Open Web Services: new tools for Medievalist Historians to manage and share their research work

Daniela Carrion[1], Federica Migliaccio[1], Guido Minini[1], Cynthia Zambrano[2]

[1] Department of Civil and Environmental Engineering (DICA), Politecnico di Milano,

Piazza Leonardo Da Vinci 32, Milano, Italy

[2] E-VIA S.p.a, Retelit Group, Milano, Italy

Abstract

The research project presented in this paper regards a joint cooperation between Medievalist Historians and Geomatics experts. The idea was to study if new tools based on Geomatics technologies could provide Historians with new methods to develop their studies, archiving digital data in a geo-database and “spatializing” the information sources to produce maps in a GIS environment. So, in the first part of the project, a GIS was designed and implemented based on data collected from documents preserved in the Italian National Archives. Afterwards, the problem of data integration and sharing among research groups working on Medieval History was dealt with, in order to make data available for the consultation and query by several research groups. Three approaches based on a client-server architecture have been explored: one is typical of the WebGIS architecture; the second one is based on OGC Web services and the third one exploits a Web page while the GIS tools are provided by a Desktop GIS installed locally on a PC. In the paper, the different approaches will be described, in order to underline advantages and disadvantages with respect to the Historians' requirements.

Keywords

Historical GIS, Medieval Geodata, Web Service, FOSS, QGIS

1 Introduction

The research project presented in this paper is the result of many years of research and cooperation between Geomatics experts and Medievalist Historians. The main aim of the cooperation was to provide the Historians with a modern tool to manage, visualize and share their sources exploiting Geographic Information Systems. The first phase of the work consisted in understanding the characteristics of the Medieval data to be organized in a GIS environment, in order to design a suitable database structure (Carrion, Migliaccio, Minini, & Zambrano, in press) for the GIS. This project has been dealt with by using a cognitive approach, aimed at identifying the key issues and the appropriate tools to help the Historians in satisfying their research

needs.
At first, data collected from historical sources have been organized into table archives; these data have been associated with spatial information (point coordinates), corresponding to the place names found in historical sources: knowing the spatial reference it is possible to import data into a Geographic Information System (GIS), to display their location on the area of interest and to overlay them with existing base maps. The data used for the project date back to the late Middle Ages (mid-Fifteenth Century) and represent fiscal information, related to various kinds of taxes paid to different Kings or Princes in the Kingdom of Naples. Then, a relational database has been designed and populated with the data. The relational database design has been achieved through the definition of the Entity-Relationship Diagram (ERD).
In the second phase of the project, that is the main focus of this paper, the problem to be resolved was how to allow Historians to share and query their data exploiting the Web. Three kinds of approaches have been explored. The first one is typical of the WebGIS architecture, in which all the data and the GIS analysis tools are stored in a server and published on the Web, so that the only client-side software requested is a Web Browser. The second one is a Web service, compliant with the OGC standards (http://www.opengeospatial.org/standards), namely a Web Mapping Service, (WMS), a Web Feature Service (WFS) or a Web Processing Service (WPS). The third approach is again based on a client-server architecture, but in this case only the data, collected into a DB and managed through a Database Management System (DBMS), are stored in a server, while the GIS tools are provided by a Desktop GIS installed locally on a PC.
In the next chapters these approaches will be discussed in more detail, with the purpose of understanding which of them can be more effectively applied in the case of the proposed historical application, taking into account the specific requirements of this project.
Regarding the GIS technologies to be exploited, it has been decided to use only Free Open Source Software: this goal raises from the need to develop and maintain the system for a long period with limited economical resources.
Another important objective pertains the ease of use of the GIS, because in this instance the end users are not GIS experts, so the system must be complete but also simple and intuitive to use.

2 State of the art

GIS potentiality is spreading to many disciplines, including History (Schlichting, 2008). At the international level, many works have to be mentioned, see e.g. (Gregory, 2002), (Gregory, Bennet, Gilham, & Southall, 2002), (Berman, 2005) and (Gregory & Healey, 2007), being the main references for this kind of research.
Many examples of historical data organized into national atlases can be found, see e.g. (Pawson, 1997) and (Black, 2003); in some cases data are structured into relational databases, to be managed into a GIS environment, see e.g. (Ardissone & Rinaudo, 2005); in most cases they are related to census data and rarely they date before the XVI century: (Boonstra, Collenteur, & van Elderen, 1995), (De Moor & Wiedemann, 2001), (Fitch & Ruggles, 2003).

