Automatic assignment of plots from the Spanish National Forest Inventory to municipalities using Linked Open Data

Crespo, N. [1], Vega-Gorgojo, G. [1,2] and Ruano, I. [1,3]

[1] Sustainable Forest Management Research Institute, iuFOR, Universidad de Valladolid & INIA, Spain.

[2] Department of Signal Theory, Communications and Telematics Engineering, E.T.S. of Telecommunications Engineering, Universidad de Valladolid, Valladolid, Spain.

[3] Department of Plant Production and Forest Resources, E.T.S. of Agriculture Engineering, Universidad de Valladolid, Palencia, Spain.

Abstract

The Spanish National Forest Inventory (SNFI) is one of the key large-scale forestry datasets in Spain. It is commonly employed for the management and research of national forest ecosystems, as well as for their conservation. SNFI data is organized at provincial level and it is gathered using a systematic grid of permanent sample plots. The most recent version of SNFI is the third edition (1997-2007); currently the fourth version is in its final stages of production. In our previous work, we created the Cross-Forest dataset, an integrated repository of forestry sources from Spain and Portugal, including the Third SNFI, which is available as Linked Open Data (LOD) at https://crossforest.gsic.uva.es/sparql. LOD is a set of best practices for publishing structured open data on the web in a standardized manner, fostering interconnectivity among data sources. LOD utilizes RDF and SPARQL, formats and protocols developed by the W3C, to provide seamless data integration. The main purpose of this work is to enhance the Cross-Forest dataset with the Spanish municipalities, since this information is especially relevant for forestry studies and management that rely on municipalities as territorial division. Furthermore, we aim to automatically associate SNFI plots to their respective municipalities, since only 42% of the plots in the third SNFI have

this information. To accomplish this, we first developed a pipeline to convert the 8217 municipalities (dataset stored in shapefile format provide by The Spanish National Geographic Institute) into RDF format, using the Well-known text (WKT) representation defined by the OGC. Then, in order to obtain the assignment of plots to municipalities, we initially attempted to use the GeoSPARQL implementation of Virtuoso Open Source Edition 7.2.8.3235-pthreads, the triplestore employed for publishing the Cross-Forest dataset. However, due to false positives, we discarded this approach and opened an issue in the Openlink Virtuoso-opensource GitHub (https://github.com/openlink/virtuoso-opensource/issues/1098). Ultimately, we opted for working with QGIS as a solution. Thus, using the Cross-Forest dataset, we were able to easily export the SNFI plots with their coordinates in CSV format and convert it into shapefile format in QGIS. We then carried out a spatial join between the municipality and plot layers, obtaining the desired assignment. Visual inspection allowed us to assess the accuracy of the plot-municipality allocation. Moreover, we compared our results with the municipality assignments in the Third SNFI and found 86% of coincidence, 9% of plots we cannot locate due to the missing coordinates registered in this inventory edition and 5% assigned to the wrong municipality at the SNFI. Finally, we have uploaded the obtained assignment and municipality data to the Cross-Forest dataset, enriching the information already available. Besides, by adding the municipality information to the forest inventory plots, we made it useful for a range of purposes and applications, such as Forest Explorer (https://forestexplorer.gsic.uva.es/) that allows the visualization of the Cross-Forest dataset through an interactive map. In our future work, we plan to update Forest Explorer to show the municipality data available.

Keywords

Spanish National Forest Inventory, plots, municipalities, QGIS, Linked Open Data, GeoSPARQL