Unlocking Geographical Information from Academia: an Open Source WebGIS Solution

B. Neves[(1)], A. M. Rodrigues[(1)], T. Santos[(1)] and S. Freire[(1)]

[(1)] e-GEO Research Centre for Geography and Regional Planning, Faculdade de Ciências Sociais e Humanas FCSH, Universidade NOVA de Lisboa, Avenida de Berna 26-C, P 1069-061, Lisboa, (brunomaneves, amrodrigues, teresasantos, sfreire @fcsh.unl.pt)

RESUMEN

Los Centros de Investigación de Geografía son por lo general productores de un gran volumen de Información Geográfica (IG), los cuales generan tanto proyectos financiados como iniciativas de investigación individuales. El Centro de Estudos de Geografia e Planeamento Regional (e-GEO) ha estado involucrado en varios proyectos a escala local, regional, nacional e internacional. Recientemente, dos cuestiones fueron objeto de debate. Una de ellas fue el hecho de que la información espacial obtenida a partir del desarrollo de tales proyectos de investigación no ha tenido la visibilidad que se esperaba. En la mayoría de las veces, la IG de estos proyectos no estaba en el formato adecuado para que los investigadores -o incluso el público en general o grupos de interés- pudieran pesquisar fácilmente. La segunda cuestión era sobre cómo hacer que estos resultados pudieran ser accesibles al alcance de todos, en todos los lugares, fácilmente y con los mínimos costes para el Centro, teniendo en cuenta el actual contexto económico portugués y los intereses de e-GEO. Estas dos cuestiones se resuelven con una sola respuesta: la puesta en marcha de un WebGIS en una plataforma Open Source. En este trabajo se ilustra la producción de un instrumento para la difusión de las indicaciones geográficas en el World Wide Web, utilizando únicamente software libre y freeware. Esta herramienta permite a todos los investigadores del Centro publicar su IG, la cual aparece como plenamente accesible a cualquier usuario final. Potencialmente, el hecho de permitir que este tipo de información sea plenamente accesible debería generar un gran impacto, acortando las distancias entre el trabajo realizado por los académicos y el usuario final. Creemos que es una óptima manera para que el público pueda acceder e interpretar la información espacial. En conclusión, esta plataforma debería servir para cerrar la brecha entre productores y usuarios de la información geográfica, permitiendo la interacción entre todas las partes así como la carga de nuevos datos dado un conjunto de normas destinadas a control de calidad.

Palabras clave: *WebSIG, SIG, Información Geográfica, Open Source*

ABSTRACT

Geography Research Centres are usually producers of a large volume of Geographical Information (GI), which results both from funded projects, and individual research initiatives. The e-GEO Research Centre for Geography and Regional Planning has been involved in several projects, at local, regional, national and international levels. Recently, two issues came up for discussion. One was the fact that the spatial information resulting from these projects and its outcomes had not the visibility that was expected. The GI from these projects, most of the time, was not in the proper format for the researchers to look up easily, or even the general public or stakeholders. The second issue was on how to make these outcomes available to everyone, in all locations, in pleasant and clarifying ways with minimum costs for the Centre, considering the current Portuguese economic context and e-GEO's interests. These two issues were solved with one single answer, the set-up of a WebGIS in an Open Source platform. This paper illustrates the production of an instrument for the diffusion of GI over the World Wide Web, using only Open Source and Freeware. This tool allows all the researchers at the Centre to publish their GI, which becomes freely accessible to any end-user. Potentially, making this type of information widely available will have a great impact by shortening distances between the work made by academics and the end-user, and we believe it is a better way for the public to access and interpret spatial information. Ultimately, this platform will serve to bridge the gap between producers and users of GI since these systems may allow the interaction between all parties, allowing the upload of new data given a set of rules aiming at quality control.

Keywords: *WebGIS, GIS, Geographical Information, Open Source, Freeware.*

INTRODUCTION

The e-GEO, Research Centre for Geography and Regional Planning from the Faculdade de Ciências Sociais e Humanas, Universidade NOVA de Lisboa, was funded in 1981, and since its beginning, researchers have been participating in numerous projects which end up with the production of Geographical Information (GI).

At the present time, the Centre has a total of around 90 researchers, between Full Members and Collaborator Members, divided in two Research Groups: Geographical Modeling, Cities and Spatial Planning (GMCSP), and Environmental and Socio-Economic Dynamics: Management for Sustainability (ESDMS), involved in both National and International projects, Ph.D. Theses, and Msc. Dissertations, partnerships and individual research.

All the GI produced along these, nearly, 31 years of research had not and it is still not having the proper recognition, either from an internal perspective (among e-GEO researchers), or for the public in general.

