

# The Experience of reusing WebGIS application templates applied to the integration of statistical and geospatial information<sup>1</sup>

Pina Grazia Ticca\*

\*Istat, Italy, ticca@istat.it

**Abstract:** The paper describes the experience of re-using WebGIS application templates applied to the integration of statistical and geospatial data and metadata. That has been done within a project co-funded by the Eurostat and Istat, the Italian NSI. Statistical data have been extracted from Eurostat statistical DB (Census Hub and DB-Regio) and linked to geographical European NUTS2 and NUTS3. The reusable template is oriented to perform comparisons, interacting with maps, for the geospatial and temporal dimensions of statistical indicators.

Using geospatial ETL procedures ad-hoc designed, new geo and statistical information has been produced.

The ETLs could be later on applied to other thematic datasets; that ensures the possibility of further reuse for the WebGIS application template.

The prototype has been published through the Istat SDI, Gistat at the Istat official website, <http://gisportal.istat.it>

## 1 Istat and geospatial information

Istat is the Italian National Statistics Institute. Istat is active in the field of producing and disseminating geospatial statistical information since 1995, when the census mapping cartography was firstly digitized; then it was linked to the census data for dissemination purposes. In 20 years' experience geospatial processing and analysis applied to statistics had a significant growth and also the Istat GIS unit. Moreover specific GIS expertise was developed and GIS key roles have been designed to promote and support the use of geospatial information inside the NSI.

A geographic information system, named Gistat, has been developed and enhanced to support 2001 and 2011 censuses. Today Gistat is available on the Internet and used to publish Web MapServices of geostatistical data, available to programmers for a machine-to-machine access, and to end users through interactive web apps.

Nevertheless the lack of standards and coded procedures, or simply the inability to use traditional statistical systems, still leave to the GIS experts the use of geoprocessing tools and geospatial data; a real integration in the statistical processes has not still been reached.

<sup>1</sup> The project is carried out within the framework of collaboration between ISTAT Regional Office for Sicily and ISTAT Territorial and IT Directorate. Working team is coordinated by A.P.M. Mirto e P.G. Ticca and is formed by R. Abbate, M. Arcasenza, V. F. Bellafiore, F. Consentino, L. D'Alessandro, A. Gigantino F. Roberto, S. Scialanca

With the development of Gistat, several directions have been explored to promote the use of geospatial data and tools in the attempt to involve and stimulate statistical producers and users. Gistat was firstly created by harmonizing the census mapping cartography (enumerations areas and administrative boundaries) that Istat designs and updates to support the population and housing censuses. It is now a powerful GIS designed to support Istat processes that make use of georeferenced data. The spatial-temporal geodatabase stores, at various time periods, census mapping cartography as well as many layers created using geoprocessing tools such as: “commuting areas”, urban aggregates, election districts,..., and also data coming from external sources such as street networks, and aerial photos. The 1KM population grid is also present and it has been built according to EU specifications to support the cross-border statistics in the ESS (European Statistical System).

Istat is moving towards the establishment of the register of places and of geographical units; following the European and international recommendations, the objective is the definition of a point-based framework to have new territorial dimensions in the statistical production.

## 2 The GRANT project

The launch in 2016 of a GRANT program from Eurostat on “Merging statistics and geospatial information” was an opportunity for Istat to propose the development of an application entitled “A reusable webGIS application and a geospatial database schema for the EU countries comparison”. In August 2016 with the award of the mentioned GRANT, a project co-funded by Eurostat and Istat, was started. Istat designed a project based on the use of previous experiences in building WebGIS applications of georeferenced statistical data. The project will last 18 months.

The proposed project is related to the implementation of a Web GIS application based on an already used application template; it is composed up to 4 interactive independent map panels. In each geographic panel the users could compare data after choosing the statistical indicator, the year and the territorial zoom.

