OSGEO UN COMMITTEE EDUCATIONAL CHALLENGE: A USE CASE OF SHARING SOFTWARE AND EXPERIENCE FROM ALL OVER THE WORLD

S. Franceschi 1, K. Adoch 2, H.K. Kang 3, C. Hupy 4, S. Coetzee 5, M. Brovelli 6
1 HydroloGIS srl, Bolzano, Italy - silvia.franceschi@hydrologis.eu
2 Mara Arwot GeoSpatial, Kampala, Uganda - ketty@mara-geo.com
3 Korea Research Institute for Human Settlements, Republic of Korea - hhkang@krihs.re.kr
4 Hupy UAS LLC, Indiana, USA - christs.hupy@gmail.com
5 Centre for Geoinformation Science, Department of Geography, Geoinformatics and Meteorology, University of Pretoria, South Africa - serena.coetzee@up.ac.za
6 Department of Civil and Environmental Engineering, Politecnico di Milano, Italy - maria.brovelli@polimi.it

Commission IV, WG IV/4

KEY WORDS: Open Geospatial Data, Open Source Software, Open Geo-Analytic Functions, Processing Toolbox

ABSTRACT:
The paper presents the outcomes of an Educational Challenge launched by the OSGeo (Open Source Geospatial Foundation) United Nations (UN) Committee in 2018. The Committee promotes the development and use of open source software that meets UN needs and supports the aims of the UN. The Challenge supported the UN OpenGIS Initiative, a project "...to identify and develop an Open Source GIS bundle that meets the requirements of UN operations, taking full advantage of the expertise of mission partners including partner nations, technology contributing countries, international organizations, academia, NGOs, private sector". The UN OpenGIS Initiative is organized into working groups, called 'Spirals'. The OSGeo UN Challenge called for the development of training material that can be used for training UN staff working on Spirals 1 and 3. Spiral 1 focuses on a new open source web platform for data collection and Spiral 3 is related to new functionalities needed by UN staff during their field operations. The material developed for the challenges is now openly available for anybody, reaching a wider audience than only UN staff members. This paper describes the challenges and the training material developed for them. Expertise from all over the world was pulled together in designing, mentoring and developing the material.

1. INTRODUCTION

In February 2006, leading teams in the free and open source geospatial world joined efforts to create the Open Source Geospatial Foundation (OSGeo), a not-for-profit organization with the mission to support the collaborative development of open geospatial technologies and data, and to promote their widespread use. On the one hand, the formal birth of such an organization reveals the maturity that different open source "tribes" had reached by realizing that synergizing the development of their geospatial software tools was required to achieve higher levels of interoperability; on the other hand, the connected internet environment called for new requirements and consequently products, and at the same time, geospatial data and applications exited their niche status.

From the beginning, the main benchmarks were interoperability and the choice of Open Source Initiative (OSI) certified licenses allowing different technologies not only to work together, but also to integrate and exchange pieces of code among them. At the same time, OSGeo guarantees the quality of software by following a process during which each software project goes through an incubation process, guided by a specific committee of the organization, before it is released and licensed.

Within the Foundation, Special Interest Groups are encouraged to gather around existing (or, in case, forthcoming) projects. Some projects are related to one open source package (e.g. GRASS, GeoNode, QGIS); others are more transversal and refer more to an application field than to a specific technology. The latter is the case for the OSGeo UN (United Nations) Committee, which aims is to foster the use of open source geospatial software within the UN and projects related to the UN SDGs (Sustainable Development Goals).

In line with its mission, the OSGeo UN Committee supports the UN Open GIS Initiative (Eom, 2017), which was designed during the FOSS4G (Free and Open Source For Geospatial) annual OSGeo Conference in Seoul, Korea in September 2015 and which officially started in February 2016 in Brindisi at the UN Logistic Base for Global Field Support. The goal of the UN Open GIS Initiative is "to identify and develop an Open Source GIS bundle that meets the requirements of UN operations, taking full advantage of the expertise of mission partners (partner nations, technology contributing countries, international organizations, academia, NGOs, private sector)."