In order to make this type of information available we, as a group of researchers, Members of the Centre, joined in a brainstorming reunion to find a solution on how to gather all the existing information and make it accessible. Research Coordinators of both Groups and the Director have agreed with us that the best way of going forward with this project would be through the development of a WebGIS using *Free and Open Source Software* (FOSS).

In the first stage of this work, information was gathered from five distinct projects and prepared to be included as vector layers in the web-based platform. The remaining of the article is organized as follows: after a short presentation of the software and library codes used, the workflow of the project is discussed. Next, the five case-studies chosen will be presented. After the display of the mashed-up information, the article concludes.

FREE AND OPEN SOURCE SOFTWARE (FOSS)

The main concern was on the use of *Free and Open Source Software* (FOSS) also named as *Free/Libre Open Source Software* (FLOSS) technologies [1] and [2]. Which software should we work? Nevertheless this was a choice that would always have to be taken, either for FOSS or *Proprietary Software*. It was made considering the eventual costs that proprietary software would bring to e-GEO with the implementation of the project and its maintenance.

However, care has to be taken in terms of semantics. The term "Free" in FOSS does not mean that it is free of cost but it is referring to the software freedoms that are addressed to FOSS software [3]. These freedoms (0 to 3) are stated in the Free Software Foundation (FSF) Website [4] and GNU Operating System Website [5] and refer to:

(0) *The freedom to run the program, for any purpose*.

(1) *The freedom to study how the program works, and change it so it does your computing as you wish. Access to the source code is a precondition for this*.

(2) *The freedom to redistribute copies so you can help your neighbor*.

(3) *The freedom to distribute copies of your modified versions to others. By doing this you can give the whole community a chance to benefit from your changes. Access to the source code is a precondition for this*.

Quantum GIS and OpenLayers

In order to accomplish the goals stated above, we are using FOSS to work the GI with Quantum GIS (QGIS), which is an Open Source Geospatial Information System [6], licensed with a GNU General Public License (GNU GPL), allowing its redistribution but only under the same conditions protecting its freedoms. It means that changes or improvements in the original code are allowed, however the license cannot be changed. In other words, the license will still be GNU GPL after all additional tools and changes brought to the software [3] and [7].

QGIS can run in a variety of Operating Systems (OS) allowing a larger number of users according to their OS preferences and needs. We, at the Centre have it running in Linux, Mac OSX and Windows, [6].

Besides its innumerous functionalities concerning Geographical Information Systems (GIS), QGIS has several "*Plugins*" in its standard menu and allows adding new ones from third parties. One of the plugins with major interest is the "*OGR2Layers*", which permits the creation of a HyperText Markup Language (HTML) page with OpenLayers.

The FOSS application OpenLayers is a JavaScript library developed by the Open Source software community that able the displaying of GI in an Internet Browser [8].

Together, these FOSS accomplish the task of making possible to researchers and the public it might interest in viewing the vector GI created or being worked in QGIS on the Internet.

Gathering the GI and make it available using OpenLayers is a job far from finished and yet, needs constant updates; it is consider as an ongoing project at e-GEO.

METHODOLOGY

The adopted methodology to run this project has in consideration the following aspects:

- Internet Geographical Information formats;
- Researchers Hardware, Software, Operating Systems;
- Geographical Information researchers produce and its formats.

All these aspects end up being compiled, standardized and, using QGIS for representing the information with the proper format and layout. This GI is then exported using the previously mentioned QGIS *plugin*, the "*OGR2Layers*", turning possible the addition of GI to the Internet (Figure 1), giving visibility to the work being done by e-GEO researchers through research projects and individual research.

