Interlinking and Visualizing Linked Open Data with Geospatial Reference Data

Abdelfettah Feliachi[1], Nathalie Abadie[1], Fayçal Hamdi[2], and Ghislain Auguste Atemezing[3]

[1] IGN, COGIT, 73 Avenue de Paris, 94165 Saint-Mandé, France
[2] CEDRIC, CNAM, F-75141 Paris Cedex 03, France
[3] EURECOM, Multimedia Department, Campus SophiaTech, France

1 Context and purposes

An increasing number of thematic datasets are published as RDF graphs and linked to other datasets by identifying equivalent resources in other relevant datasets. Among the set of properties usually used as data linking criteria, geolocation (addresses, locations, coordinates) remains one of the most commonly used.

However, resources that actually refer to complex topographic features are generally described by very simple geolocation properties, such as a position defined by coordinates (long, lat). On the other hand, geographic reference datasets provide more precise geometric information about geographic features. Interlinking thematic linked open datasets with geographic reference datasets would enable us to take advantage of both information sources to link independent thematic datasets and create rich cartographic applications for data visualization.

This data linking task is generally performed by comparing properties values of each resource of a given data set, with homologous properties of the resources described in other datasets [3]. In the field of geographic databases, data matching is also performed by comparing properties, and especially complex geometries (curves, lines, polygons) that are used to represent the shape and the location of geographic features. This task is usually based on distance measures chosen according to the type of the geometric primitives that must be compared [1, 2, 4, 5]. We aim at combining both approaches to link both thematic and geographical reference data and exploit the generated links in a data visualization application.

2 Approach and use case

In order to take advantage of existing data linking tools, we have converted geographic shape data and stored them into a RDF triple store. This task has been achieved by using the Datalift[4] platform that also enables to perform the linking process with external published datasets, through the use of Silk [5] linking tool. Our linking approach is mainly based on geolocation properties comparison.

[4] `http://datalift.org/`
[5] `https://www.assembla.com/spaces/silk/wiki/Silk\_Workbench`

Thus we have added to Silk more GIS distance measures for computing the shortest distance between any geometric primitive and simple position locations used in thematic datasets.

The result of this interlinking process is a list of `owl:sameAs` links between entities of each datasets, at a given threshold. These links are used to extend the geographic reference data set with information queried on the fly from the external thematic datasets through the visualization interface. We have applied this approach on a geographical reference dataset about buildings and data about historical monuments extracted from French DBpedia[6], on the area of Paris.

Fig. 1. DBpedia points locating historical monuments linked with polygons describing buildings in a geographic reference dataset.

3 Conclusion

The use of links between thematic and reference data could be further investigated to enable data visualization at different level of detail, and visual detection errors during matching process of geodata.

References

1. Mustire, S. et Devogele, T. Matching networks with different levels of detail. GeoInfor-matica, paratre en 2008
2. Olteanu, A.-M. Appariement de donnes spatiales par prise en compte de connaissances imprcises. Thse de doctorat. Universit de Marne-La-Valle, 2008
3. Scharffe, F., Euzenat, J.: Mthodes et outils pour lier le Web des donnes. RFIA 2010: Re-connaissance des Formes et Intelligence Artificielle (2010)
4. Voltz, S. An Iterative Approach for Matching Multiple Representations of Street Data. In : Proceedings of ISPRS Workshop, Multiple representation and interoperability of spatial data, Hanovre (Allemagne), 22-24 fvrier 2006, p. 101-110
5. Walter, V. et Fritsch, D. Matching Spatial datasets: Statistical Approach. International Journal of Geographical Information Science, 1999, 13(5), p. 445-473

[6] `http://fr.dbpedia.org`