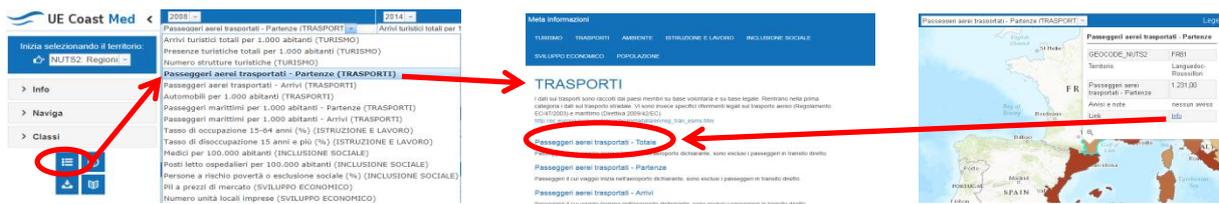


**Fig 1** The prototype application

Geospatial procedures to transform and harmonize the data (ETL) have been designed and used to process the geo and statistical information. The development of ETLs will ensure the possibility of re-use again the application, in the context of dissemination of other thematic data.

In the prototype developed, see the figure 1, geographic data are referred to NUTS2 and NUTS3 levels according to the availability of historical series of the chosen statistical indicators. In the prototype the data are limited to the EU Mediterranean area; only a subset of indicators and time series have been used, the localization to English has been designed but not still fully implemented. The statistical indicators used concern different thematic areas, such as: tourism, environment, transportation, population, economic growth, education,...

Furthermore a metadata template site has been used and customized to provide documentation information. That documentation information is accessible in different ways from the application (see figure 2) so it can constantly be queried by users; it is related to the data, to the application, to the representation techniques and classification methodology used,....



**Fig 2** the prototype application: example of metadata navigation

As already mentioned the users can interact with the application choosing the statistical indicators, the year of reference and the territorial level of analysis. The interaction is possible through GIS standard tools such as pan, zoom, objects clicking,.....

Special care has been used in creating representation intervals for the thematic maps; the domain of the yearly range of values was processed to understand the best method for breaking the classes. The colours have been chosen and optimized to assure the readability of the maps in comparing the data.

The project is still ongoing and the main outputs foreseen at the completion of the projects are:

- the Web GIS application available on Gistat platform

- the documented schema of the geodatabase, including relationships to integrate statistical data
- the geographical and statistical datasets at EU level for the coastal countries integrated to statistical indicators (available for download)
- the ETL procedures to populate the geodatabase after extracting data from the Eurostat statistical DB (Census Hub and DB\_Regio) and producing links to NUTS2 and NUTS3 datasets.

### **3 The experience of re-using the WebGIS application templates**

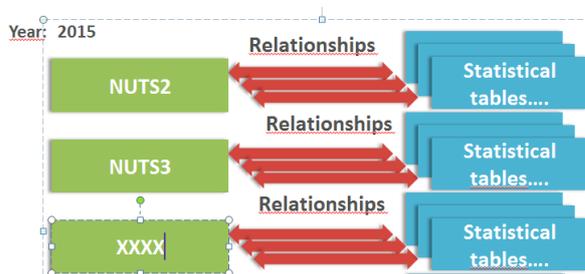
In the last five years, through Gistat platform, Istat published many Web GIS applications. Through the use of WebMapServices, geospatial information linked to statistical data are available to be accessed and analysed from users. Users can create and print thematic maps of statistical indicators for different geographic levels and time periods. Users can view, access, spatially query statistical data; they can overlay maps and perform geographical comparisons.

The web applications development generally has been done starting from scratch. For every application the design phase was oriented to establish the GIS needs for statistical users; this is not an easy task. Moreover the data preparation phase has always been very long and complex. Each application required its own geodatabase linked to statistical data.

In the recent developments the reuse of experiences, data models, web application templates has been promoted in order to reduce the preparation steps and especially the process of updating the data. Indeed very often, many products were abandoned due to the efforts needed to update the data or to support the IT evolution. Generally it is more convenient a new development instead of the maintenance of the existing products, both in terms of data, both in terms of application upgrading and enhancement.

With the opportunity offered by the GRANT project, Istat GIS unit tested the reuse of a data model designed to link statistical and geographic data and already used in other applications. That data model is suitable to design and publish WebMapServices optimized to be consumed by WebGIS applications. Moreover the production of those WebMapServices made also feasible the reuse of WebGIS applications templates oriented to the dissemination of georeferenced statistical information.