The possibility to share georeferenced data on the Internet (Tait, 2004) is precious to improve the cooperation among scientists. It is possible to find some examples of historical data organized into a GIS and published on the Internet, such as the China Historical GIS (http://www.fas.harvard.edu/~chgis/) or the northern Italy cadastral map WebGIS (Brovelli, Minghini, Giori, & Beretta, 2012). Another interesting example published on the Web is constituted by the "E 179 database", (http://www.nationalarchives.gov.uk/e179/) which contains records relating to lay and clerical taxation and which is included into the United Kingdom National Archives. However, in this case the geographic component is not made explicit into a map.
It must be also mentioned that, when publishing geodata on the Web, the attention to interoperability is essential (see Maguire and Longley, 2005; Johnson et al., 2011).

3 The historical data

One of the aims of the project is to create a GIS that can be useful for Historians to draw maps from historical data in order to study the territorial dynamics emerging from Medieval sources, achieving a cartographic representation of the spatial distribution of historical information.
Currently three historical sources have been collected into the DBMS:

- the *Liber Focorum Regni Neapolis*, produced by the financial offices of Alfonso V of Aragon in the 1440s, after his conquest of the Neapolitan kingdom;
- the *Quaterni declaracionum*, produced between 1446 and 1463 by the leading Treasury officials (the *magistri rationarum*) of the last Prince of Taranto, Giovanni Antonio Orsini del Balzo, the greatest feudal vassal of the kingdom (Pizzuto, 2009);
- the *Quaternus decimarum*, drawn up in 1478 by Paolo Vassallo, bishop of Aversa, that registers the payments of the *Decima*, a tax related to ecclesiastical benefits (Mangia, 2013).

From the data collected by Historians, several entities have been defined following the rules of relational databases. A fundamental entity was the one referring to the historical place names, which could then be connected with present-day place names, allowing for the geo-referencing of information by means of map coordinates in UTM WGS84 mapping and reference system. The geo-referencing of the location has been performed by the Historians with the support of the Geomatics experts. Another crucial entity for Historians is the source, which must be always linked to the data stored into the database. Besides, entities regarding census data, types and amount of taxes due and collected, historical characters names and functions, and other, have been defined for the purposes of the project.
Historians' know how was fundamental during the phase of extraction of data from historical archives and translation of the information into digital format (typically into a spreadsheet). Also the conceptual model of the database has been planned and designed in cooperation with Historians, in order to understand their needs and the meaning of the collected information (in particular the preparation of a data dictionary was necessary).

The possibility to easily visualize on a GIS-based map the locations linked to the data collected from the historical sources is a remarkable achievement for Medievalist Historians. From the Geomatics point of view the spatial component of the information is really simple: once the Historians have defined the location corresponding to the present placement, it translates into point coordinates. What is challenging is to allow the scholars with an efficient tool to manage the complexity of tables which are linked to those coordinates with the relationships designed in the ERD.

4 Data publication on the Web

Also during the discussion on issues related to the publication of data the participation of Historians was again fundamental, because they represent the end users of the GIS, so it is important to understand what tools can be useful for them. The database design and the historical sources organization into a GIS has been presented in Carrion et al. (in press).
For easy data sharing among the Medievalists, the choice of the Web architecture is very important, considering the elements which are crucial in this peculiar case:
i) reduced costs for software acquisition, system implementation and maintenance;
ii) ease of use for non expert users;
iii) ease of maintenance over time;
iv) scalability, ease of integration of new data.
As mentioned in the previous paragraph, the data considered in this project show their complexity in the database structure. So, another key point is to allow the management of the relationships among the database tables.
We can then summarize the system requirements as follows:

- exploitation of Open Source Software, in order to comply with the low budget available;
- access to the database tables, including the ones not containing georeferenced data;
- access to the relationships between entities;
- possibility to display the results of the table queries on maps and vice-versa;
- user friendliness;
- simplicity (to reduce the resources needed both for the system development and its maintenance).

In the following three possible approaches for Medieval data sharing through the Web will be presented and discussed, showing their advantages and disadvantages.

4.1 Historical WebGIS

The first type of GIS architecture approach adopted and implemented to publish the Medieval data has been a Historical WebGIS. The WebGIS (Carrion, Migliaccio, Minini, & Zambrano, 2014) is based only on Free Open Source Software; its architecture is shown in Figure 1.
The server side is stored in an Ubuntu Linux virtual machine and is composed

by three modules: a DBMS, a GIS Server and a Web Server.
The DBMS used is PostgreSQL, an Open Source Software that can be integrated with a PostGIS extension, which manages the geographical information stored in the database and allows to interact with QGIS Mapserver, a WMS server that takes advantage of the QGIS libraries and uses the .qgs projects created with QGIS Desktop. The GIS server is invoked by the Apache Web Server, whose task is to publish web pages.
The client side is entrusted to a Web Browser, which interprets the web pages written in HTML5 language. The three components of this language are HTML, CSS and JavaScript: they manage contents, customization and behavior of web pages. The JavaScript toolkit GeoExt was used: it is composed by the ExtJS library, which allows to improve the graphical user interface and to integrate them with grids, buttons and toolbars, and the OpenLayers library, which allows to include web-mapping functionalities.
The Historical WebGIS, named *Geografie Medievali* (Medieval Geographies) and published into a Website, can be accessed through a "WebGIS" mode or a "Show table" mode (see Figure 2 and 3).