The main commitment of OSGeo in the project has been related to Spiral 2 of the project, i.e. capacity building. With the help of volunteers from GeoForAll (the Education Outreach of OSGeo), the suitability for UN staff to use existing training material about the most used open source packages was evaluated. It was decided, in agreement with the UN, to focus the capacity building on QGIS and PostGIS. Several courses were organised, again with the help of volunteers, and by using the education platform of the Politecnico di Milano. In the meantime, the project progressed and new pieces of software were developed by other working groups, specifically, a web platform for data collection as part of Spiral 1 and new functionalities were customized for UN operational field activities as part of Spiral 3. As a result, it was necessary to develop training material for these new pieces of software. The OSGeo UN Committee decided to launch an Educational Challenge for that. In the next section, we provide background about the 2018 Challenge; subsequent sections present the two
challenges and training material developed for them. In the conclusion, some lessons learned are discussed.

2. THE 2018 OSGEO UN CHALLENGE

The challenge was announced on various platforms to make sure that it would reach a global audience. Apart from the two challenges described in this article, a third challenge focused on the current lack of training material for using open source software together with freely available high resolution global geospatial datasets for environmental, social and economic analysis in support of UN SDGs. Providing training material for working with freely available open data and software can make a significant contribution to monitoring SDGs in developing countries where data is often scarce and resources for buying software are limited. The training material developed for the third challenge provides a step-by-step guide on how to calculate the state of SDG indicator 9.1.1, Proportion of the rural population who live within 2km of an all-season road. Anyone in anywhere in the world can follow the steps in the training material to measure progress on SDG 9.1.1 in their area of interest (Ilie et al. 2019).

Applicants had to submit their CVs together with a five-page outline of the proposed geospatial educational material, including learning outcomes, prerequisite knowledge, educational approach, content and a schedule of important milestones for the development of the material. We allowed applicants to submit proposals for more than one challenge. However, one proposal had to be for one challenge only. Only proposals by individuals were accepted (no proposals by teams).

Most of the submissions were from Europe and Africa. Evaluation call is for the development of education material that teaches the user how to apply the GeoSHAPE platform to assist in a real-life scenario involving a relevant theme in today’s world. Themes could include humanitarian aid, disaster management, emergency management, or sustainable development. The scenario could be an ongoing, recent, or historical event or could be created for the purposes of the challenge. If the scenario is created, it should be as realistic and relevant to today’s challenges as possible and use data relevant to a geographic location on the earth. The instructional materials should teach the user how to perform specific workflows relevant to the scenario. Criteria for the content included the suitability of the proposed material for the target audience and requirements, feasibility of the proposal, as well as the gender, relevant experience, qualifications and language proficiency of the proposer. The OSGeo UN Committee Educational Challenge Judging Commission comprising representatives from OSGeo, Boundless and UN Open GIS from five different continents evaluated the proposals based on the following criteria:

- Proposer is an OSGeo Charter Member: 10
- Proposer’s documented experience in education and training: 20
- Proposer’s age (20 if <= 40 years old; 0 if >40 years old): 20
- Experience with respect to what is proposed by the specific challenge: 10
- Proposal (max): 40

With these criteria we gave preference to early career OSGeo charter members with the aim of contributing to their career development. At the same time, we wanted to make sure that suitable educational material of acceptable quality would be developed. The winners of the three challenges were guided by mentors from Boundless, the UN Open GIS Initiative and OSGeo to ensure that the material met the requirements specified in the challenge.

3. OSGEO UN SPIRAL 1: GEOSHAPE PLATFORM FOR DATA COLLECTION

3.1 The challenge

The Spiral 1 challenge called for the development of educational material that teaches the user how to apply the GeoSHAPE platform to assist in a real-life scenario involving a relevant theme in today’s world. GeoSHAPE (Geospatial capabilities for Security, Humanitarian Assistance, Partner Engagement) is a platform designed to enable collaboration on geospatial information between mission partners in connected and disconnected operations. GeoSHAPE was built utilizing open source software and open standards in order to ensure maximum availability to partners and maximum interoperability with other software and data. The core software projects that form the foundation of GeoSHAPE are a geospatial web server (GeoServer), a tool for distributed versioned editing of geospatial data (GeoGig), a platform for developing geospatial information systems and deploying spatial data infrastructures (GeoNode), and basemaps (OpenLayers 3). The interface of GeoSHAPE is an integration of a geospatial portal (GeoNode), a web mapping client (MapLoom), and a mobile application (Arbitre), that leverages the infrastructure provided by the core geospatial server and database components.

