**INTEGRATION OF GEOSPATIAL DATA WITHIN THE STATISTICAL PRODUCTION PROCESS – GeoINQ**

Mário Lucas, Statistics Portugal, [mario.lucas@ine.pt](mailto:%20mario.lucas@ine.pt)

Ana Santos, Statistics Portugal, [ana.msantos@ine.pt](mailto:ana.msantos@ine.pt)

Francisco Caldeira, Statistics Portugal, [francisco.caldeira@ine.pt](mailto:francisco.caldeira@ine.pt)

Bartholomeus Schoenmakers, Statistics Portugal, [bart.schoen@ine.pt](mailto:bart.schoen@ine.pt)

**Abstract**

*The use of cartography has supported census data collection at Statistics Portugal since 1981.*

*Since 2006, with the production of the enumeration areas dataset to support the 2011 census, Statistics Portugal has been developing a Spatial Data Infrastructure (SDI) that is currently being used, in a transversal way, to promote the integration of the geospatial data in the statistical production process, in order to achieve efficiency and accuracy, within sampling process, data collection or dissemination.*

*In 2011 Statistics Portugal built a national geodatabase comprising all the georeferenced buildings from the 2011 Census (BGE). BGE is a point based coverage that is being continuously edited in a internal quality control process and updated by the Municipalities which provide Statistics Portugal, on a monthly basis, all the completed buildings and buildings permits including X,Y location and addresses. After the 2011 Census, Statistics Portugal evaluated the possibility of implementing a geographic tool that would allow the visualization of buildings and dwellings units of survey samples.*

*For this purpose, GeoINQ was developed to allow the visualization of the location of sample buildings and to provide management and control functionalities of the data collection process. In addition, GeoINQ has editing tools to update the BGE by field interviewers.*

*It is a Geographic Information Systems (GIS) WEB solution developed in order to integrate geospatial data into the production process of official statistics in an innovative way. It allows greater efficiency and rationalization of the resources especially in the household’s surveys by supporting the data collection process.*

*GeoINQ is integrated, via webservices, with Statistics Portugal Global Survey Management System (SIGINQ-IE), and consumes a set of INSPIRE geographic data services.*

**Keywords:** GeoINQ, Geospatial data, Webservices, Surveys, Statistics

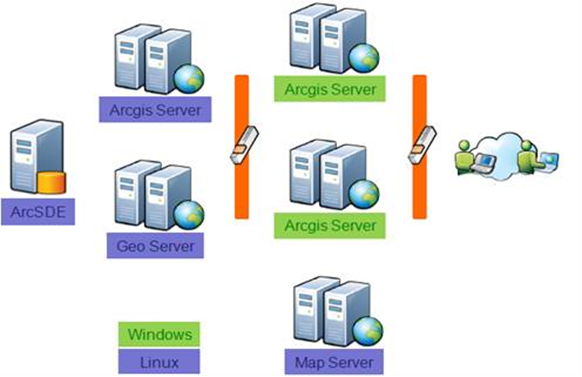
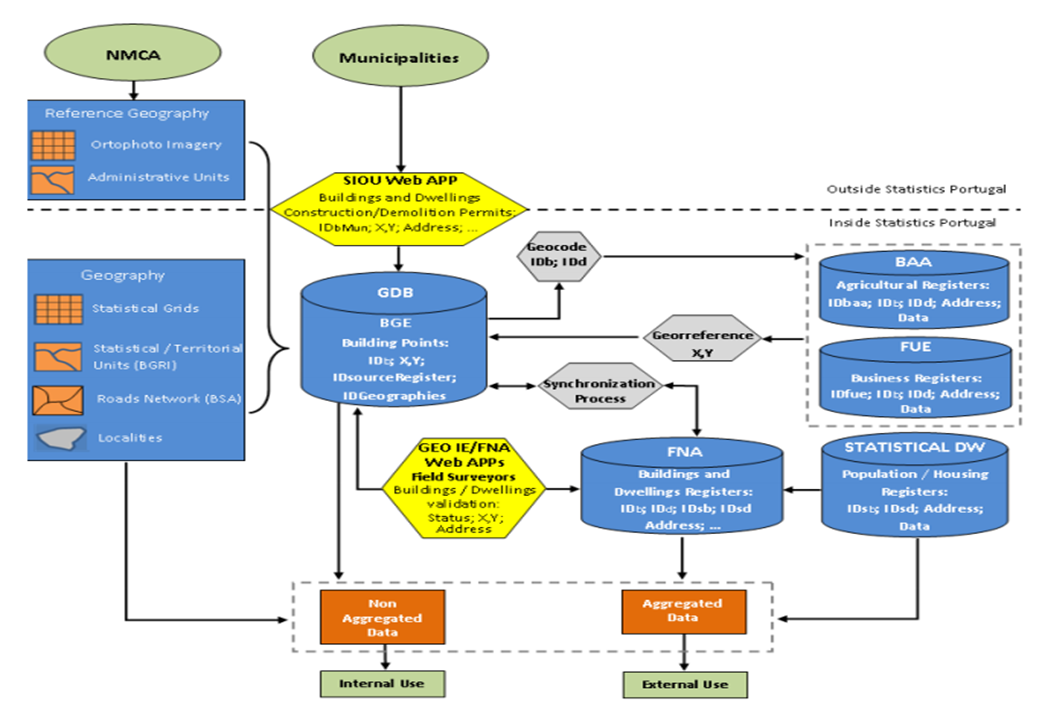
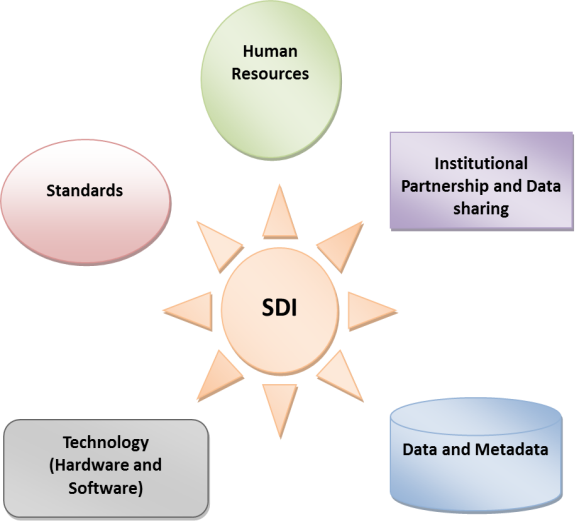
**1. Context**