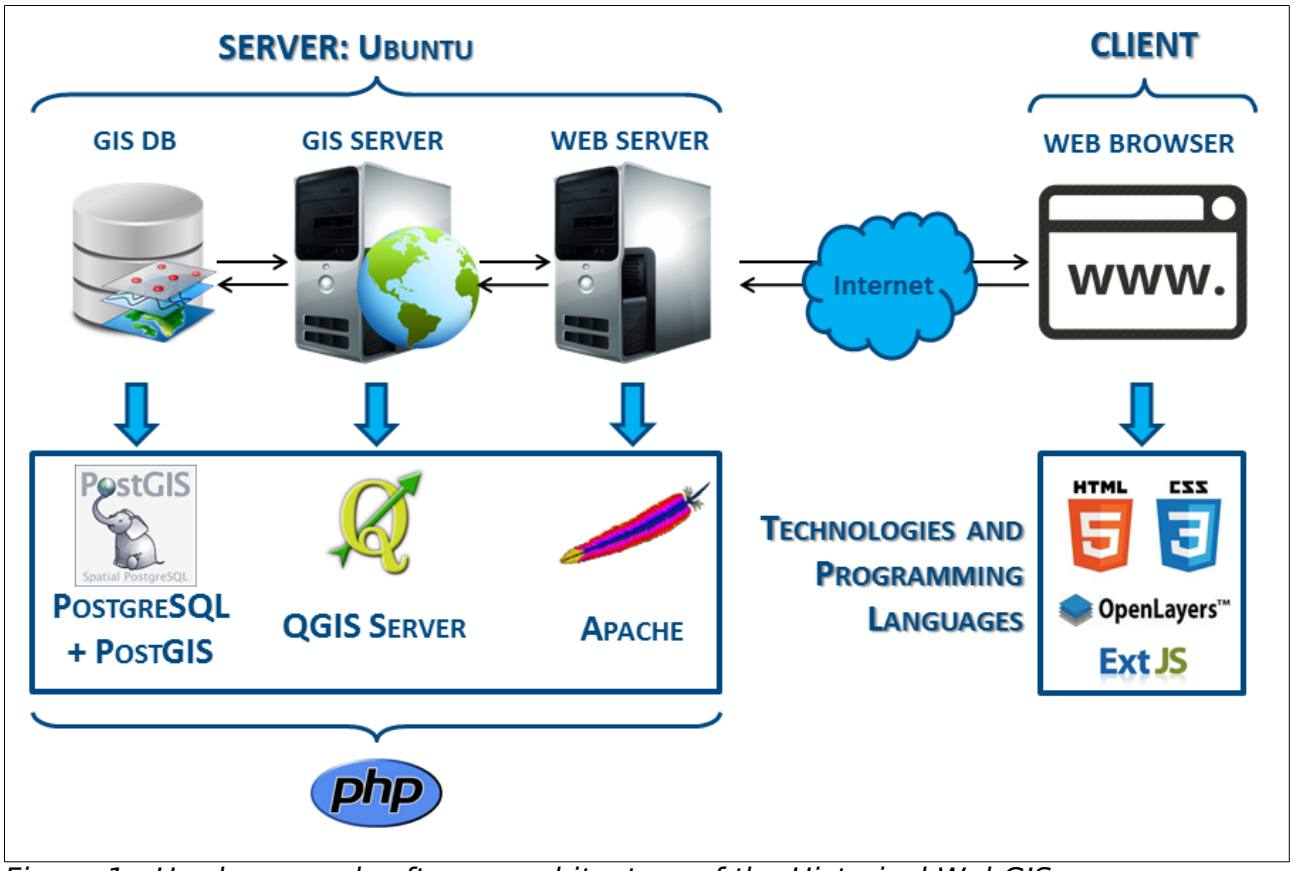


Figure 1 - Hardware and software architecture of the Historical WebGIS.

However, it must be underlined that the WebGIS only represents the georeferenced entities of the database and is based on the architecture just described, while the "Show Table" page displays all the entities, georeferenced and not, and some views that combine the information of two or more entities in one table, exploiting the relationships between them; this page is realized

again with ExtJS libraries and allows to query tables and views with a query builder tool.
The typical advantage of this kind of architecture is that the installation of GIS software is not required on the client-side, because all the GIS functionalities are implemented into the server. On the other hand, in this case, the GIS tools provided by a GIS server are less advanced than those provided by a Desktop GIS, since every tool must be integrated or developed on purpose and the resources available for this project are very limited. Then, the main disadvantage of this solution is that every change must be programmed *ad hoc.*
The idea is to integrate over time other historical sources into the database, allowing Historians to explore the connections of the spatial information over the past. In this perspective a system where every upgrade requires complex programming is not feasible, and for this reason new approaches are being studied.