The platform is built with the intention of allowing collaboration across distributed and disconnected teams. Data creation and editing can occur simultaneously by teams in the field through either the mobile editing interface (built with Arbitre) or through web-based editing (built with MapLoom). Key features of the platform include the ability to create, edit, and share critical data; provide real-time/map updates anywhere in the world; capture data and photos in the field and upload them to the map; operate in connected and disconnected environments; create notifications about changes in the map. The platform allows the sharing of critical data in an integrated and dynamic platform with the goal of increasing situational awareness in dynamic operations.

Because the GeoSHAPE platform is built with open source components it is available for free with an internet connection. The Rapid Open Geospatial User-Driven Enterprise (ROGUE) Joint Capability Technology Demonstration is a process to create virtual machines to run the platform. GeoSHAPE documentation provides details on the minimum requirements, general concepts, components used and how to set it up. These materials were built with the intention of enabling partners to use the platform without enterprise support. The purpose of Spiral 1 was to provide the next level of instruction to mission partners for application of the platform.

Given the key features of the platform and goal of its use, the challenge for Spiral 1 requested themes that would highlight its capabilities. Themes could include humanitarian aid, disaster management, emergency management, or sustainable development. The scenario could be an ongoing, recent, or historical event, or could be created for the purposes of the challenge. If the scenario was to be created, it should be as realistic and relevant to today’s challenges as possible and use
data relevant to a geographic location on the earth. The instructional materials should teach the user how to perform specific workflows relevant to the scenario. The call requested that applicants cover the following in the proposal:

1. A description of the applied scenario.
2. A list of workflows to be included in the scenario. The list should be specific enough to highlight how the platform can be used in assist in the effort. For example, distributed editing could be used to track and share the location of resources in the field in real-time.
3. A list of the types of relevant datasets to be used with a short description. The datasets should include freely available and open geodatasets. Given the nature of the challenge, datasets may need to be created for the purposes of the scenario.

3.2 The training material

The training material for Spiral 1 (Adoch, 2018) is designed as a step-by-step software learning guide for users of the UN GeoSHAPE platform for data collection. The latter was developed by the geospatial firm Boundless, a company now acquired by the aerospace and data analytics Planet company.

The material provides a guide on how to:

1. Create, edit and share critical data on an integrated dynamic map in near real time.
2. View map updates by users from anywhere in the world.
3. Use GeoSHAPE exchange in connected and disconnected environments. The course is structured with content to suit novice, intermediate and advanced users.

The primary target audience is, therefore, professionals at local, regional, national or international agencies especially those in developing countries. Three modules make up the training material. The first module comprises the core concepts of the UN GeoSHAPE platform for data collection. This module briefly explains the core concepts of the platform and how they relate to each other. Understanding how these concepts relate to each other enables the user to understand, generally, the architecture of the platform and why each of the components is relevant for a robust software system. In addition, understanding these brief explanations will allow the user to realise the full potential of the geographic information system functionality provided by the platform. The following terms relevant to Spiral 1 are explained in the training material:

- Vector data is the type of map data that is represented as points, lines or polygons on a map.
- Raster data is any sort of pixel data or digital image, where each pixel corresponds to a continuous location or area on the earth’s surface.
- Coordinate reference systems (CRS) or spatial reference systems (SRS) are used to describe real-world positional locations on a map.
- A map projection is the systematic transformation (mathematical computation) used to project a point represented as latitude, longitude on the surface of a sphere onto a location on a plane.
- Rendering of tiles refers to the process of drawing vector or raster data to produce a png or jpeg, which can then be sent to a web or mobile client (web-browser, or mobile app) for display.
- Developers build web applications on top of a web stack which is comprised of an operating system (usually in reference to the server’s operating system), programming language(s) for client and server side code, a database, and a server (such as Apache or IIS).
- THE web-client application refers to the main portion of the web application that a user interfaces with through a browser. The client-side portion of the application talks to the server (requests and sends data), performs rendering and any client-side calculations.