The use of cartography supports data collection at Statistics Portugal since 1981. In 1995, Statistics Portugal started the preparation of the cartographical infrastructure to support the 2001 Census, named as “Geographic Information Referencing Base” (BGRI 2001) which is based on Geographic Information System (GIS). Since 2006, with the production of the BGRI 2011 for the 2011 Census, Statistics Portugal has been developing the Spatial Data Infrastructure (SDI) to support other statistical activities, in a permanent effort on the creation of geospatial data at a national level.

* 1. *SDI*

Statistics Portugal SDI is the set of resources, standards, technologies, policies, and legal, administrative and organizational frames required for the effective generation, collection, handling, access, distribution, sharing and use of spatial data for statistical production and dissemination.

**Figure 1. Statistics Portugal SDI – Technology and Data**

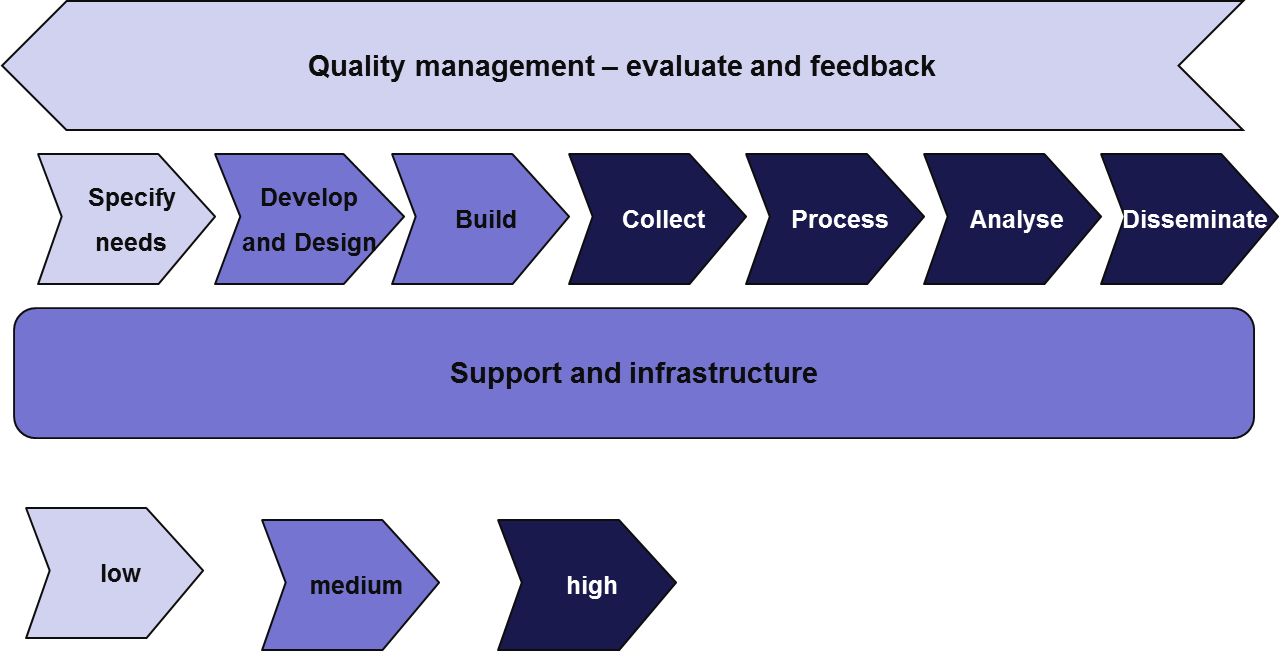


**2. Integration of geospatial data within the statistical production process**

With the SDI development, and the creation (at the 2011censuses) of a point based foundation spatial framework - supported by the BGE buildings - , Statistics Portugal has been increasing the integration of the spatial component at several stages of the statistical production process. This growing integration enhances efficiency and accuracy within several domains:

* Sampling
* Data collection
* Availability of new territorial statistical data
* Dissemination of (geo)statistical information

**Figure 2. Degree of geospatial data and GIS activities within the statistical production process**

****

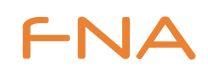
**3. GeoINQ**

GeoINQ is a Geographic Information System (GIS) WEB solution, developed between July and October 2016 in a Statistics Portugal/ESRI Portugal partnership, tailored to Statistics Portugal needs in order to fully integrate SDI spatial data with other existing systems in Statistics Portugal, in particular with the *Integrated Survey Management System – Surveys by Interview* (SIGINQ-IE).

*3.1 GeoINQ – integrating SDI Spatial Data at SIGINQ-IE*

SIGINQ-IE is Statistics Portugal surveys management system that integrates a set of subsystems (Meta-information system; Contacts Center - SICC; National Buildings and dwellings register file – FNA; Surveys by interview process management register system – GPIE-REG; Interviewers management system – ENTR; Sampling management system – SIGUA-UA)

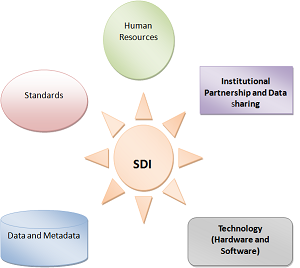
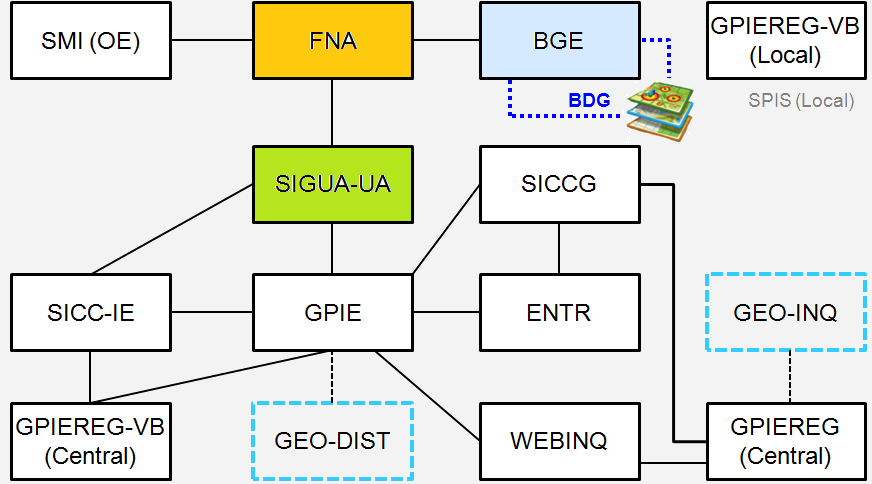
**Figure 3. SIGINQ-IE – Surveys by Interview**



Surveys by Interview

In this general framework, GeoINQ is also a SIGINQ-IE subsystem that allows the integration of the SDI Spatial Data within the Survey Management System.