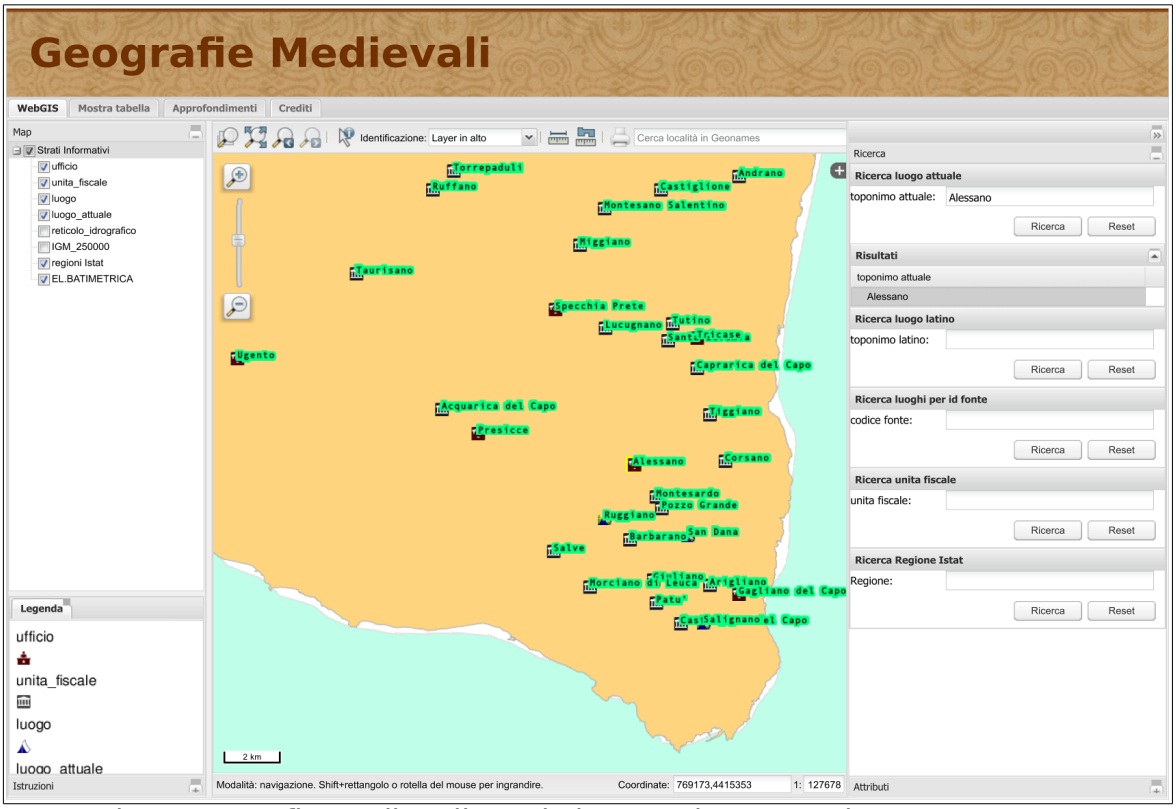
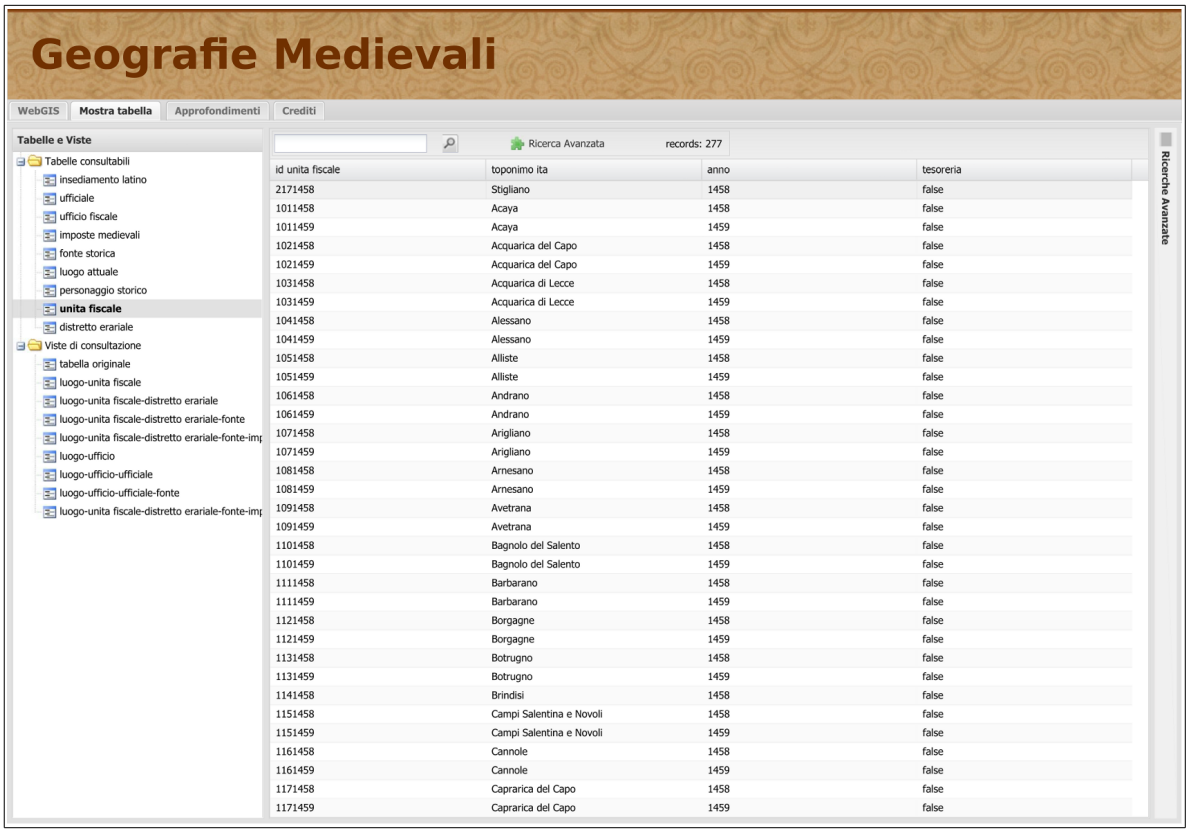


