and with few exceptions, has not required sub-national reporting. As a result, USAID cannot examine its Agency portfolio to discern patterns in a sub-country context, as the act of aggregation modulates hot spots (of success, of failure, of mitigating circumstances) within a country’s boundaries. It is clear there is a gap in M&E data management that is (1) global in nature, (2) provides high-fidelity data, and (3) can be used across multiple development sectors. This architecture addresses each of these needs.

Technical overview

The architecture developed uses a stack of open source software, starting with a PostgreSQL database that serves as the project data repository to which all other technologies connect (Figure 1 and Table 1).

Tools used for the prototype

PostgreSQL. PostgreSQL is a stable, enterprise class relational database. While it is serves as the foundation for the Boundlesss OpenGeo package, it does have some limitation in edit capabilities, which is discussed in the lessons learned section.

OpenGeo (Boundless n.d.): Boundless’ OpenGeo architecture provides Geoserver and the PostGIS extension for PostgreSQL in an out-of-the-box package pre-configured to work together. Geoserver is used to serve the data to users who are not advanced geospatial analysts, but would like to see and perform limited analysis/queries on the data in a map.

Open Data Kit (ODK): Open Data Kit is used for mobile data collection and provides the option for data submission over wifi instead of transmitting via cell tower. It communicates with PostgreSQL databases. Because it requires the TomCat web server and servlet, while Boundless’ package requires a Jetty installation, this architecture utilizes two ports on the same computer to allow both installations to run concurrently.

OpenOffice Base: OpenOffice Base is used for desktop data entry and data review and approval processes. This software is the most ubiquitous form of a file database in the open source environment. It has a reliable reputation for serving as a front-end database for database servers such as PostgreSQL. OpenOffice Base uses a Java Database Connectivity (JDBC) driver to connect to PostgreSQL.

Quantum GIS (QGIS): QGIS is a desktop application used for advanced geospatial analysis. Temporal analysis, raster and vector support, and geospatial correlation analysis are available tools and analytical options with QGIS. The “OpenGeo Suite Explorer for QGIS” plug-in allows more complex layer styles to be uploaded to PostGIS and Geoserver directly.

Data preparation and restructuring

The data sources for this prototype are the State Department’s Standard Foreign Assistance Master Indicator List (State Department 2012) and AidData.org’s Malawi Geocoded Activity-level data (Peratsakis 2012). Both datasets required some substantial data manipulation in order to create a normalized, relational database with a single table for indicator results data entry.

The State Department’s Standard Indicators list is an excel spreadsheet of every standard indicator and its associated disaggregation requirements, if any. The challenge arose in that some indicators have disaggregations, while others do not. Some indicators have a natural geographic component where a project activity is taking place, while others are more abstract (ie: Percentage of days per month that selected interdiction vessels are available for operations). Once the data relationships were normalized, 1,128 data entry-ready indicators or disaggregates were identified. 957 of those are indicator disaggregates, while the remaining 171 are indicators that have no disaggregates. Of those 1,128 indicator/disaggregation combinations, 1,029 of them are spatial in nature and have an element of sub-national geo-location feasibility.

More than 25% of the 438 indicators are flagged by the State Department as “outcome” indicators. These indicators cannot be directly associated with a single project activity, and are typically gathered in a survey that assesses change in economic prosperity, disease rates, or other related development topics over time. The data schema portion of this prototype does not cover outcome indicators or indicators that require surveys or geospatial analysis prior to data entry. However, it is entirely possible and highly advised to use a similar geospatial approach to manage outcome indicators in order to perform correlative analysis between output and outcome indicators and other ancillary data. Simple statistical analysis may not conclude any performance patterns regarding specific development hypotheses, but the uneven performance that confounds conclusion during statistical analysis may have clear correlations to other datasets when analyzed geospatially.

The Malawi AidData.org dataset is an excel file with 2,035 records available on AidData.org’s website. It was ingested into a relational database, and normalized until there were groupings of umbrella “Subprojects” in a lookup table, and their associated “Activity” locations contained in a separate table. It should be noted that AidData.org did not collect indicator data as it was not available. As a result, the “results” data that this prototype uses is illustrative only.

Each of the final master data table structure follows table normalization guidelines for relational database design. As illustrated in Figure 2, subprojects are a parent table to activities, which in turn is a parent table to results, which is a parent table to results photographs. Each results record is based on a time stamp in order to enable temporal analysis.