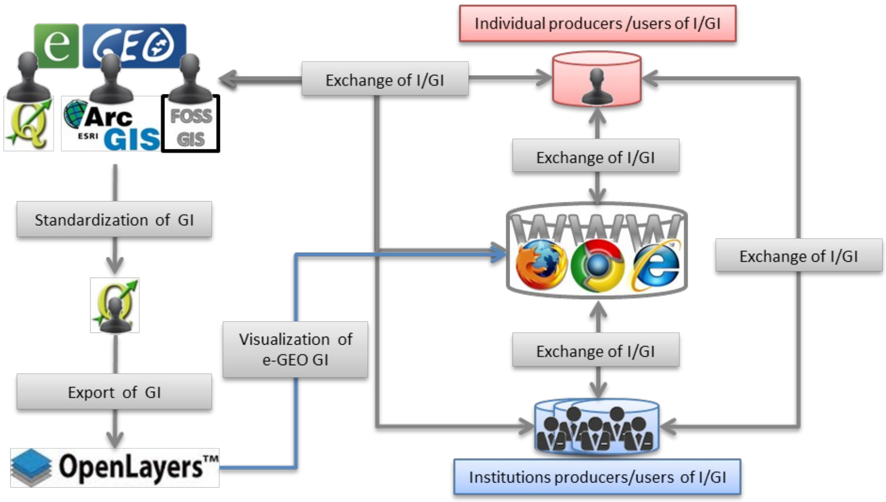

Figure 1: *Interaction of information between e-GEO and external structures*

E-GEO CASE-STUDIES

As mentioned above, part of the present article presents five case-studies, which are representative of the type of GI produced by the Centre. All information shown was later uploaded in the WebGIS platform.

Case-study 1: Measuring peripherality in Mainland Portugal

The importance of the concept of peripherality is a consequence of the fact that humans tend to concentrate their activity in particular locations - urban nodes, and that these nodes are hierarchically organized [9]. In this study a set of indexes were computed which quantify central-peripheral relations within a spatial surface which is characterized by a strong polarization of activity near the coastline and near two main urban nodes. Different metrics were calculated which were latter tested using a Bayesian framework comparing the observed population distribution with estimated posterior distribution parameters.

The spatial object of this paper is the surface of Mainland Portugal. Being peripheral within a territory marked by a strong and generalized polarization of human activity along the coastline means to be located far away from the Atlantic coast. Moreover, first order advantages, together with historical determinism in what initial hierarchies is concerned, caused an agglomeration of administrative functions in a small set of urban nodes. Also, since these small set of locations are not distributed homogeneously along the Atlantic coast, distance in relation to these cities becomes very important.

This project intended to synthetize the concept of spatial peripherality to its primary element: geographical position, and how it can be measured. It is argued that by creating a singular, yet rigorous indicator, an important tool is being provided to the geographer or to any social scientist whose experimental table is a spatial surface/phenomenon [10] and [11]. Its analysis together with other socio-economic indicators is useful when one wants to explain supply and demand along a closed surface.

The indicator to be built is a function of geographical distance and resident population, given the fact that the place where each agent lives is, in the limit, the summary or the epitome of a large set of natural, social and economic conditions. It is shown that, using only these two variables, the resulting spatial patterns are similar to those obtained when using a large set of variables.

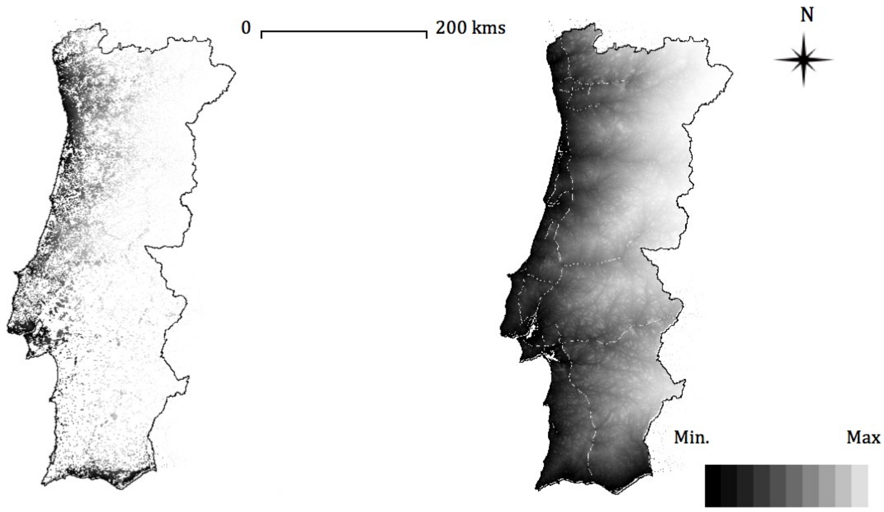

Figure 2: *Accessibility surfaces*

Figure 2 represents two accessibility surfaces computed using GRASS (Geographic Resources Analysis Support System). Distances were calculated in respect to the coastline; the surface represented on the left includes time-distances from every census tract considered as urban, whilst the right figure represents time-distances from every possible 50*50 metres cells which correspond to the whole surface.

Case-study 2: The Need For Updated Geo-Information at the Municipal Level

When characterizing the municipalities' attitude regarding GI, three major questions arise:

- Is GI used in a regular basis?
- Is it an important resource?
- Is the information's spatial and temporal attributes appropriate for the municipal activities?

