From historical documents to GIS:

a spatial database

for Medieval fiscal data in Southern Italy

Daniela Carrion (*), Federica Migliaccio (*), Guido Minini (*), Cynthia Zambrano (**)

(*) Department of Civil and Environmental Engineering (DICA)

Politecnico di Milano - Milano (Italy) daniela.carrion@polimi.it

(**) E-VIA S.p.A. - Retelit Group - Milano (Italy)

Abstract

The study presented in this paper addresses a GIS (Geographic Information System) supporting historical research, including tools for geo-referencing data extracted from medieval source documents. Working closely with historians, an effective model for fiscal and census data has been designed to supply them with tools for digitally archiving data from historical source documents in a database structure, which may suggest new criteria for studying and correlating information. The considered data had been collected from the *Quaternus Declaracionum* of the *Regno di Napoli* (Kingdom of Naples) and focuses on an area corresponding to the Principality of Taranto from 1458 – 1460.

Key words: Geographic Information Systems (GIS), Relational Database Management System (RDBMS), GIS and Society, Medieval GIS, Historical maps

Introduction and state-of-the-art

As it is well known, a GIS (Geographic Information System) and its spatial database are powerful tools for archiving and representing data that are intrinsically related to geographic positions or, in other words, are geo-referenced. As such, GISs have become essential in countless fields, ranging from pure geographic and mapping applications to land management and urban development planning as well as facilities and networks design and maintenance to emergency operations optimising.

In the field of historical research, the use of a GIS appears to be consolidated at the international level, e.g. Gregory (2002), Gregory et al. (2007), Henderson and Berman (2003), Berman (2005), Schlichting (2008). Additionally, historical data applications–which is where a relational database is implemented in a GIS environment–may be found in Boonstra, Collenteur and van Elderen (1995), De Moor and Wiedemann (2001), and Fitch and Ruggles (2003). However, GIS applications are thus far not common in the research of Italian historians. In fact, GIS applications for archiving and analysing Italian historical data are a new achievement; GIS methodologies are essentially applied to the cartographic analysis of ancient maps: examples may be found in Balletti, Guerra and Monti (2000), Balletti (2000), Baiocchi and Lelo (2002), Ardissone and Rinaudo (2005), and Balletti (2006). Furthermore, WebGIS applications for Italian ancient maps can be found; one example is the northern Italy cadastral map of Brovelli et al. (2012).

In this paper the design of a GIS database for medieval fiscal data is presented. It was developed and implemented in the framework of an interdisciplinary research study conducted by geomatics experts to support the work of Italian medieval historians using GIS tools. The spatial database has been designed according to the rules of the relational model (Codd 1970) but adheres to the requirements of the historians involved in the project. This GIS allows the medieval fiscal data to be analysed and visualised in a spatial context and projected onto a modern mapping system. This type of GIS approach to the study of data obtained from the period represents a much rarer application of geomatics to medieval humanities.

Medieval data: from the source documents to a GIS

Because data obtained from historical archives and libraries are often related to geographic places (i.e., to locations that can be detected in maps), a GIS approach can be exploited as a useful tool for analysing and mapping data for historical research, provided that the GIS database is modelled according to the purposes of the historians. For example, in this case, not all rules typically applied in computer science can be applied in designing the database and vice-versa. On occasion, new solutions must be devised, which are generally considered non-orthodox in informatics (this issue will be addressed later in the paper).

The research that is described in this paper is based on a GIS project on the medieval data of Southern Italy (mid XV century). The source that has been researched by the historian is the *Quaternus Declaracionum*, which is preserved in the *Diversi della Sommaria* collection of the Italian National Archives in Naples (Italian A.S.N. *Diversi della Sommaria*. II numerazione, reg. 248). The *Quaternus Declaracionum* contains fiscal data for the Principality of Taranto (an area corresponding to a part of present-day Apulia) from 1458 to 1460, see Pizzuto (2009).

In the *Quaternus Declaracionum*, as is typically the case in medieval documents, time is expressed according to the "indiction" system, where dates run from year 1 to year 15 of a certain 15-year indiction. Specifically, in this study, two *Quaterni Declaracionum* (contained in a unique bound volume of 465 sheets of paper) have been considered; they were compiled by the *magister rationarum* Francesco de Agello in the years of the seventh and eight indictions ranging from 1458 to 1460 A.D.. The former *Quaternus* presents a homogeneous and coherent structure because all officials, although operating in different areas of the *Terra d'Otranto* Province, directly reported their activities in the year of the seventh indiction to the *magister rationarum*. The role of the *magister rationarum* was that of an accounting auditor working for the prince. The first part of the second *Quaternus* has a similar structure and contains data referring to the year of the eighth indiction; however, this structure later changes, and the documents are divided into sections with each of them referring to a specific administrative function conducted in different indictions.