Results

Feasibility Results

The prototype was tested for interoperability among tools. It met the requirements for data collection, data transfer between data tables, and data integrity (i.e.: no duplication or loss of data). To illustrate how a team of aid workers would use this prototype, this paper steps through the process for entering a new subproject and activity locations, identifying indicators and expected results, monitoring that project and collecting results data, and finally, analyzing collective results visually.

Plan and execute a subproject

A project manager plans a subproject in a specific location. (S)he enters information about the subproject into the database and the expected results for this project, by location, via OpenOffice Base. The screenshots in Figure 3 shows three of the forms in the front end database; the first form is used for planning a subproject and activity location.

Monitor project progress, collect results

As the project is implemented, monitoring and evaluation officers will visit the subproject activities. While they are on site, they will collect F-indicator data and photographs via their mobile phones. Figure 4 shows a subset of screen forms from the ODK mobile application. In addition to the F-indicator data collected, the monitoring officer enters their name, the visit date, and collects geographic coordinates for the location.

As the monitoring officer navigates through the screens to fill out the form, (s)he fills out the *aggregate* numbers for each indicator over the lifetime of the activity. There is also an additional field to identify the percentage of completion for this subproject activity that is not part of the FAF, but is an important field used by project managers to identify activities that are struggling before larger interventions are needed.

Review data and approve for transfer to master tables

After completing the forms, the monitoring officer returns to the project office and uploads them to the database. The forms are uploaded to ODK tables. Because ODK tables cannot be altered in any way, the data are automatically transferred to “holding” tables that are designed for review and data transfer to the master tables by the monitoring supervisor in OpenOffice Base. Supervisors can also use Base to view results photos, and manage the attribute data and lat/lon points (Figure 3). Open Office base is also used to export the data tables so they can be sent to USAID on a monthly basis.

Two important functional pieces are needed to make this process work. First, underlying the approval button is a macro that copies and inserts the approved records into the master table, then deletes the corresponding records from the holding tables to save space. Second, the PostgreSQL database contains trigger functions to update the geometry field required to map the data when the lat or lon fields are updated from any front end interface.

Use views for geospatial analysis and reporting

To perform rudimentary geospatial analysis, users can use a web-based map published by Geoserver (Figure 5). The prototype allows for attribute queries, and basic visualization of the data. While Geoserver has more advanced capabilities, including customized processes, they were not explored in this prototype. Users can also use QGIS, as it has capability to utilize data housed in PostgreSQL. QGIS has a variety of plug-ins that allow for temporal analysis, conflation, and other geospatial analyses and processes.

Analytical Usefulness Results

In order to succeed, the proposed architecture must provide users with better information than what is traditionally collected and analyzed. To test the merit of this architecture, results pseudo data were randomly generated and entered into the database- the majority of which followed a pattern of increasing outputs as scheduled, while one area’s results were adjusted based on a simulated scenario of civil unrest in the city of Lilongwe during the 2001-2013 timeframe, affecting project implementation. The important question is: did the geographic representation of the data give better analytical capabilities and conclusions than what is currently being done?

Precision Improvement Test

Traditionally, implementing partners’ reporting to USAID is done per contract, per country, with no geographic disaggregation, but for the purposes of testing the precision improvement, it was assumed that Malawi’s reporting requirements are at district level instead. Analytical challenges arise when, over time, administrative boundaries change, and make comparison from year to year difficult. Table 2 depicts what the USAID mission might submit to headquarters after aggregating all results from their implementing partners in a given year.

If one indicator, “Kilometers of roads constructed or repaired with USG assistance” is used for a map graphic and compared to the same data, disaggregated to the lat/lon precision as pulled from the prototype, the modified area unit problem becomes clear. Lat/lon precision better identifies were gaps in road rehabilitation projects are covered. And while lat/lon data can be aggregated up for more traditional report, aggregate data cannot be disaggregated down to higher fidelity.

This becomes more important as clusters within districts are localized rather than evenly distributed throughout the administrative boundary. Note the orange Mangochi district on the southern tip of Lake Malawi, where in the district map, it appears there is good, dense coverage throughout the administrative boundary. But when the higher precision lat/lon points are added to the middle map, it appears that this coverage is limited to a few areas along the lake. Improving the precision improves decision-makers’ context- it is clear there are pockets within each district that lack road construction projects. In Figure 6, gaps in road construction projects are more easily identified when precise lat/lon data is gathered.