Figure 2 - The "Geografie Medievali" Website: "WebGIS" mode.



id unita fiscale	toponimo ita	anno	tesoreria
2171458	Stigliano	1458	false
1011458	Acaya	1458	false
1011459	Acaya	1459	false
1021458	Acquarica del Capo	1458	false
1021459	Acquarica del Capo	1459	false
1031458	Acquarica di Lecce	1458	false
1031459	Acquarica di Lecce	1459	false
1041458	Alessano	1458	false
1041459	Alessano	1459	false
1051458	Alliste	1458	false
1051459	Alliste	1459	false
1061458	Andrano	1458	false
1061459	Andrano	1459	false
1071458	Arigliano	1458	false
1071459	Arigliano	1459	false
1081458	Arnesano	1458	false
1081459	Arnesano	1459	false
1091458	Avetrana	1458	false
1091459	Avetrana	1459	false
1101458	Bagnolo del Salento	1458	false
1101459	Bagnolo del Salento	1459	false
1111458	Barbarano	1458	false
1111459	Barbarano	1459	false
1121458	Borgagne	1458	false
1121459	Borgagne	1459	false
1131458	Botrugno	1458	false
1131459	Botrugno	1459	false
1141458	Brindisi	1458	false
1151458	Campi Salentina e Novoli	1458	false
1151459	Campi Salentina e Novoli	1459	false
1161458	Cannole	1458	false
1161459	Cannole	1459	false
1171458	Caprarica del Capo	1458	false
1171459	Caprarica del Capo	1459	false

Figure 3 - The "Geografie Medievali" Website: "Show table" mode.

4.2 OGC Web services

The second approach which has been explored, however not yet implemented, is the possibility to publish the database through OGC Web services. The services which correspond to the OGC standards that have been considered are WMS, WFS and WPS. WMS is too simple for our case: in fact it only allows showing a map. To take into account the complexity of the designed database a WFS or a WPS would be necessary. However, a strong implementation work is needed also in this case, in particular if Free Open Source Software is used. A strategy is suggested in Peng and Zhang (2004). It is important to underline that what is crucial for the Medieval data considered in this work, is the possibility to be able to exploit the complexity of the database structure. From this point of view, a significant drawback of the OGC services approach for this project is the impossibility to display tables corresponding to not-georeferenced entities.

The advantage of this solution is again the fact that no software is needed to be installed on the client-side, apart possibly for some plug-ins. Moreover, the interface can be designed on purpose, allowing to make it as simple as possible, for non expert users.

The main disadvantage is still very similar to that of the WebGIS approach: the heavy programming (Rautenbach, Coetzee, & Iwaniak, 2013) which is needed, both during the implementation and the maintenance phases of the project, which makes this solution less appealing, for this low-budget project. In addition, usually a Web Service provides much less functionalities with respect to GIS Desktop software.

4.3 Web page and QGIS client architecture

The third approach presents an architecture (shown in Figure 4) that is much simpler than the one of the WebGIS: in this case only the GIS database is stored into the Server; the Client must be provided by GIS software e.g., QGIS Desktop, that natively supports the connection with a PostgreSQL database through the Internet.