**Figure 4. GeoINQ – SDI Spatial Data integration at SIGINQ-IE**

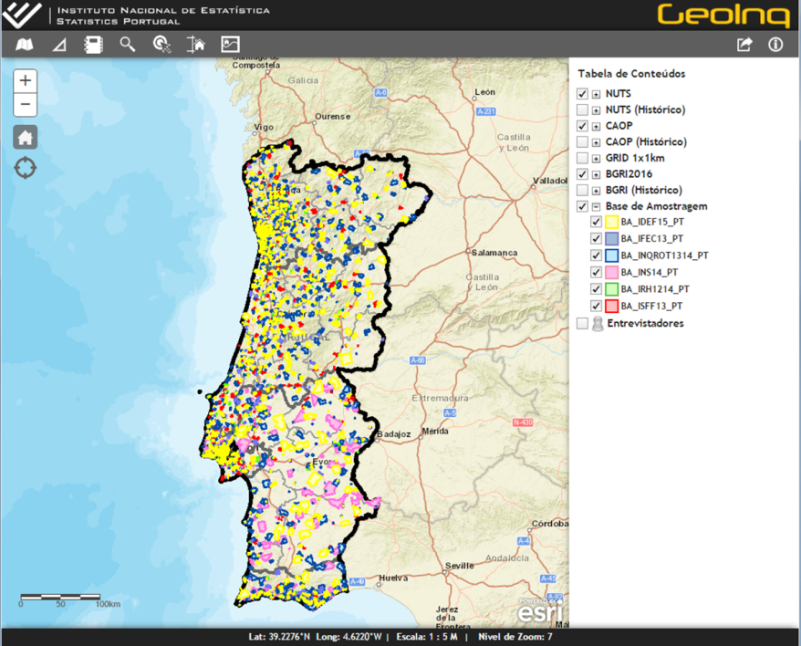


*3.2 GeoINQ - Requirements*

GeoINQ was developed under a set of basic requirements:

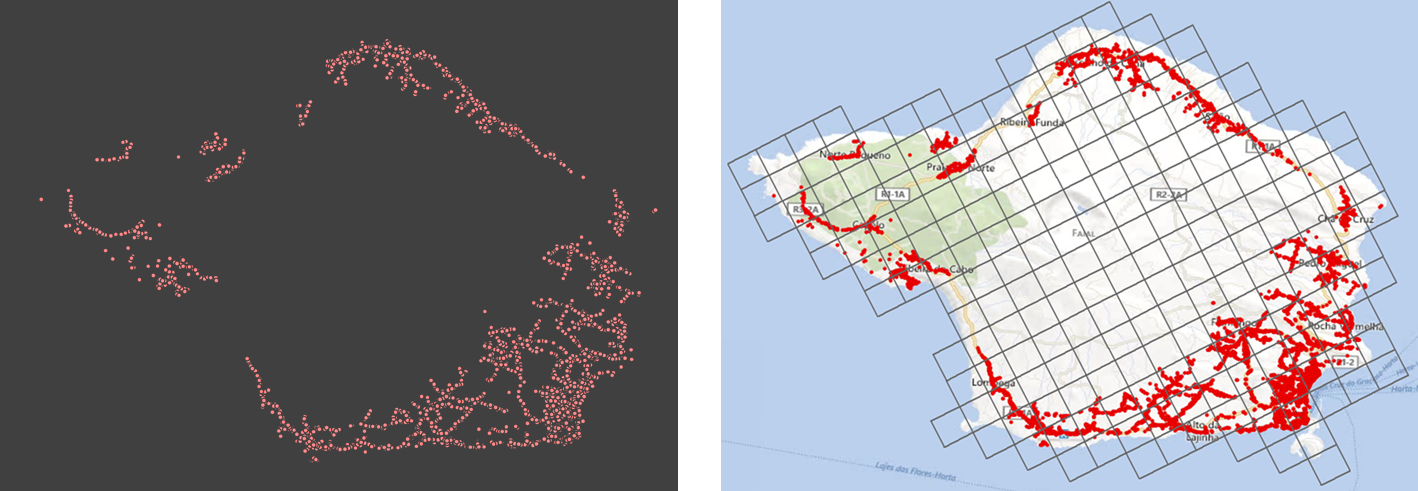
* ensure full integration of GeoINQ with other Statistics Portugal information management systems;
* GeoINQ is a subsystem of SIGINQ-IE and can only be invoked and depends on the authentication credentials of SIGINQ-IE;
* Interoperability between GeoINQ and different information systems is guaranteed by a set of Web services that ensure the integrity of the information of all the systems;
* GeoINQ can be used in mobile devices;
* Users can only access the features and geographical layers defined in the user’s profile according to their role in the SIGINQ-IE system (interviewer, supervisor, regional coordinator, national coordinator, methodology expert) and their analysis/geographic information needs;
* Must have functionalities to search, visualize, query and edit geographic data and related attributes.