Identify Aid Implementation Issues Test

Sometimes, issues during implementation arise- and these issues may seem like isolated events, but current data collection methodologies do not allow users to determine whether this is the case. The results pseudo data in the Lilongwe area of the country is based on a fictional case of civil unrest that prevented implementing partners from implementing shovel-ready infrastructure grants in the city. When looking at the results data (Figures 6 and 7), the cluster of low-performing activities is prominent, making it clear there was an issue inside the city. This understanding would have been lost if they had been aggregated to the district or country-level boundaries.

Further Research Options and Lessons Learned

During development of this prototype, several challenges arose, including tool limitations, version incompatibility, and bugs that have not yet been addressed. For example, PostgreSQL does not support editing joined table views, and as a result, Open Geo has no edit functionality. Even with these issues, the final prototype is functional. In order to bring the prototype into full production, some additional research and development are required:

* Data review and approval must still have human processes in place to discourage fraud. The system is capable of implementing security best practices, and these should be implemented in a live system. But fraud can only be prevented through checks and balances throughout the process of data collection and management. This architecture supports better data review processes by using more precise geographic coordinates and contributes to better quality control.
* Definitions for acceptable precision must be created and required as a standard. AidData.org has been a leader in identifying this issue (AidData 2014). To use the most precise data available is advised- especially if this data can be collected during monitoring visits during project implementation instead of after a project closes out.
* An additional assessment should be conducted to identify how these tools can be absorbed into an organization’s current project management software, if any exists. Additional fields and tables within the data structure should be considered to make the prototype more appealing for project management use. Creating tables related to project management (such as travel authorizations, budgets, resource management, etc. would create a one-stop shop for managing project activities. It is critical for this system to provide actionable data to project managers in order for it to be used successfully.

Conclusions

High-fidelity activity results data are necessary to ensure quality control throughout the project lifecycle. They are a critical link for evaluating USAID’s development hypotheses in data-driven analyses. This paper proves that the technology is present and affordable and the data can be structured efficiently for data transfer and aggregation. Improvement in the accuracy and precision of results data provides the aid community with another project management resource that provides early warning for challenging activities while building a historical context that allows analysts to evaluate return on aid investments. International assistance is important work, and the data exists- why not leverage the data into something useful?

Figures and Tables

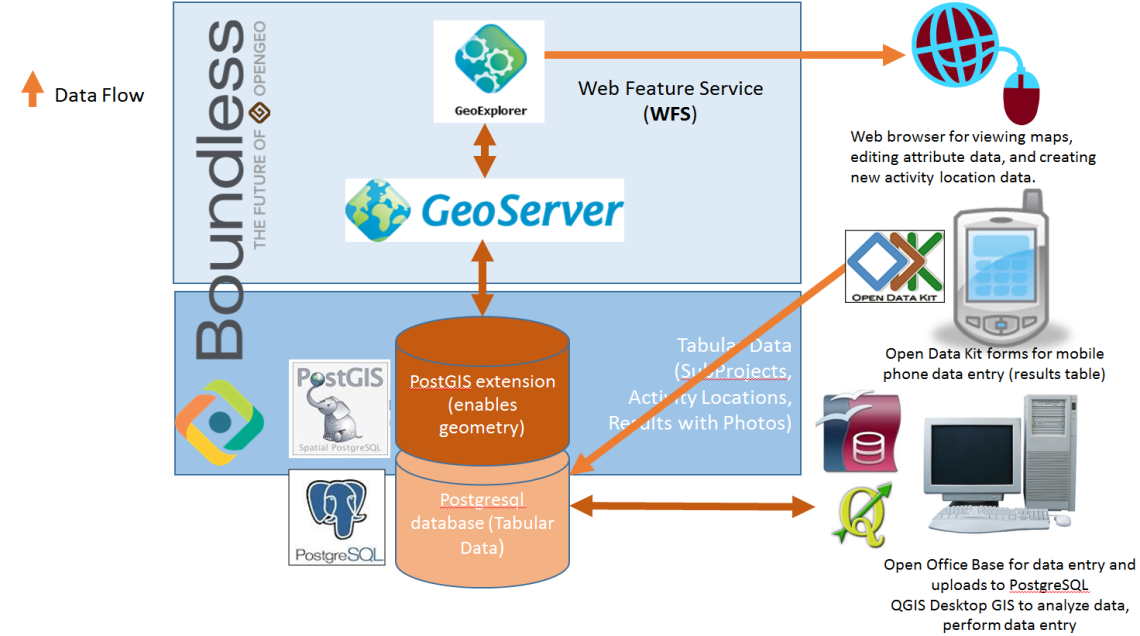


Figure : An overview of the architecture developed to manage performance data from USAID.