The second module describes the components of the GeoSHAPE platform (GeoSHAPE, 2019). The module briefly describes the key building blocks needed for the operation of the platform. UN GeoSHAPE for data collection is the integration of a geospatial portal, GeoNode; a web mapping client, MapLoom; and a mobile application UN AnyWhere, previously called Arbiter (Arbiter, 2019), that are based on a few other components, such as GeoServer, GeoGig and Tileset. The following infrastructure components are described in the training material:

- GeoNode is the main user facing web application and geospatial portal.
- Maploom is the web mapping client which has been embedded in Geonode. Maploom allows you to view, create and edit geospatial features. It also supports connecting to any other server that supports OGC standards: wms, wfs, wcs. GeoServer sits behind GeoNode and serves all the vector and raster data.
- GeoGig is a distributed version control system inspired by Git/Github which has been designed for geospatial vector data. GeoGig is a datastore in Geoserver.
- Geoserver supports numerous datastores one of which is postgres/postgis or GeoGig. During upload, you can choose to upload your layer into a GeoGig datastore instead of PostGIS.
- UN AnyWhere is an Android mobile application (see Figure 1) designed to work with GeoSHAPE. The GeoSHAPE server can work without being connected to the internet and each mobile device can also work without being connected to the GeoSHAPE server.
- Tileset is a package that contains all tiles (images) for an Area of Interest. For example, if you want to use UN AnyWhere in Yumbe, Uganda, your area of interest will be the boundaries of Yumbe. The tileset will need to contain every single image at every zoom level. There are typically 18 zoom levels.

The third module describes the main steps for using GeoSHAPE platform for data collection. Data from various sources and at different scales, collected at various moments in time, are integrated for use within the system, to serve a specific data collection use case. For purposes of writing this paper and also as a way to mirror a real life scenario in the use of Spiral 1, the following sample data was used: water point data and administrative boundaries from OpenStreetMap were downloaded using the Overpass API. The area of interest is Yumbe district in the Northern region of Uganda. Both are open datasets in shapefile format, represented as points and polygons respectively.

Among other steps in using Spiral 1 in the field, the third module takes the user of the platform through the steps for creating a new map, adding layers to Geoshape, viewing the history of a feature/layer, managing conflicts, synchronizing with other repositories and distributed servers, merging two repositories, creating a layer in QGIS, GIS administrator functionality, and GeoSHAPE mobile (UN AnyWhere).

4. OSGEO UN SPIRAL 3: FUNCTIONALITIES FOR UN OPERATIONS

4.1 The challenge

Spiral 3 aims to develop spatial analysis functionality in support of UN Peacekeeping Operations. The main tasks of Spiral 3 can be divided into three categories: 1) to collect requests from UN operations; 2) to develop spatial analysis functions; and 3) to develop training material for the spatial analysis functionality. The development of spatial analysis functions comprises two sub-tasks: firstly, to develop primitive spatial analysis functions; and secondly, to develop an integrated environment in which users can conveniently use the primitive spatial analysis functions like a spatial analysis modeler.

The primitive spatial analysis functions were implemented into a library file (jar) with the Java compiler, in compliance with international standards, such as the Open Geospatial Consortium’s (OGC) Web Processing Standard (WPS) (WPS Standards, 2019). More than 100 primitive functions for spatial analysis were developed with a user manual (Spiral 3 Manual, 2019). The jar library file as an outcome of Spiral 3 that can be downloaded from the Source Forge website (Spatial 3 library, 2019). It can be used in Web, desktop and application development environments. For the Web environment, the Spiral 3 library file is an extension of OGC WPS and can be included as a GeoServer WPS extension.

All of the analytic functions are executable in a desktop environment plugin the Spiral 3 library file into the open source desktop software, uDig, which is compatible with Java language. Users can import the spatial analysis functions in the form of a library itself, similar to other functions in the ‘GeoTools’ library when developing source codes.