Based on the answers to the prior questions, considerations on the need and value of updated GI for municipal use can be drawn, as well as assessing the priorities and perceptions of the local government regarding geographical data. With this goal in mind, a survey was designed to characterize the actual level of GI use at the municipal level, and to investigate the needs and how Portuguese municipalities value GI [12]. The survey was planned for providing valid data for developing accurate estimates for the entire population of municipalities. The execution involved several steps. Firstly, the survey text was designed based on the cited objectives. Then it was put online and an invitation e-mail was sent for all the municipalities. After receiving the completed questionnaires, the data was summarized and statistical analyses were performed. From the results of that investigation, several indicators were assessed and are here presented on the form of a WebGIS platform [13].

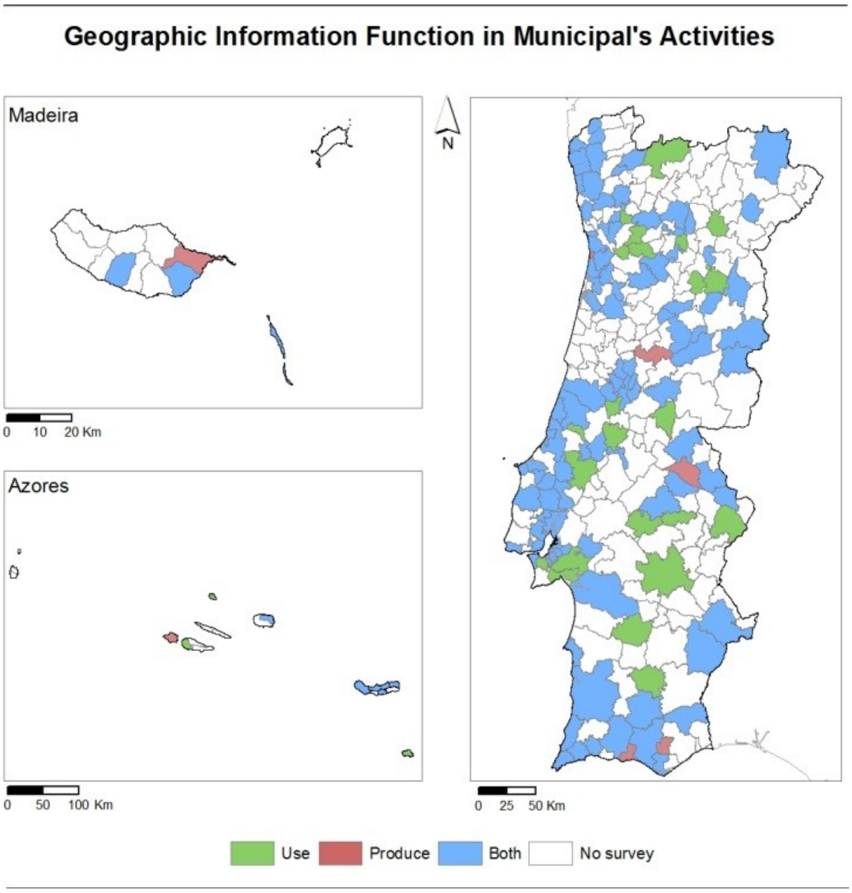

Figure 3: *Relation between geographic information and the municipalities in Portugal*

Case-study 3: Creating Web-based Solar Maps

The application consists of a methodology that applies altimetric data to the evaluation of the potential for incorporating solar power systems into buildings in a city neighborhood. The use of LiDAR data can play an important role in analyzing the suitability of buildings for receiving solar systems. Solar mapping takes advantage of GIS and visualization technologies, and offers a solid knowledge base on solar resources and best practices. Solar maps also offer a comprehensive planning tool to evaluate energy reduction opportunities for new and existing buildings, to plan future energy consumption and supply or to monitor compliance with energy and greenhouse gas goals.

This work, conducted in an area of 625 ha located in the heart of the city of Lisbon, Portugal, analyzed the suitability of rooftop areas for the installation of solar energy systems, and performed a brief technical analysis that considered the optimal location for solar Photovoltaic (PV) systems [14].

Identifying the incoming solar energy at rooftop level entails the modeling of solar radiation incident in each location. Two inputs are required – a Digital Surface Model and the buildings' footprints. With these data, modeling the solar radiation can be done in a GIS environment.