Statistical data have the time series and the chosen data are serialised in years; therefore the designed model for geodata is based on that time units. In the geodatabase geographical entities are defined and stored by year (if changes happened in the territory). The relationships to the blocks of indicators are based on the yearly keys of statistical data that are the same of the corresponding keys in the geographic layer per year (see figure 3).



**Fig 3** the data model for the geodatabase

The model has been then translated into a physical schema, implemented inside GISTAT geodatabase, where the data have been loaded. Naming rules have been used for the field names and aliases have been set to guarantee the readability of the indicators (figure 4). Even relationships names have been coded using naming rules, so that nothing is hardcoded in the web application.

The preparation of the WebMapServices is also guided by Python procedures; a WebMapService per year has been dynamically generated. The authoring schemas are produced and dynamically linked to the geodatabase tables (figure 5).

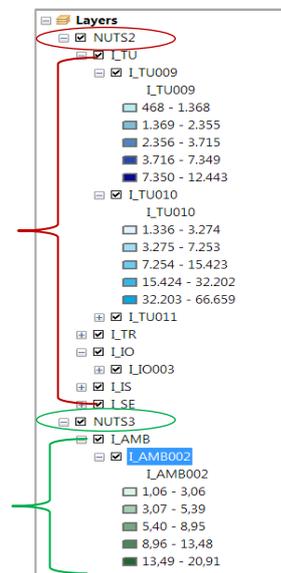
Everything has been automated, where possible, following the logic of reusable components.

Field Name	Data Type
GEOCODIE_NUTS2	Object ID
Indirizzo	Text
I_IO004_2008	Double
I_IO004_2009	Double
I_IO004_2010	Double
I_IO004_2011	Double
I_IO004_2012	Double
I_IO004_2013	Double
I_IO004_2014	Double
I_IO004_2015	Double
I_IO004_Annis_e_note2008	Text
I_IO004_Annis_e_note2009	Text
I_IO004_Annis_e_note2010	Text

Field Properties:

Alias	Tasso occupazione 15-64 an
Allow NULL values	Yes
Default Value	

**Fig 4** the geodatabase table



**Fig 5** schema of a WebMapService

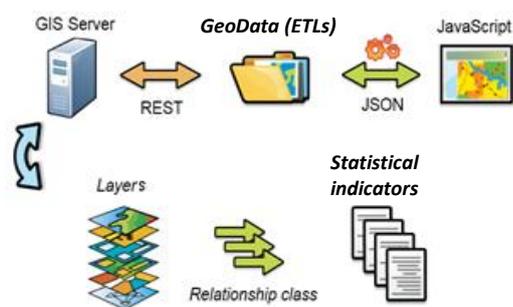
## 4 Conclusions

The lack of standards and guidelines for sharing integrated statistical and geospatial information is evidently a weakness in the process of enhancing the use of geospatial information in the statistical dissemination and production processes.

The described project, still ongoing, is for Istat a good opportunity to proceed in the direction of integrating statistical and geospatial information; that has been done re-using knowledge and previous geospatial experiences. That is not “*using geospatial standards*”, but it could lead to the definition of internal guidelines and to the production of reusable tools, templates and geodata models that surely will make the integration of the geospatial dimension easier in the statistical production, analysis and dissemination systems.

In this project a template GIS application to compare statistical maps on different panels was re-used. The idea came from the success of publishing other WebGIS application on GISTAT website, where statistical indicators could be compared in space and time, using more than one panel. In this way statistical users could integrate geospatial tools in analysis and dissemination processes.

As already mentioned, the project was not fully based on *geospatial standards*, but the reuse of the previous experiences has been a key factor in the success of the project itself. The time to setup the requirements and the entire project was optimized. The choice of a data model already implemented was a guide for the design of ETL procedures and WebMapServices. The experience in producing WebGIS applications and the availability of a production environment (GISTAT platform) completed the setup of a re-usable geospatial workflow for Istat.



**Fig 6** Re-usable geospatial workflow