

GeoNode Importer Progress Report

Ariel Núñez – August, 2013

Background

The Open Data for Resilience Initiative (OpenDRI) has facilitated the setup of geospatial data repositories (GeoNodes) in more than 20 countries.

Several worldwide data repositories exist with free access to population, infrastructure and hazard data but accessing them requires one or many of the following elements:

- a) High bandwidth
- b) Understanding of custom datastructures/formats.
- c) Diverse software tools with difficult to install dependencies.
- d) Diverse cropping mechanisms (10 degrees grids, per continent, per city)
- e) Make dumps for popular countries but give only the option to download the planet wide dataset for non popular countries. (Most of the ones we work with are in the non popular category).

The current body of work address the needs for GeoNode deployments to access and use data relevant to one country without downloading more data than needed or having to do additional post-processing.

Scope

During FY14, we are setting up the infrastructure and working on two initial datasets:

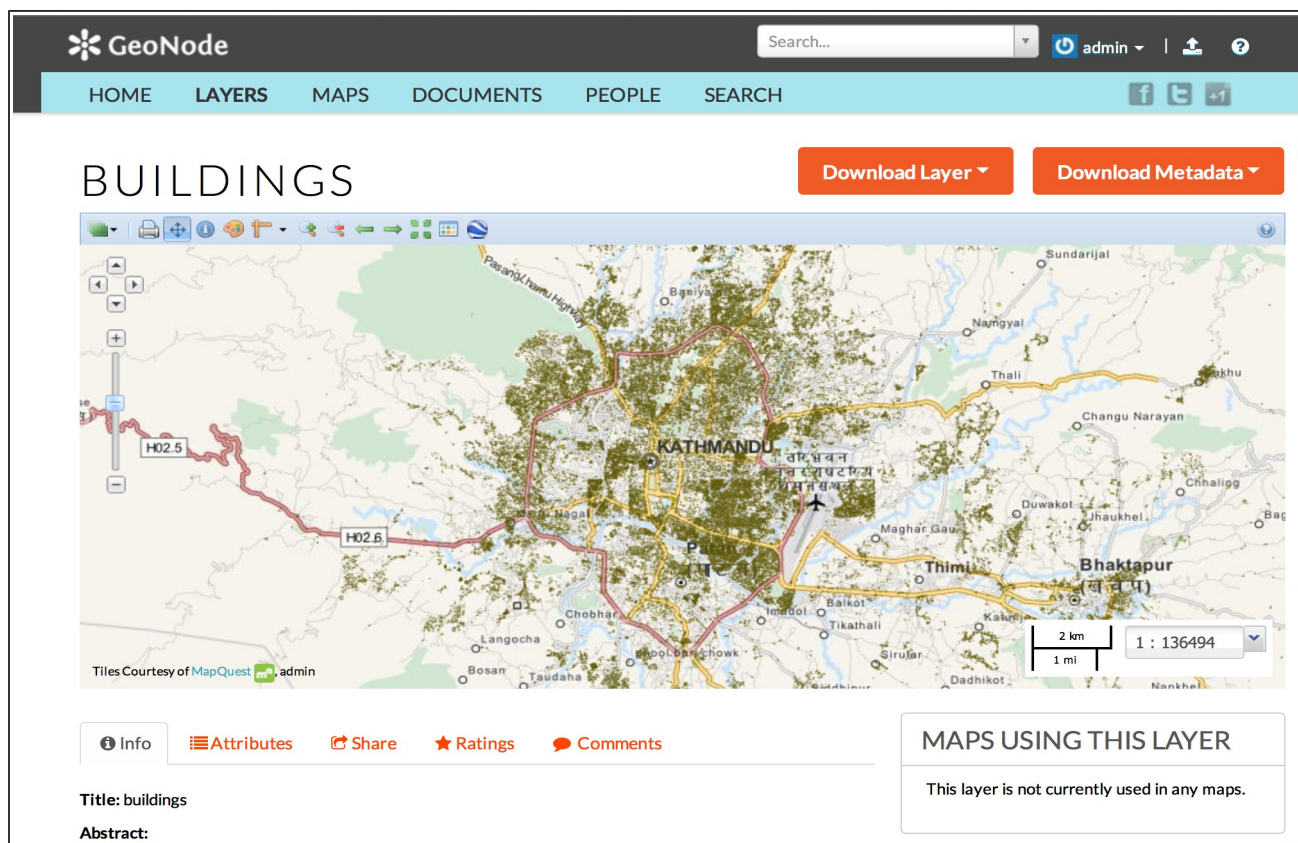
- a) Buildings from the Open Street Map dataset collected by many worldwide initiatives, including OpenDRI in Nepal, Sri Lanka and Indonesia.
- b) Flood maps from the NASA Goddard Laboratory, obtained in partnership with the Dartmouth Flood Observatory.

In order to minimize the operations on the GeoNode servers and the bandwidth use, the architecture of the importer involves a server that preprocesses and splits the datasets by countries. For the OSM datasets, the Humanitarian OpenStreetMap Team accepted to process and host the required artifacts as part of their exports website. (<http://export.hotosm.org>). Similar conversations are underway with the NASA Goddard Team for the Flood datasets.

Progress

The focus to date has been the creation of a proof of concept with the OSM Buildings dataset. The current prototype takes 36 hours to process the world wide OpenStreetMap (~200GB uncompressed) dataset performing the following operations:

- 1.) Filter the ways that are tagged with 'building=*' using a tool called Osmosis. (2 hours)
- 2.) Create dumps for each of 177 countries. (30 hours)
- 3.) Delete the geometries that have less than 3 vertices (buildings are supposed to have area). (0.5 hours).
- 4.) Convert the resulting files to Postgis SQL so it can be loaded in GeoNode(3 hours). Adding columns for specific keys like Building Roof, Wall type, Levels, and others.
- 5.) Add companion SLD and XML metadata files (0.05 hours).



Screenshot of the buildings dataset for Nepal loaded into a test GeoNode.

Next Steps

- Coordinate with the HOT systems administrator to deploy the current codebase.
- Release a geonode-importer (name to be decided) third party extension for GeoNode 2.0.
- Get feedback from potential users, in particular Amrit Karmacharya in Nepal, Srimal Samansiri in Sri Lanka and Maning Sambale in the Philippines.

Other Links

Original Technical Server Concept note: <https://gist.github.com/ingenieroariel/5485939>

Current code: <https://gist.github.com/ingenieroariel/6056187>

Contact

Ariel Núñez

Mail: ingenieroariel@gmail.com

Phone: +57 301 633 9948

Skype: ingenieroariel

Timezone: GMT -5