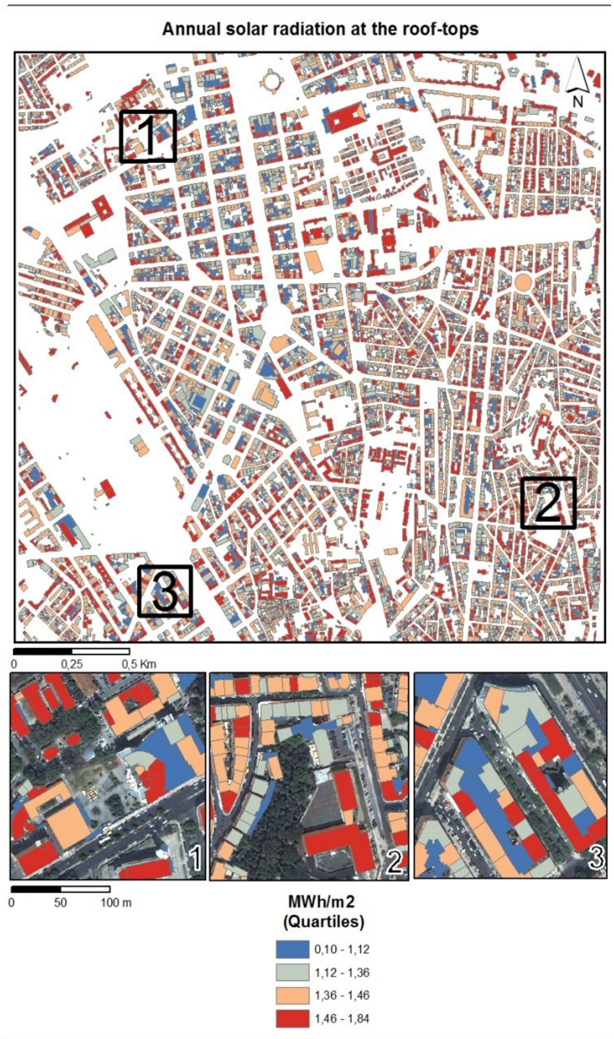

Figure 4: *Mean annual solar radiation available at the roof-tops*

Case-study 4: Recent segregation in the Lisbon area: exploring relations between the social mix and urban dynamics

This studied focused on the pattern of socio-spatial differentiation in the Lisbon area. The main goal was to analyse data distribution patterns with spatial statistics and to interpret its geographic significance in respect to recent trends of urban changes.

Groups of homogenous areas were identified and heterogeneity was quantified using data associated with census database in order to infer whether the city of Lisbon has become more socially segregated.

The starting point was a global analysis of spatial autocorrelation patterns for social and physical indicators. The evolution of these indicators over the considered period provided a general idea of changes in the city's social mix. Later, a detailed analysis of local spatial patterns allowed the identification of significant clusters which constitute evidence of segregation behaviour.

The dataset used comprises the period from 1991 to 2011 and refers to three population census. The study area is the Lisbon city. Social and demographic indicators were used together with indexes of urban dynamics in order to quantify the effect of recent construction and new neighbourhoods on the traditional social mix. Moreover, it's important to study how new developments have affected the social structure of nearby areas. Two geographical levels were be used: local administrative areas - *freguesias*, and census spatial units - *sub-secções*. Whilst the local administrative units provide a stable and long established level of administration, the census tracts provide the greatest level of possible detail when using census data.

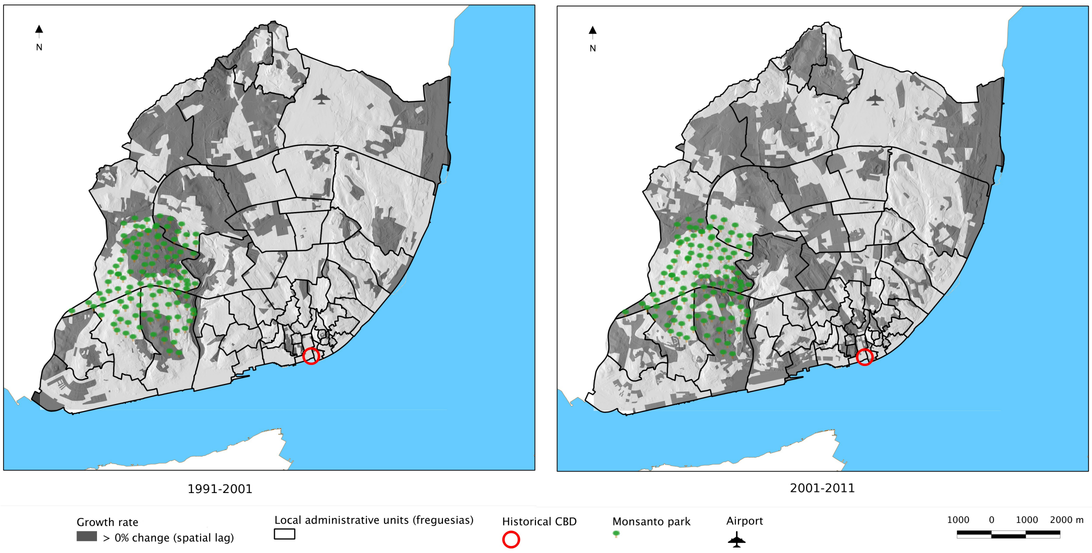

Figure 5: *Lisbon resident population (Growth rate). The values shown represent the spatial lag of the original series, computed using a binary contiguity spatial weights matrix*