*3.3 GeoINQ – Geospatial Data*

GeoINQ contains a geographical framework with several geospatial feature layers like NUTS, Portugal Official Administrative Boundaries (CAOP), European 1km2 GRID, Census Statistical Units (BGRI), sampling frames (BA), Interviewers residence building and Base maps, including the Portuguese Mapping Agency (DGT) orthophotomaps.

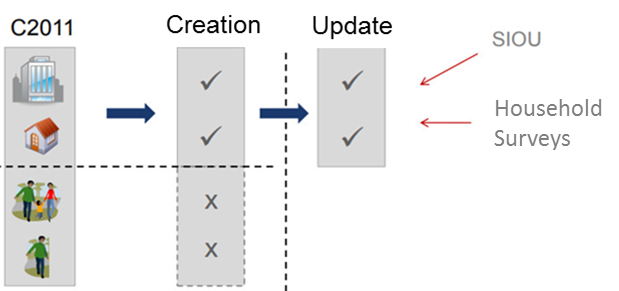
**Figure 5. GeoINQ – geographical framework**

The Buildings Geographical Database (BGE) contains the points (X,Y coordinates) of all the buildings georeferenced in the Census 2011 Operation.

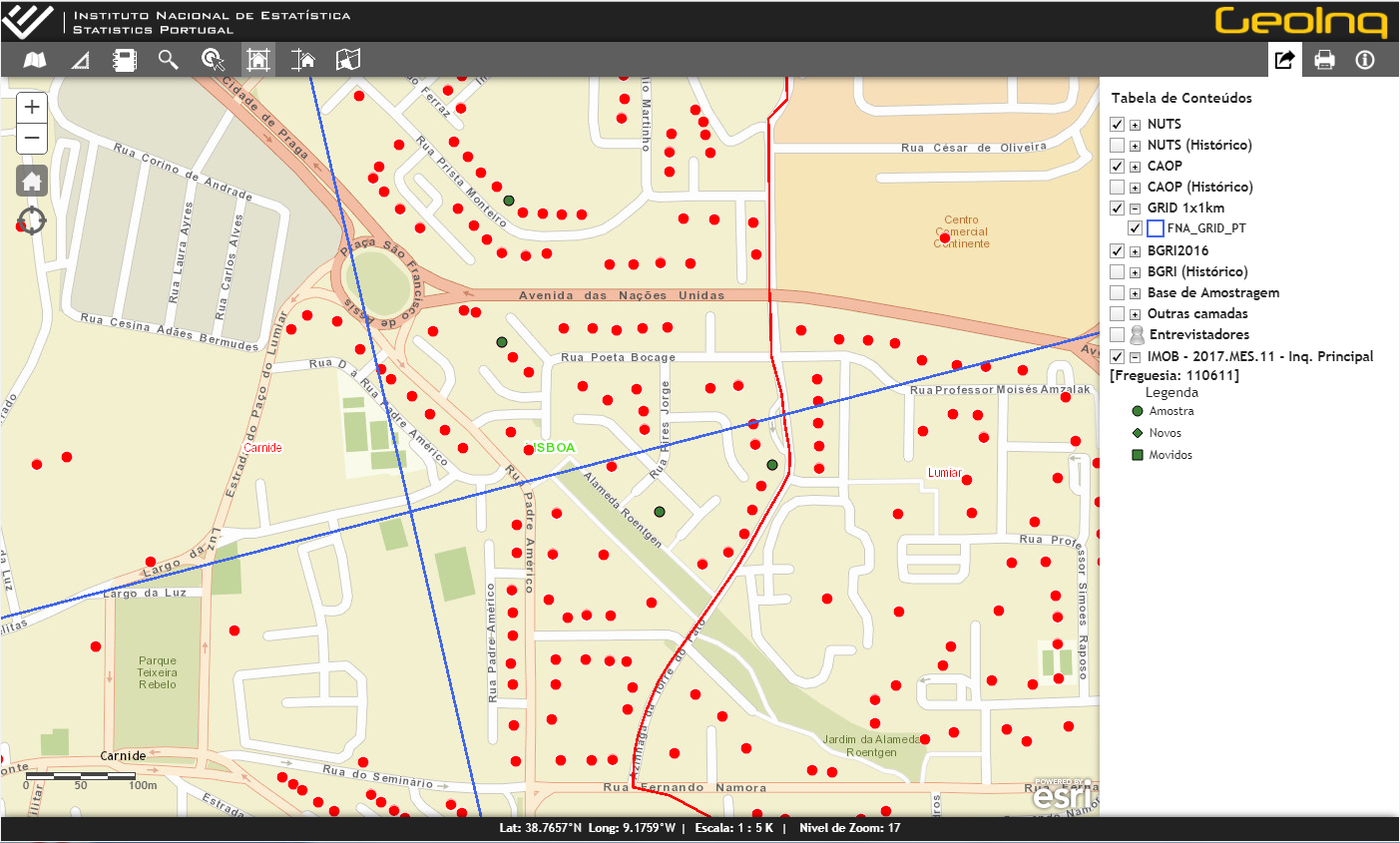
**Figure 7. BGE Building Points**

BGE is the national buildings and dwellings register (FNA) spatial component. FNA is an exhaustive file of buildings and dwellings collected in the last Census (2011). FNA has been continuously updated (buildings demolitions/new buildings, buildings location, address and several other attributes) from that moment onwards by several sources in particular by the Urban Operations Indicators System (SIOU – Building Permits and completed buildings from municipalities) and by Statistics Portugal Household Surveys.

**Figure 6. FNA – creation and update strategy**



The main GeoINQ geospatial features are the surveys samples assigned to interviewers - FNA dwelling units that are materialized by the BGE buildings points (points in green in the figure below).