For this, dozens of experts participated in Spiral 3 activities. Specifically, UN GIS staff from the UN Mission in South Sudan (UNMISS), the UN Mission in the Democratic Republic of Congo (MINUSCO), the UN Global Service Center (UNGSC) and the UN Headquarters requested spatial analysis functionality for their work. In terms of technical development, experts from MangoSystems Lids and the Korea Research Institute for Human Settlements, as a member of Korean Research & Development Group called OpenGDS (Korean...
Open Source for Geospatial R&D Group - OpenGDS, (www.opengds.re.kr), led the development of spatial analysis functions by collaborating with UN GIS experts and global contributors through the open source development platform (Geotools Spatial Statistics 2019).

In terms of Spiral 3 training materials, OSGeo experts (OSGeo, 2019, OSGeoUN, 2019) supported the development of educational material and the dissemination of the material through the OSGeo online educational network ‘OSGeoLive’. Figure 4 shows the collaborative governance structure of Spiral 3. The environmental analysis based on imagery data (OSGeo Educational Challenge 2) to be presented in this paper was a part of collaboration with OSGeo to develop training materials for Spiral 3.

4.2 The training material

The following requirements were specified for the training material to be prepared for the Spiral 3:

1. The material must be developed for the Processing Toolbox library of uDig developed by Mango System.
2. The processing must be replicable anywhere on the globe, thus the datasets selected must have global coverage and be available for free under an open license.
3. The tutorial should use raster and vector data for environmental analysis related to ecology and ecosystems identification.

The training material (Franceschi, 2018) - see Figure 5 - commences with a brief overview of the uDig Processing Toolbox (see Figure 6), serving to introduce the user to the different algorithms in the geo-analytic library and showing what is possible. It can be adopted as a starting point, a learning guide to approach the software and to make one’s own analysis.

Figure 6. The workspace of the tutorial in uDig: the Processing Toolbox View and the dataset loaded in the Map View.

Geo-analytic functions in the 'Processing Toolbox' library are divided into four categories. First, the General Tools support I/O, visualization, and primitive geometry functions, such as extract, clip, aggregate and dissolve. Second, the Spatial Statistics Tools provide geo-statistical analysis functions such, as ordinary least squares (OLS) and ZonalStats (Figure 7).

Third, the Raster Tools support raster data analysis functions, such as radial line of sight. The training material describes how to use some of the commands for environmental analysis of raster and vector data with the uDig Processing Toolbox. In particular, some specific algorithms are explained and applied to the given dataset. The application shows how to fill in the GUI (Graphical User Interface) and what the result should look like when loaded into the map view.

Natural features of the area around Seoul were used as test dataset. A link in the training material is available for downloading the complete package as a compressed folder. Also included are reference links to the source databases where the user can download the same information for another area of interest.

The tutorial starts with some preliminary operations in uDig to download, install and set up the application and the data. Since the information can be collected from different sources, before starting with the analysis the different kinds of data need to be harmonized to a common projection for the defined study area. The result of the activities of this section is a homogeneous dataset for the area of interest where the data are in the same projection (characteristic needed in some of the processing tools) and clip over a common, limited area in order to limit the uninteresting information and speed up the analysis as much as possible. Information used in the process of ecosystems
localization, and in general for environmental sciences, is very often are derived from raster data. The tutorial shows some possibilities for extracting variables and indexes from two different kinds of raster data, namely satellite images (Landsat8) and digital terrain models (DTM).

The Normalized Difference Vegetation Index (NDVI) is a simple indicator that can be used to analyze remote sensing measurements and assess whether there is live green vegetation on the surface of the area. The tutorial shows how to calculate the NDVI from Landsat8 images, by reclassifying the resulting map to standardized values and colours, and finally, by calculating statistics on the raw and derived data. The Processing Toolbox contains other algorithms to work with satellite images and in particular the possibility to merge raster bands in order to extract indexes or to enhance the visualization of a particular aspects contained in the maps.

The DTM data are more suitable for environmental analysis. The tutorial shows how to approach general operations, such as zonal statistics or spatial relationships like the OLS (ordinary least squares) on this kind of data, together with a first basic approach to the morphological analysis. Surface analysis covers the elaboration of the main morphological variable, like for example, the slope, the curvature and the roughness, together with some morphological indexes. The user is guided to calculate a specific index, the Terrain Ruggedness Index (Figure 8), which is of relevance for the localization of ecosystems. The final step of the analysis of raster data is to aggregate this information and associate it to points representing a specific feature on the terrain or the barycenter of a homogeneous or the whole observed area. This last operation represents one possibility to join raster and vector based information.