Case-study 5: Feature extraction from satellite imagery and LiDAR to update municipal building's map

Urban planning and management are fundamental municipal activities, which should be based on detailed and up-to-date geo-information. In the city of Lisbon, Portugal, the last comprehensive large-scale mapping effort dates from 1998. This work constitutes an approach to automatically updating the city's building map, using LiDAR (Light Detection And Ranging) data and QuickBird satellite imagery [15].

The study area, located to the northeast of the downtown of the city of Lisbon, occupies approximately 24 ha in the former grounds of the 1998 Lisbon World Fair (Expo '98) and has been suffering intense development and land use changes since then.

The methodology developed for the updating of the Municipal 1998 Buildings map is based on a change detection procedure. Using the LiDAR-based normalized Digital Surface Model (nDSM), the 1998 building map is assessed for the detection of changes occurring between that year and 2006. The process is initiated with the segmentation and vectorization of the nDSM grid into features having at least a height of 1 floor (i.e. 3 meters). These 'tall features' are overlaid with the 1998 buildings to evaluate and detect those buildings that disappeared, those that remain, and the 'tall features' that are new in 2006. After retaining only those features wider than 5 meters, those that correspond to trees are dismissed. Identification of trees is performed using the NDVI (Normalized Difference Vegetation Index), computed using the Red and Near-Infrared bands of the 2007 QuickBird multispectral image [16]. The obtained buildings are subject to an 'adjacency test' with the remaining 1998 buildings from the initial assessment, in order to identify the 1998 buildings that are adjacent to new buildings, and the new buildings that are isolated structures. The remaining 1998 buildings which are not adjacent to new ones are deemed "Unchanged Buildings". Those that are adjacent are merged with their contiguous new buildings and are squared up, resulting in "Modified Buildings". The new isolated buildings are squared up and labeled as 'New Buildings' on the Updated Buildings map.

Figure 6 shows the Updated Buildings map for the study area, updated at the building block level.

The initial 1998 Municipal map included 95 individual building structures in 32 building blocks in the study area. The Updated Buildings map includes 34 building blocks, of which 6 are 'Unchanged' from 1998, 11 were 'Modified', and 17 are considered 'New' in 2006.

Compared to producing a building map for 2006 using direct extraction of buildings from the LIDAR data, this approach to updating the 1998 Building's map based on change detection has the advantage of preserving the original footprint of unchanged buildings while assessing the amount of changes. Such a procedure can be used to quickly obtain an updated map for urban analysis and management before another, more costly, topographic database is available.

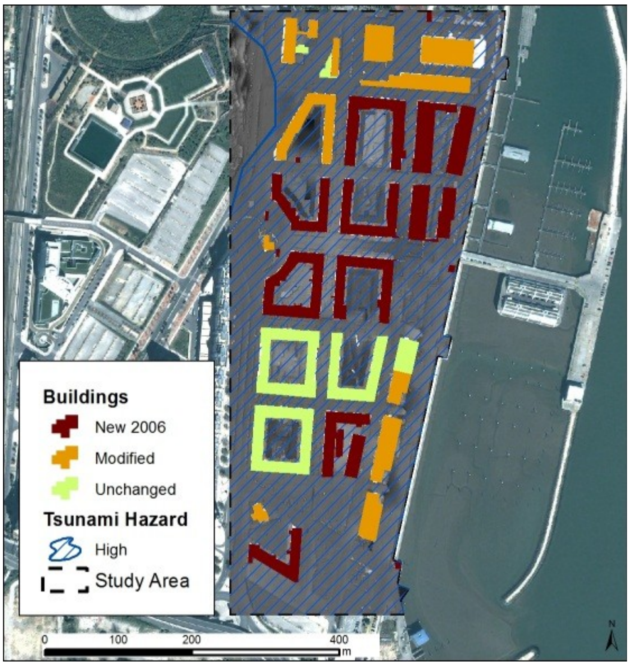

Figure 6: *Updated map of buildings for 2006, overlaid on the nDSM and the 2007 pansharpened QuickBird image.*

CONCLUSIONS

Researchers, when working with spatial information, often produce new GI. These outcomes represent *new geographies* which have the potential to shape human's perception of the landscape being observed.