Table : The open source software utilized in the architecture

|  |  |  |
| --- | --- | --- |
| **Software** | **Version** | **Purpose** |
| Boundless  - PostGIS  - Geoserver  -Jetty | 4.0.2 | * Geospatial extension for PostgreSQL * Web mapping * Required for Geoserver installation |
| PostgreSQL | 9.3 | * Data repository for subproject, activity, results, and photo data |
| OpenOffice Base | 4.0.1 | * Front-end allows users to access the PostgreSQL database, approve and migrate data entry from monitoring visits. |
| JDBC4 PostgreSQL Driver | 9.3-1102 | * Allows connection between PostgreSQL and Base |
| TomCat | 6.0 | * Required for ODK installation |
| Open Data Kit Aggregate | 1.4.3 | * Mobile data collection functionality - server |
| Open Data Kit Collect | 1.4.3\_rev1042 | * Mobile data collection functionality - android application |
| QGIS | 2.2.0 | * Desktop geospatial analysis |
| QGIS Time Manager | 1.0.2 | * Temporal Analysis |
| QGIS OpenGeo Suite Explorer for QGIS | 4.1 | * Allows integration between QGIS and PostGIS/Geoserver |

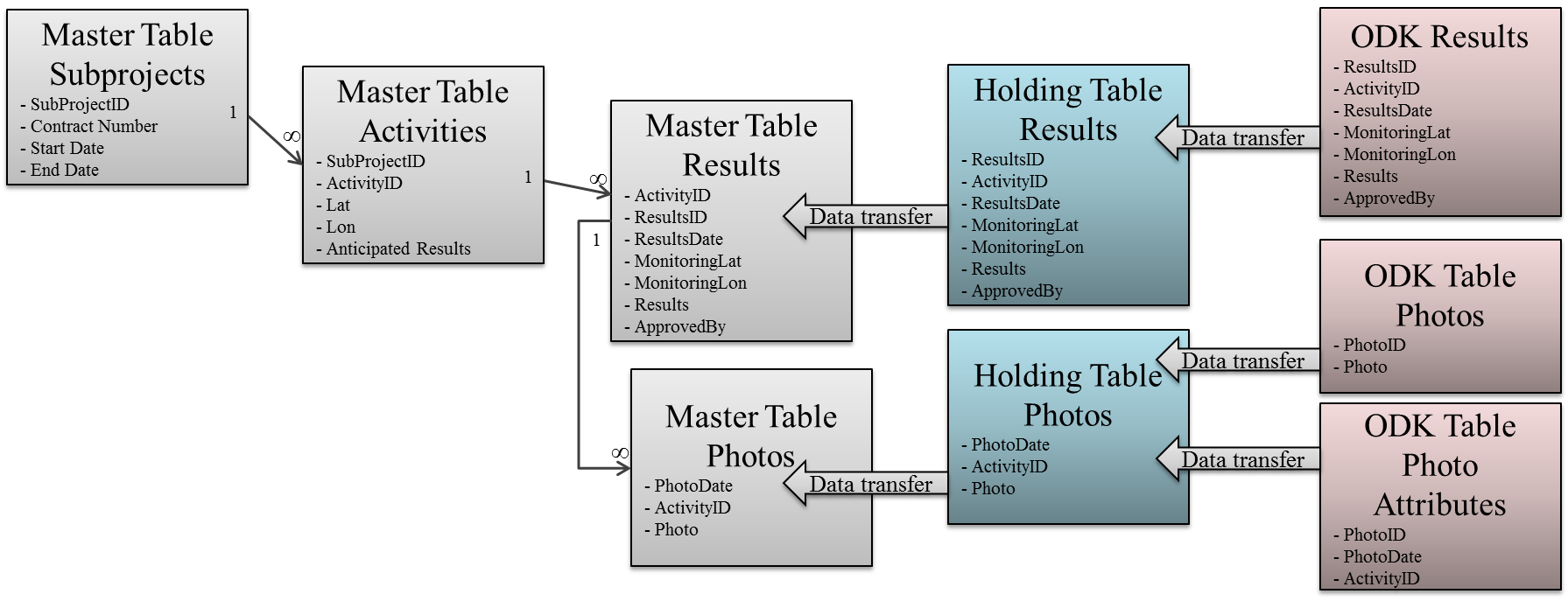


Figure : The entity relationship diagram for the PostgreSQL back end, combined with a data flow diagram showing the data transfer from ODK’s tables to the master table. The data is immediately transferred from the ODK tables to the Holding Tables, but requires a manager to approve transfer from the Holding Tables to the Master tables. Note that in full production, the ODK table data should be archived to protect against data loss.

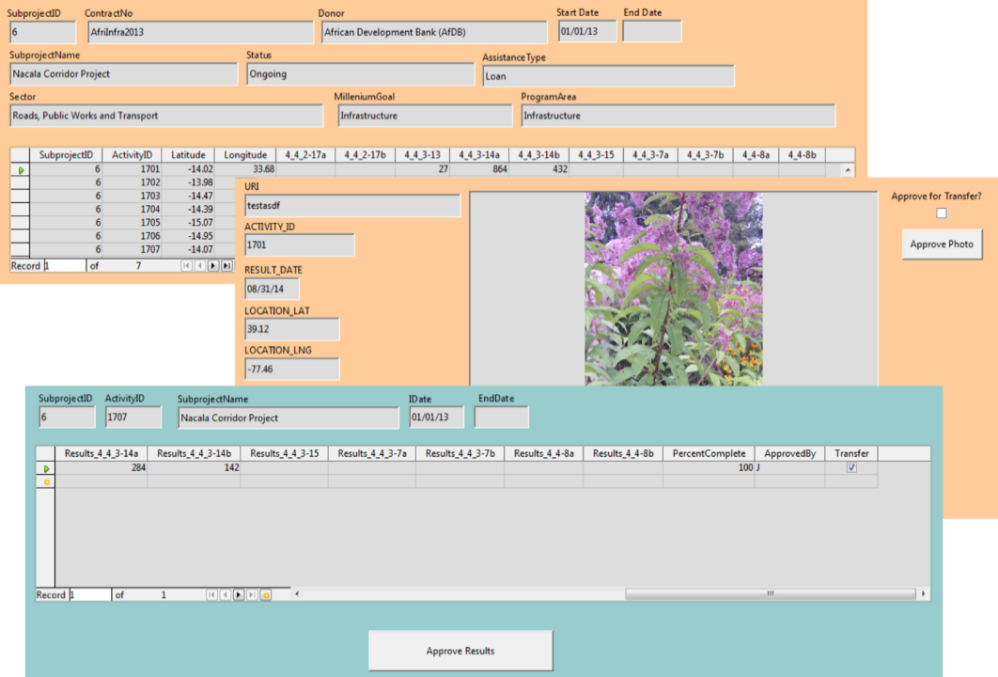


Figure : The OpenOffice Base front end, where managers create subproject data, and approve results data and photos from monitoring visits.

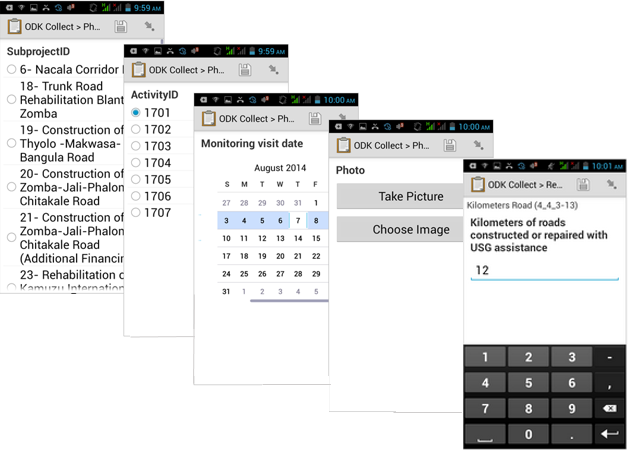


Figure : A subset of screenshots illustrating the Open Data Kit interface, customized for a monitoring and evaluation visit data collection exercise.