The last section of the training material is about vector analysis. Some vector operations are performed on single and multiple vector layers. Analysis on vector data can be performed on geometries or on the attributes and, in general, are not as memory intensive as raster analysis. The first operations shown are on the geometries of single vector layers. Two examples of point pattern analysis are reported, the Nearest Neighbor Index and the Quadrant Analysis, both performed on the distribution of the natural sites near Seoul. A density algorithms, kernel density, was applied to the same layer. The second part of the vector processing contains operations between different layers, an example is the Nearest Neighbor Distance which is part of the proximity analysis. Finally the Processing Toolbox contains a set of tools which can be used to create a statistics report on the results obtained for the different analysis, create charts like a histogram or a scatter plot with a statistical summary of the content of the maps.

4.3 The Future Plan of Spiral 3

As of now, most of the tasks planned on the first phase in March 2016 were completed except user-friendly interfaces for Web environment integrated with other spiral outcomes. Developing a primitive spatial analysis function are completed. Part of training materials for Spiral 3 has been developed in collaboration with OSGeo, for example, like the environmental analysis based on imagery data (OSGeo Educational Challenge 2) presented in this paper. New training materials will be developed in 2019. Spiral 3 will also provide training programs in conjunction with Spiral 2 activity and FOSS4G 2019.

5. CONCLUSIONS

In this paper, the training material developed for the two UN related challenges of the 2018 OSGeo UN Educational Challenge launched by the OSGeo UN Committee was described. The Challenge called for the development of training material that can be used for training UN staff working on a new open source web platform for data collection and on new functionalities needed by UN staff during their field operations. The Challenge successfully contributed to the OSGeo UN Committee’s aim of promoting the development and use of open source software that meets UN needs and supports the aims of the UN. However, the material is freely available to anyone anywhere in the world, and can be used by a much wider audience. Experts from all over the world collaborated in the design, mentoring and development of the material. The documentation has been written by experts already with full knowledge of the base software (GeoShape and uDig) and of the science behind the application. This requirement seems to be one of the most important aspects to approach in a productive way a new set of tools like the ones contained in the tutorials. The second important aspect considered for facilitating the use and the spread of the material and the tools is the development of a quickstart rather than a complete manual. Using this, the user starts understanding how the tools are designed, where to find documentation and what some menus and options means in an automatic way. Reading a full manual from beginning to end in order to understand the bigger picture is impossible for a typical user to do. Because of the risk of wasting time without success, they often leave the software for something less performant but better known or supported. Furthermore, the success of a well known open source tool mostly depends on the collaboration between developers and users in order to find bugs, to add useful tools and to develop a user friendly GUI even for scientific or very specialized applications. Such collaboration is nowadays desirable and
possible using standard technologies. The experts involved in this project worked on different continents and in different time zones, all communication happening online, yet, successfully produced training material without a single face-to-face meeting. Open source software and open data with global coverage were used, so that anyone can repeat the tutorial for their own area of interest in their particular part of the world.

Finally the challenge proposed by the OSGeo UN Committee can be evaluated as successful from the OSGeo point of view:
• Excellent training material is available for UN staff in field operations (and everyone interested in the usage of these pieces of software), and to further improve the material, UN staff will participate as mentors in the next challenge. Being a partner in the UN Open GIS Initiative, this means that OSGeo was able to positively contribute to such a project.
• Cooperation between the training material developers and the mentors was good and fruitful, strengthening the networking between members of the wide open source community and opening possibilities for future collaborations.

**ACKNOWLEDGEMENTS**

The authors wish to thank the OSGeo UN Committee for supporting the initiative. Boundless (Spiral 1) and OSGeo (Spiral 3) sponsored prizes for the challenge. Parts of the Spiral 3 research were supported by a grant (14NSIP-B080144-01) from National Land Space Information Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.

**REFERENCES**