In e-GEO, it was realized that our work, as scientists, was not over with the publication of results in scientific periodicals. A web-based geographical information system, where new GI may be overlayed with base tiers of information is an important tool to bridge the gap between researchers and the community in general. The greater the degree of interactivity, the easier it will be to create and update geographical data.

It is expected that, in the short-term, the platform will become an important vehicle for the exchange of information, discussion of results and the occurrence of knowledge spillovers between individuals or groups with distinct backgrounds.

E-GEO WEBGIS LAYOUTS

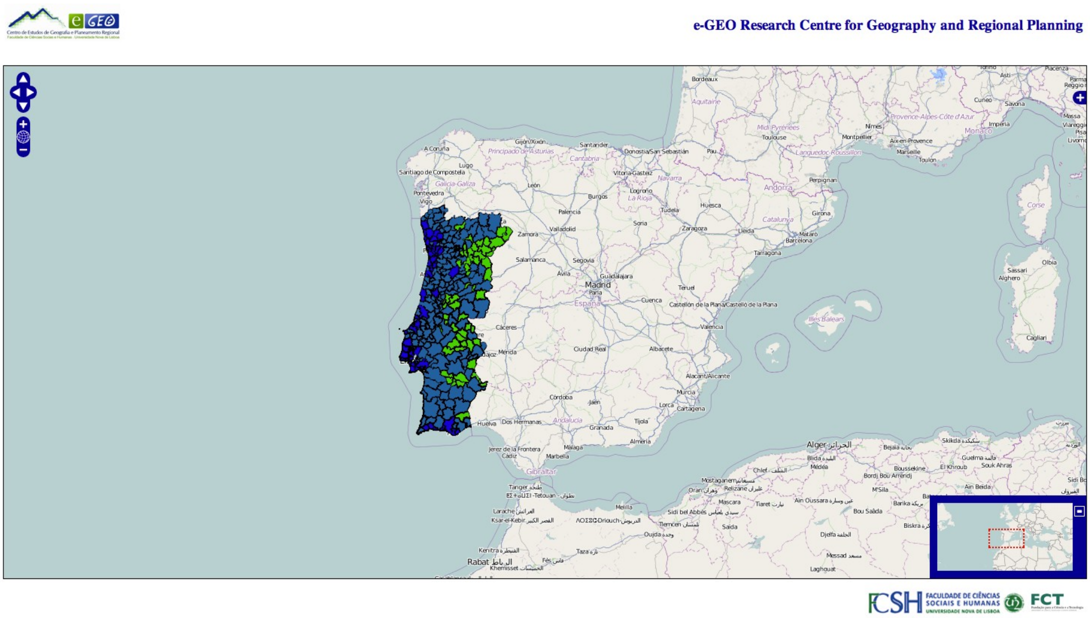

Figure 7: Case study 1

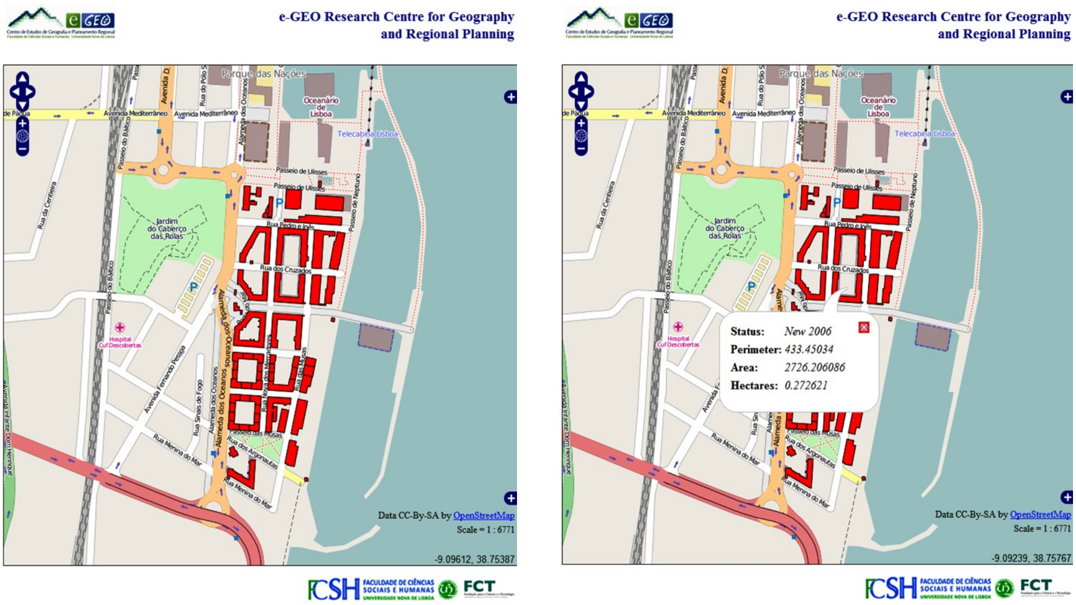

Figure 8: Case study 5

REFERENCES

- [1] DE PAOLI, S.; MISCIONE, G. (2011), "Relationality in geoIT software development: How data structures and organization perform together", *Computers, Environment and Urban Systems*, Volume 35, Issue 2, March 2011, Pages 173-182, ISSN 0198-9715, http://dx.doi.org/10.1016/j.compenvurbsys.2010.09.003
- [2] CROWSTON, K.; LI, Q.; WEI, K.; YELIZ ESERYEL, U.; HOWISON, J. (2007), "Self-organization of teams for free/libre open source software development", *Information and Software Technology*, Volume 49, Issue 6, June 2007, Pages 564-575, ISSN 0950-5849, http://dx.doi.org/10.1016/j.infsof.2007.02.004
- [3] STEINIGER, S.; HAY, G.J.; (2009), "Free and open source geographic information tools for landscape ecology", *Ecological Informatics*, Volume 4, Issue 4,

September 2009, Pages 183-195, ISSN 1574-9541, http://dx.doi.org/10.1016/j.ecoinf.2009.07.004

- [4] FREE SOFTWARE FOUNDATION (n.d.), Free Software Foundation – what we do, In *Free Software Foundation*, Retrieved March 10, 2012, from http://www.fsf.org/
- [5] GNU OPERATING SYSTEM (2012), What is free software?, The Free Software Definition, In *GNU Operating System*, Retrieved March 10, 2012, from http://www.gnu.org/philosophy/free-sw.html
- [6] QUANTUM GIS (n.d.), Welcome to the Quantum GIS Project, In *Quantum GIS*, Retrieved March 10, 2012, from http://www.qgis.org/