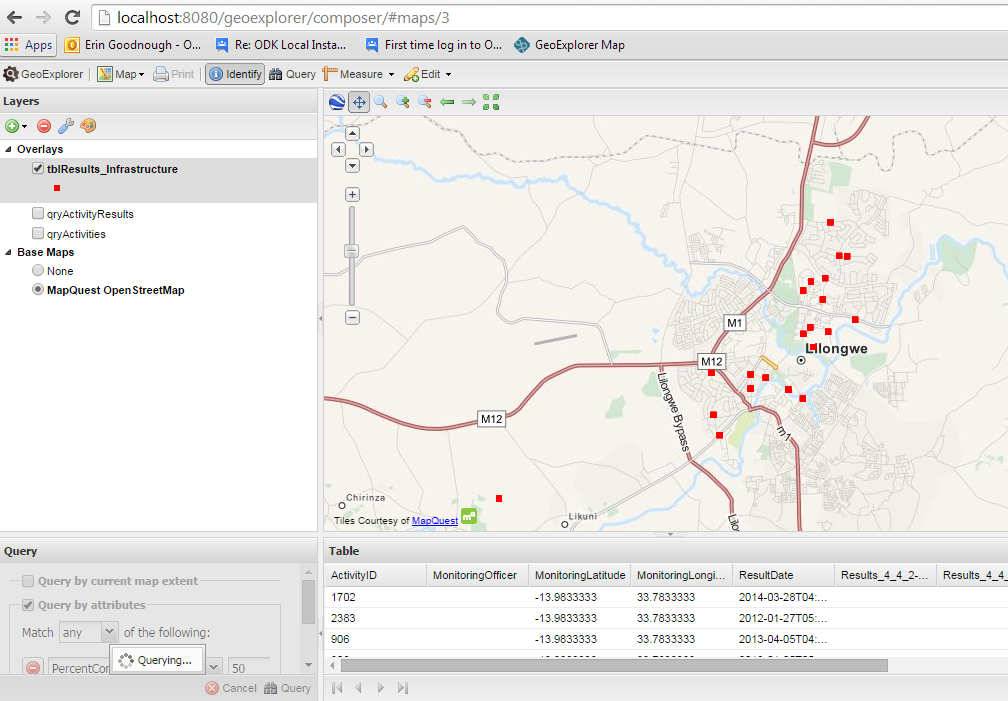


Figure : Simple query functionality in the web map supported by the Boundless package.

Table : The first 9 of 27 districts’ results.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Admin 2 Name | Sum of Results 4\_4\_2-17a | Sum of Results 4\_4\_2-17b | Sum of Results 4\_4\_3-7a | Sum of Results 4\_4\_3-7b | Sum of Results 4\_4\_3-13 | Sum of Results 4\_4\_3-14a | Sum of Results 4\_4\_3-14b | Sum of Results 4\_4\_3-15 | Sum of Results 4\_4-8a | Sum of Results 4\_4-8b |
| Area under National Administration |  |  | 1232 | 1232 |  |  |  |  | 1232 | 1232 |
| Balaka |  |  | 4736 | 4736 | 212 |  |  |  |  |  |
| Blantyre |  |  | 27316 | 27316 | 844 |  |  | 100 | 9256 | 9256 |
| Chikwawa |  |  | 9960 | 9960 | 164 |  |  |  | 6124 | 6124 |
| Chiradzulu |  |  | 2320 | 2320 | 124 |  |  |  |  |  |
| Chitipa |  |  | 1720 | 1720 | 52 |  |  |  |  |  |
| Dedza |  |  | 11580 | 11580 | 248 |  |  |  | 6028 | 6028 |
| Dowa |  |  | 8412 | 8412 | 192 |  |  |  | 5560 | 5560 |
| Karonga | 36 | 36 | 4220 | 4220 | 192 |  |  |  |  |  |