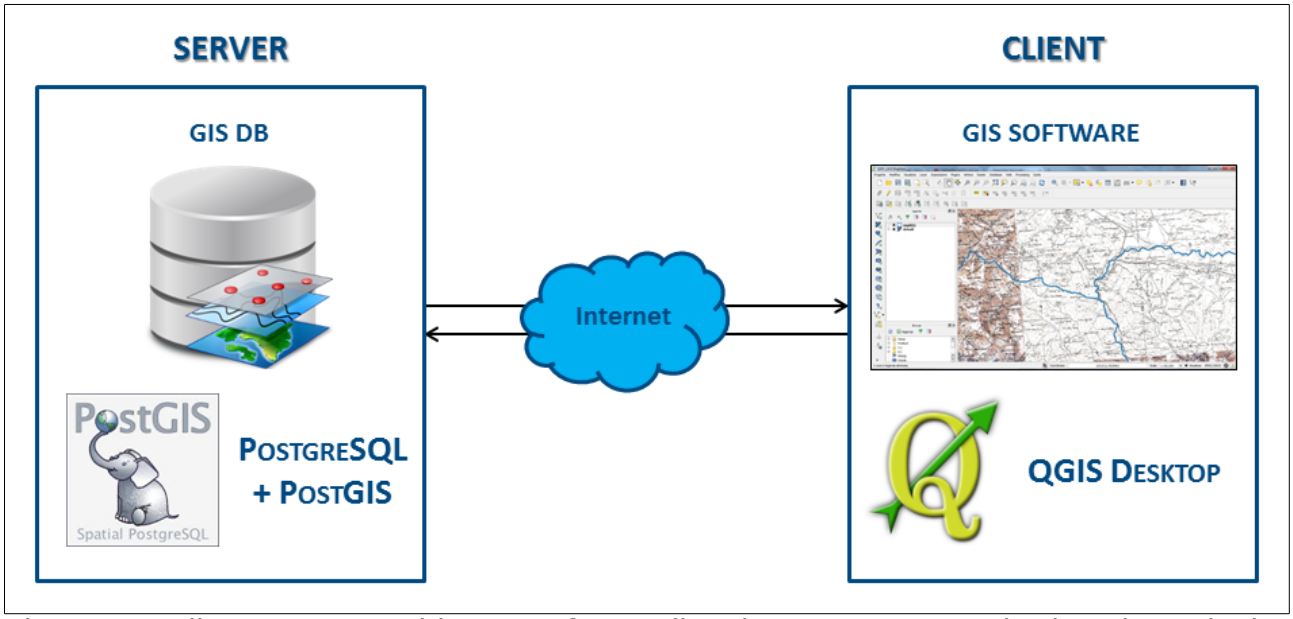


Figure 4 - Client-Server architecture for Medieval Geo-resources sharing through the Web.

Of course, GIS Desktop software offers many tools that are not usually available in a WebGIS architecture (like the one proposed in section 4.1), unless they are specifically designed; moreover, QGIS Desktop is an Open Source Software that does not require licensing costs and, in recent years, is catching up with commercial software like ESRI ArcGIS from the point of view of completeness and performances.
The main problem that remains in this approach is that the relationships implemented into the database are not recognized by QGIS, so when the connection with the DBMS is realized, the entities imported into QGIS, shown as layers or attribute tables, are not associated through relationships. This is a significant limitation when working with a database in which the non-spatial tables play an important role, like in our case.
One possible solution is the creation of views into the database, that contain the information stored in two or more entities. The main disadvantage is its rigidity and the increasing in number of attributes contained in a single table, when joining many entities, which makes the consultation of data much more difficult and less intuitive.
The solution which is being investigated at this moment, since it could prove to be more advantageous, is to exploit the possibility of creating some scripts that work into the QGIS environment and allow the software to take into account the relationships among entities established in the database, as it has been proposed by Cho, Bellemans, Janssens, and Wets (2014).

The entities imported into QGIS from the Historical Database and their cartographic representation are shown in Figure 5. They are overlaid onto a basemap of the Italian Military Geographic Institute (IGM) made available as WMS service on the Italian National Geoportal.