- [7] SMITH, B. (2012), A Quick Guide to GPLv3, The Foundations of GPL, In *GNU Operating System*, Retrieved March 10, 2012, from http://www.gnu.org/licenses/quick-guide-gplv3.html
- [8] OPENLAYERS (n.d.), OpenLayers: Free Maps for the Web, In *OpenLayers*, Retrieved March 10, 2012, from http://openlayers.org/
- [9] DIJKSTRA, L.; POELMAN, H. (2008), "Remote Rutal Regions: How Proximity to a City Influences the Performance of Rural Regions". *Technical Report 01*. Directorate-General for Regional Policy.
- [10] GOLLEDGE, R.G.; SIMPSON, R.J. (1997), "Spatial Behavior: A Geographic Perspective". *The Gilford Press*, 1997.
- [11] SCHUMANN, C.; TALAAT, A. (2002) "The European Peipherality Index." *42nd Congress of the European Regional Science Association (ERSA)*, Dortmund.
- [12] SANTOS, T.; TENEDÓRIO, J.A.; FREIRE, S.; NEVES, B.; FONSECA, A. (2011), "Utilização de informação geográfica pelos Municípios Portugueses: resultados de um inquérito nacional. *VII Conferência Nacional de Cartografia e Geodesia*. LIDEL, Lisboa.
- [13] SANTOS, T.; TENEDÓRIO, J.A.; NEVES, B.; RODRIGUES, A.; FREIRE, S. (2011) "Disponibilização de Informação Geográfica em Ambiente Web 2.0: Resultados de um Inquérito à Administração Local", *VIII Congresso da Geografia Portuguesa, Repensar a Geografia Para Novos Desafios: Competências, Investigação, Acção*. Lisboa. ISBN: 978-972-99436-4-5
- [14] SANTOS, T.; FREIRE, S.; FONSECA, A.; TENEDÓRIO, J.A. (2011), "Producing a building change map for urban management purposes". *EARSeL eProceedings*, 10(1): 56-65. ISSN 1729-3782
- [15] FREIRE, S.; SANTOS, T.; TENEDÓRIO, J.A. (2011), "Feature Extraction from Satellite Imagery and LiDAR to Update Exposure to Tsunami and Improve Risk Assessment in Dynamic Urban Areas", in (Ed.): J.M. Tavares, R.M. Natal Jorge, *Computational Vision and Medical Image Processing - VipIMAGE 2011*, Boca Raton, FL: CRC Press, pp. 337-342.
- [16] ROUSE, J.W.; HAAS, R.H.; SCHELL, J.A.; DEERING, D.W. (1973), "Monitoring vegetation systems in the Great Plains with ERTS", *Third ERTS Symposium*, NASA SP-351 I: 309-317.