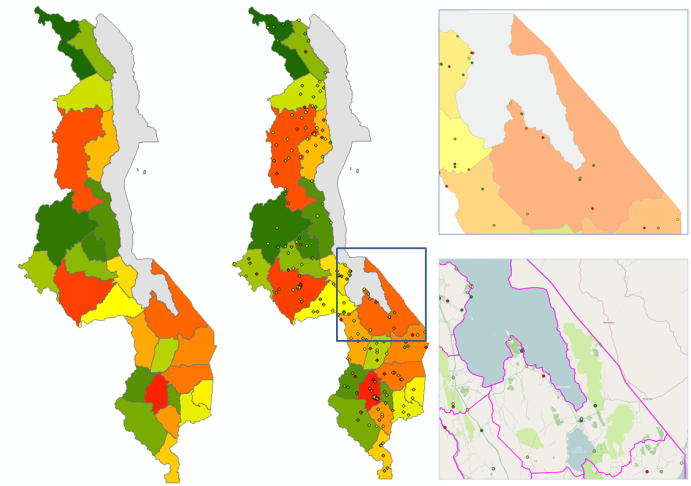
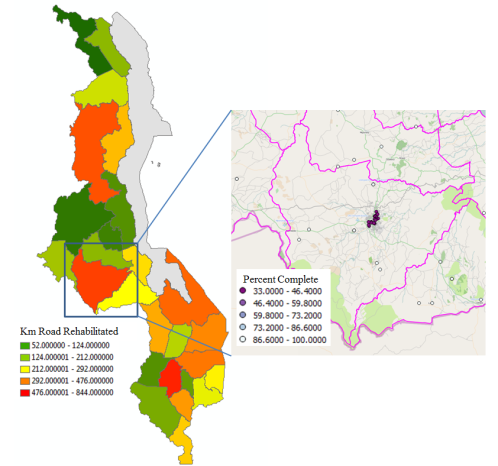
 

Figure : The number of kilometers of roads rehabilitated per administrative level 2 boundary. Dark green colors depict smaller values, while dark red colors depict higher values. The points in the second map depict the data disaggregated to lat/lon precision. While Lilongwe is dark red indicating a higher number of kilometers of road rehabilitated, several road construction projects were not completed due to the simulated “productivity depression” storyline. This context, using traditional methods of reporting, would have been lost, but were picked up using data from the prototype. Administrative data was sourced from the Malawi Portal (MASDAP 2014)

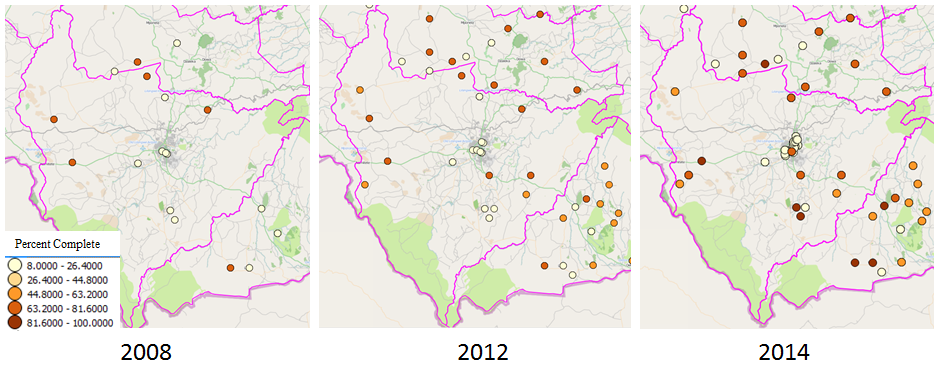


Figure : Using Time Manager to visualize project implementation, users can see that the percent complete field in the Lilongwe city proper is not making progress, but other activities in the district are. Note that this data uses AidData’s lat/lon fields from its activity dataset, but also uses results pseudo data fields.

ACKNOWLEDGEMENTS

This project would not have been possible without guidance and encouragement from my Capstone advisor, Dr. John Kelmelis, data from AidData.org and technical advice from many of my current colleagues at MDA Information Systems LLC. Former colleagues who have inspired and/or supported this project in its infancy include Heather McHugh and Bob Bouvier.

Works Cited

AidData. "Aid Data User Guide." 2014. https://docs.google.com/file/d/0B1I9MfL0ynr\_VUlaRTM0ckM4b2M/edit (accessed May 2014).

—. "Open Data for International Development." 2014. http://aiddata.org/ (accessed April 2014).

Bothwell, Carol. "Global Aid, Development, and Conservation Immersion Summit: Catholic Relief Service's GIS for Monitoring and Evaluation." ESRI Federal User Conference, 2014.

Boundless. *OpenGeo Suite Architecture.* n.d. http://boundlessgeo.com/whitepaper/opengeo-architecture/ (accessed April 2014).