**Figure 8. Surveys samples**

**4. GeoINQ – Functionalities**

GeoINQ app has a set of functionalities, supported in map services, which allow users to perform several geospatial operations – from the basic tools (zoom, search, query, etc.) to the more complex geoprocessing tools and creation of thematic maps.

*4.1 Authentication*

GeoINQ has an authentication system with individual credentials for each user. Authentication and identification of each user survey project are done via GPIEREG (one of SIGINQ-IE subsystems).

*4.2 Search and visualize*

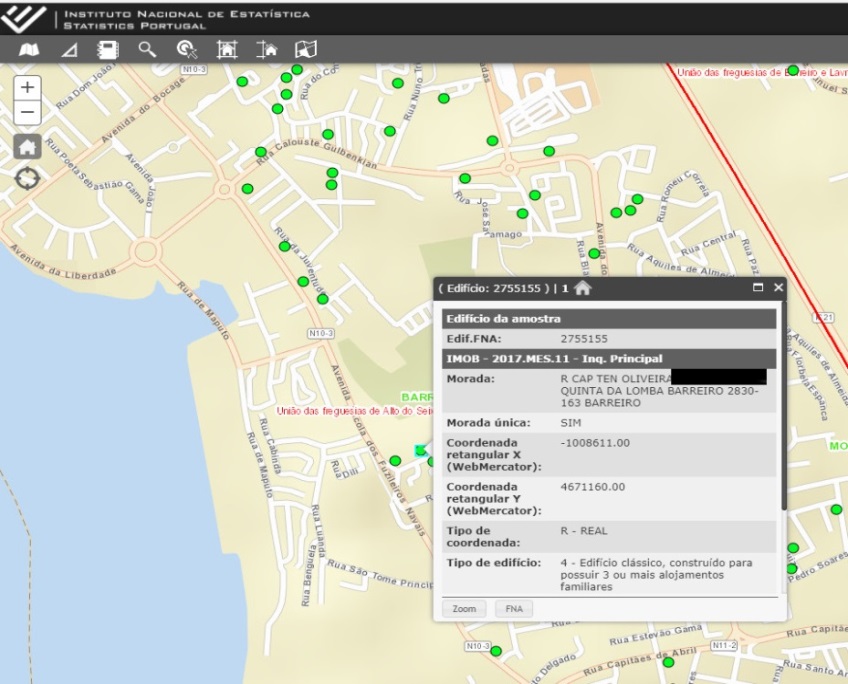
GeoINQ users can apply spatial filters to search and visualize statistical units and samples from the different surveys.

Users can only access the geographic and analysis information previously defined in their user profile. For example, only interviewers have permission to view the buildings of an entire sample from a specific survey that is assigned to them.

* 1. *Query*

All users can query and obtain information about the attributes of BGE/FNA buildings that integrate social/household surveys samples and their respective FNA dwellings

**8. GeoINQ – Query function**



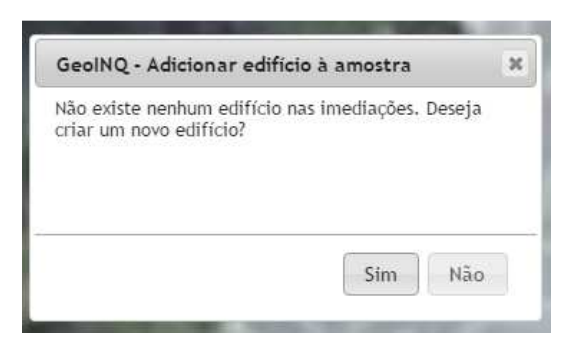
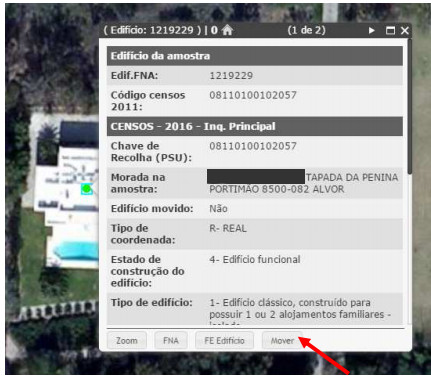
* 1. *Edit*