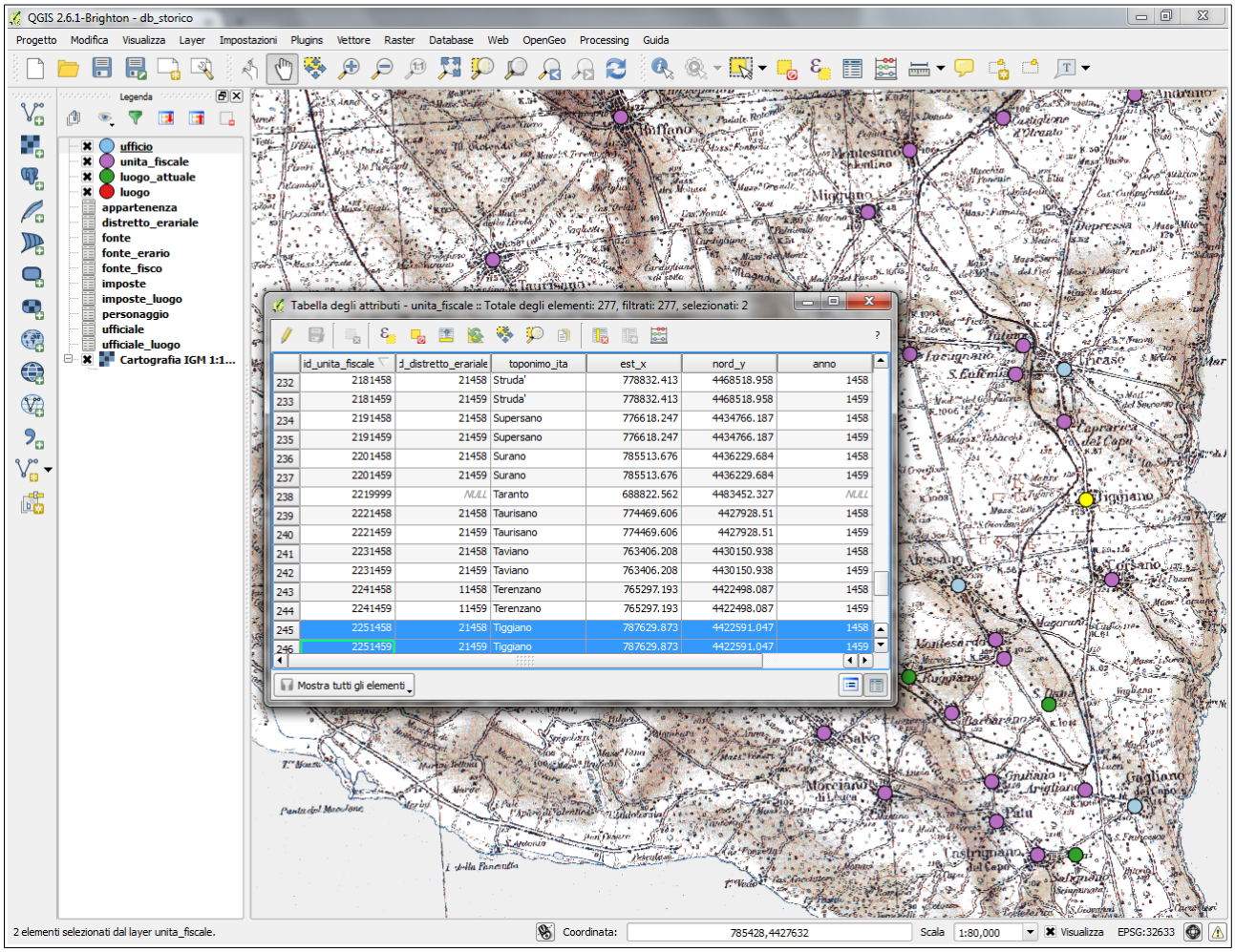


Figure 5 - The Historical Database imported into QGIS.

5 Conclusions

The research presented in this paper has the final objective to implement tools allowing to share a Historical Medieval database on the Web and to manage it through a GIS environment, exploiting all its potentiality and completeness (all the entities related with each other, as represented in the ERD design). The goal is to supply Historians participating to the project with an instrument that makes it possible to realize complex queries on Historical data and to show the outcome of spatial queries on maps.

Different approaches have been explored, namely: a WebGIS, which has been implemented and tested by Historians, OGC Web Services (such as WMS, WFS and WPS) and a Web page – QGIS Desktop combination with *ad hoc* implemented tools. In Table 1 the compliance of the proposed approaches with respect to the system requirements listed in section 4 is shown.

	WebGIS	Web Services	Web page +Desktop GIS
Open Source Software	Yes	Yes	Yes
Access to the database tables	Yes	No (WMS) Partial (WFS or WPS)	Yes
Access to the database relationships	Yes, but it needs ad-hoc implementation	No	Yes, but it needs ad-hoc implementation
Possibility to see the results of the table queries on the map and viceversa	Yes, but it needs ad-hoc implementation	Yes (WFS or WPS)	Yes
User friendliness	Yes	Yes	Yes
Maintenance simplicity	No	Yes	Yes

Table 1 – The proposed approaches against the requirements of the project.

According to the evaluation presented in Table 1 the approach based on the Web page – Desktop GIS combination with *ad hoc* implemented tools has been considered as the most convenient with respect to the Historians' needs and is being implemented. In the foreseen scenario, geodata will be stored on a server and published on a Website, so that the data sharing features of the system will be maintained. Moreover, thanks to the user friendliness of the last versions of QGIS, a very large set of functionalities will be available to Historians with a limited training effort on their side.
Many studies based on the application of GIS technologies to Historical studies have been and are being carried on, however the case discussed in this paper presents quite a unique character, since it deals with data from the Medieval period, for which such applications are still rather uncommon.