Campagne, Michele, ed. *GIS for Sustainable Development.* Taylor and Francis. Boca Raton, FL, 2006.

Custer, Samantha. "Do Strong Monitoring and Evaluation Systems and High Levels of Staff Supervision Make World Bank Projects More Effective?" *AidData 3.0.* January 12, 2012. http://aiddata.org/blog/do-strong-monitoring-and-evaluation-systems-and-high-levels-of-staff-supervision-make-world-bank (accessed March 12, 2014).

Government Accountability Office. "Lack of information sharing could create the potential for duplication of efforts between U.S. agencies involved in development efforts in Afghanistan." *Government Accountability Office.* Februrary 01, 2011. http://www.gao.gov/products/GAO-11-318SP#mt=e-report&f=/ereport/GAO-11-318SP/data\_center/International\_affairs/Lack\_of\_information\_sharing\_could\_create\_the\_potential\_for\_duplication\_of\_efforts\_between\_U.S.\_agencies\_involved\_in\_development\_efforts\_in\_Afgha (accessed April 2014).

Khan, Khalid, Amin Akhter, and Zulfiqar Ahmad. "GIS as Project Manager in Geophysical Applications Software." *Scientific Research.* April 2011. http://www.scirp.org/journal/PaperInformation.aspx?PaperID=4664#.VAY2m\_ldVqs (accessed June 2014).

MASDAP. "GAUL Malawi Administrative Boundaries." *Malawi Spatial Data Portal.* 2014. http://www.masdap.mw/layers/geonode:gaul\_malawi\_poly (accessed September 2014).

MEASURE Evaluation. "Geographic Tools and M&E: Linking data to support M&E." *Slideshare.* September 2012. http://www.slideshare.net/measureevaluation/geographic-tools-and-me-linking-data-to-support-me (accessed July 2014).

Merry, Sally Engle. "Measuring the World: Indicators, Human Rights, and Global Governance." Edited by Marina Welker, and Rebecca Hardin Damani J. Partridge. *Current Anthropology* ( The University of Chicago Press on behalf of Wenner-Gren Foundation for Anthropological Research) 52, no. S3 (2011): S83-S95.

Ott, Thomas, and Frank Swiaczny. *Time-Integrative Geographic Information Systems: Management and Analysis of Spatio-Temporal Data, Volume 1.* Berlin: Springer-Verlag, 2001.

Peratsakis, Christian, Joshua Powell, Michael Findley, Justin Baker and Catherine Weaver. "Geocoded Activity-Level Data from the Government of Malawi's Aid Management Platform." Washington, DC: AidData and the Robert S. Strauss Center for International Security and Law, 2012.

State Department. "Standard Foreign Assistance Indicators." 2012.

United States Department of State. "Standard Foreign Assistance Indicators." *U.S. Department of State.* 2012. http://www.state.gov/f/indicators/index.htm (accessed April 2014).

USAID. "Afghanistan Performance Monitoring Plan." n.d. http://www.usaid.gov/afghanistan/performance-monitoring-plan (accessed July 2014).

—. "Country Development Cooperation Strategy Guidance PDF." *USAID.* n.d. http://www.usaid.gov/sites/default/files/documents/1870/CDCS\_Guidance\_V3.pdf (accessed September 2014).

USAID Demographic and Health Surveys. *The DHS Program - Quality Information to Plan, Monitor and Improve Population, Health, and Nutrition Programs.* 2014. http://www.dhsprogram.com/ (accessed 08 07, 2014).

USAID. *Where We Work.* n.d. map.usaid.gov (accessed August 2014).

USAID/COMPASS. "2nd Quarter 2007 Report : Community Partnerships for Sustainable Resource Management (COMPASS II)." March 2007. http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&cad=rja&uact=8&ved=0CEgQFjAF&url=http%3A%2F%2Fwww.rmportal.net%2Flibrary%2Fcontent%2Fframe%2Fcompass-ii-quarterly-progress-report2014-2nd-quarter-2007%2Fat\_download%2Ffile&ei=cIAHVJruG9imggSM7 (accessed September 2014).

World Bank. "World Development Indicators." n.d. http://data.worldbank.org/data-catalog/world-development-indicators (accessed September 02, 2014).

1. This statement is based on my experience as an aid worker using the database. Activity locations were collected at village level lat/lon coordinates, but indicator data was collected at the district or province administrative level. [↑](#endnote-ref-1)