Interviewers that are in the field, collecting data in-person, have permission to edit BGE buildings geography and addresses directly in GeoINQ. With GeoINQ geoprocessing tools interviewers can move, add or delete BGE/FNA buildings.

GeoINQ is integrated with the FNA system and it also allows users to edit FNA dwelling units attributes.

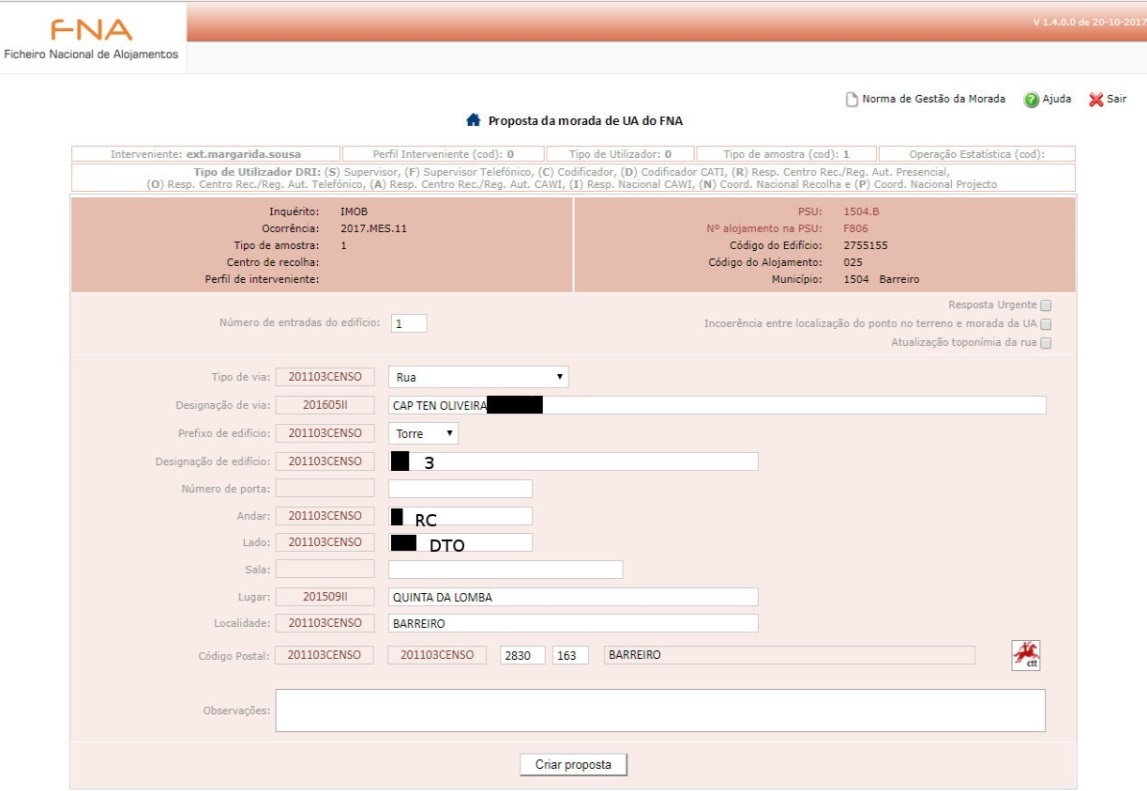
Edits to the BGE buildings geography, done directly at GeoINQ, or to FNA buildings/dwellings address, done at FNA app via GeoINQ, proposed by interviewers are validated by the Geoinformation and FNA Management Units.