References

- ✔ Ardissone, P., & Rinaudo F. (2005). A GIS for the management of historical and archaeological data. In *Proceeding of the CIPA 2005 XX International Symposium*. Torino.
- ✔ Berman, L. M. (2005). Boundaries or networks in Historical GIS: Concepts of measuring space and administrative geography in Chinese History. *Historical Geography, vol. 33, Special issue: Emerging trends in Historical GIS*. Baton Rouge.
- ✔ Black, J. (2003). Mapping the past: Historical atlases. *Orbis, 47*, 277-293.
- ✔ Boonstra, O., Collenteur, G., & van Elderen, B., eds (1995). Structures and contingencies in computerized historical research. In *Proceedings of the IX International Conference of the Association for History and Computing*. Nijmegen, 1994 (Hilversum, 1995).
- ✔ Brovelli, M. A., Minghini, M., Giori, G., & Beretta, M. (2012). Web

Geoservices and ancient cadastral maps: The Web C.A.R.T.E. project. *Transactions in GIS 16*, 125-142.

- ✔ Carrion, D., Migliaccio, F., Minini, G., & Zambrano, C. (2014). Sharing historical geodata through a WebGIS. A Web application for History researchers. *Conference of Informatics and Management Sciences March, 24. - 28. 2014*, 331-334.
- ✔ Carrion, D., Migliaccio, F., Minini, G., & Zambrano, C. (in press). From historical documents to GIS: A spatial database for medieval fiscal data in Southern Italy, *Historical Methods: A Journal of Quantitative and Interdisciplinary History*
- ✔ Cho, S., Bellemans, T., Janssens, D., & Wets, G. (2014). Integrating GIS and FEATHERS: A conceptual design. *Procedia Computer Science 32*, 405-412.
- ✔ De Moor, M., & Wiedemann, T. (2001). Reconstructing territorial units and hierarchies: a belgian example. *History & Computing 13*, 71-97.
- ✔ Fitch, C. A., & Ruggles, S. (2003). Building the National Historical Geographic Information System. *Historical Methods: A Journal of Quantitative and Interdisciplinary History 36*, 41-51.
- ✔ Gregory, I. N. (2002). A Place in History: A guide to using GIS in historical research. *Oxbow*. Oxford.
- ✔ Gregory, I. N., Bennet, C., Gilham, V. L., & Southall, H. R. (2002). The Great Britain Historical GIS: From maps to changing human geography. *The Cartographic Journal*.
- ✔ Gregory, I., & Healey, R. (2007). Historical GIS: Structuring, mapping and analysing geographies of the past. *Progress in Human Geography 31*, 638-653.
- ✔ Johnson, G., Gaylord, A., Franco, J., Cody, R., Brady, J., Manley, W., Dover, M., Garcia-Lavigne, D., Score, R., & Tweedie, C. (2011). Development of the Arctic Research Mapping Application (ARMAP): Interoperability challenges and solutions. *Computers & Geosciences 37*, 1735-1742.
- ✔ Maguire, D., & Longley, P. (2005). The emergence of geoportals and their role in spatial data infrastructures. *Computers, Environment and Urban Systems 29*, 3-14.
- ✔ Mangia, M. A. (2013). *La diocesi di Benevento nel Quaternus decimarum del 1478: un contributo alla realizzazione di un GIS di cartografia storica sul Mezzogiorno bassomedievale*. Master's degree thesis, Università del Salento.
- ✔ Pawson, E. (1997). The New Zealand Historical atlas. *Journal of Historical Geography 23*, 496-499.
- ✔ Peng, Z., & Zhang, C. (2004). The roles of geography markup language (GML), scalable vector graphics (SVG), and Web feature service (WFS) specifications in the development of Internet geographic information systems (GIS). *Journal of Geographical Systems 6*, 95-116.
- ✔ Pizzuto, S. (2009). Il Quaternus declaracionum di Francesco de Agello (1450-1461). Un contributo allo studio della geografia politica del Principato di Taranto in età orsinana. In *I domini del principe di Taranto in età orsiniana (1399-1463). Somaini, F. and Vetere, B. eds,*

Congedo Editore, 61-76.

- ✔ Rautenbach, V., Coetzee, S., & Iwaniak, A. (2013). Orchestrating OGC web services to produce thematic maps in a spatial information infrastructure. *Computers, Environment and Urban Systems 37*, 107-120.
- ✔ Schlichting, K. (2008). Historical GIS: New ways of doing history. *Historical Methods 41*, 191-196.
- ✔ Tait, M. (2005). Implementing geoportals: Applications of distributed GIS. *Computers, Environment and Urban Systems 29*, 33-47.