**9. GeoINQ – Edit geography function**



**Click to add a new building**

**10. GeoINQ – Link to FNA App Function - Edit FNA dwelling attributes (address)**

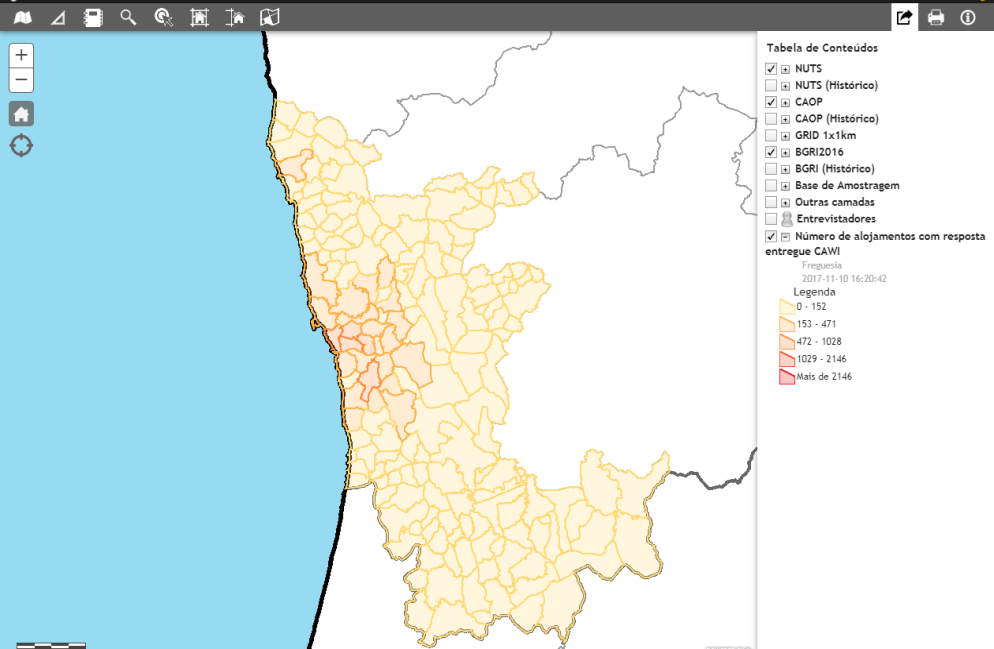
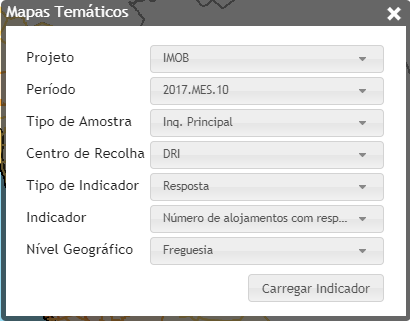


* 1. *Thematic maps*

GeoINQ allows users to create thematic maps for different geographic levels, with indicators from the different surveys samples in order to plan, monitor and analyze the data collection progress, and to support interviewers.

It is possible to create real time thematic maps with several indicators like the total number of dwellings from a specific survey sample, total number and percentage of dwellings with response, total number and percentage of dwellings with CAPI (Computer Assisted Personal Interview) response, total number and percentage of dwellings with CATI (Computer Assisted Telephone Interview) response, total number and percentage of dwellings with CAWI (Computer Assisted Web Interview) response.

**10. GeoINQ – Thematic Maps Function**



**5. GeoINQ – Main Users**

The **Statistical Methods Unit** is responsible for the surveys sampling system and for developing methodologies to monitor the respondent statistical burden. Statistical Methods Unit uses GeoINQ geographic information in the sampling process (as FNA and BGE led to a completely new approach and methodology for the constitution of sampling frames and samples based on geographic information and spatial analysis) and to analyze the geographical dispersion and overlap of samples in the national territory.

The **Data Collection Department** is responsible for the collection of statistical information by several methods (Web, phone, face to face). The interviewers reporting unit includes a hierarchical structure of organization to support data collection. The Data Collection Department use GeoINQ geographic information to assign samples to interviewers and to support interviewers that are collecting the data. It also uses GeoINQ thematic maps to monitor statistical operations and data collection progress.

**Interviewers** are responsible for the data collection and are, by far, GeoINQ main users because they represent around 90% of GeoINQ total users.

Before GeoINQ interviewers only had access to a table with the dwellings address, name and contact of the reference person of their samples. Some statistical units don’t have an address or the address is incomplete and unable to locate. Interviewers use GeoINQ and geographic information to solve difficulties in locating statistical units they need to survey. Interviewers also use GeoINQ to plan and monitor the data collection and to edit and update geographic information from BGE (Buildings) and the corresponding FNA buildings/dwellings attributes, contributing in this way to keep both databases up to date.

**6. GeoINQ – Conclusion**

GeoINQ app development required a wide and complex analysis of processes, for the integration of Statistics Portugal Spatial Data Infrastructure with SIGINQ-IE and other statistics Portugal information systems. GeoINQ allows the use of geographic information, in an innovative way, at several stages of Statistics Portugal statistical production process mainly in the data collection stage contributing to a rationalization of resources and a most efficient data collection and data collection management.