These documents represent a rich source of information on the dominion of the Orsini, who ruled the Principality of Taranto, and cover a large period of time. Notes were taken by the administrators of the Principality on all incoming (*introitus*, corresponding to taxes) and outgoing (*exitus*, corresponding to expenses) sums and written in the *Quaternus*. Additionally, another purpose of these documents was to compute and record the amount of payments or goods due to the Church. The fiscal information was complemented by data that can be compared to a present-day census, i.e., recording the number of families or houses ("fires") in the settlements located in the territory of the princes Orsini.

Because a GIS approach was applied to address the medieval data, an important issue arose concerning the procedure with which the medieval place-names were geo-referenced. In the *Quaternus*, the names of the settlements were recorded by the *magister rationarum* with the Latin toponyms by which they were known at the time when the document was created. However, during the study, medievalist historians could match the place-names in the original source with present-day place-names (see Pizzuto 2009) and determine (to a certain degree of uncertainty) their location with respect to a modern mapping system (the official Italian cartography, as described later). The degree of uncertainty can be considered as a measure of accuracy of the positioning of the settlement with respect to modern maps. Essentially, medieval place-names were linked by the historians to names of present-day centers of habitation, such as towns, districts and neighbourhoods or churches and monasteries. Other scholars have recently proposed similar approaches to geo-reference place names, based on the concept of administrative unit ontology and gazeteers, see Southall (2014). The smaller the extension of the present-day sites, the lower the degree of uncertainty that could be presumed for the geo-location of the medieval settlements. Thus, a correct geo-referencing of ancient names and a corresponding measure of accuracy were achieved.

After the present-day place names had been defined, they were geo-referenced by applying a proper map projection, such as the UTM mapping system (Grafarend and Krumm 2006). This process was performed by exploiting the WMS (Web Map Service) geo-services of the Italian *Geoportale Nazionale* (http://www.pcn.minambiente.it/GN/). As a reference, the Italian official cartography produced between the 1940s and the 1980s in the Italian DATUM Roma 1940 by the IGM (*Istituto Geografico Militare*) has been considered. Afterwards, the coordinates were transformed into the widely used UTM WGS84 cartographic system.

In addition to the medieval place names and data on taxation and census, the *Quaternus Declaracionum* provides the names of officials serving in the Principality of Taranto from 1458 to 1460 (e.g., the tax collector for a district or the captain), which provides an opportunity to identify the functions of these historical figures or trace the locations where they could be found in those years.

In summary, the information derived from the historical source studied during the project can be grouped into three main categories: historical (medieval) place names, amount of taxes due to the prince and names of officials serving in the Principality of Taranto from 1458 to 1460.

Spatial database model and the GIS for medieval data

At the beginning of the project, the data collected from the source document *Quaternus Declaracionum* had to be digitized. This work was performed by historians, who stored the data in a "file-based" structure, i.e., an *Excel* spreadsheet. Then, a relational database was designed to manage the spreadsheet digital data in a more convenient way; this required identifying "entities" (physical objects or abstract concepts) that could describe the aspects of the medieval "real world" that had to be represented in the GIS. Lastly, nine entities were defined, four of which had spatial attributes and could thus be geo-referenced. ***Table 1*** and ***Table 2*** contain a brief description of the spatial and non-spatial entities defined for the medieval GIS, respectively.

Table 1 approximately here

Table 2 approximately here

All spatial entities were geo-referenced using East and North coordinates in the reference system UTM WGS84; the points correspond to the original medieval names reported in the *Quaternus Declaracionum*. The positions of the points could be identified to different degrees of accuracy in the database; the levels of accuracy are represented through an ordinal scale of values provided in ***Table 3***.

Table 3 approximately here

The Data Base Management System (DBMS) used for the historical database is PostgreSQL. After an evaluation of the available alternatives, PostgreSQL was selected for two primary reasons. First, the medieval GIS is intended as an Open Source project, thus it must be based on Free/Open Source technology. Secondly, PostgreSQL can be integrated with the PostGIS extension to manage the spatial component of entities through the special datatype GEOMETRY. The PostgreSQL + PostGIS combination provides the possibility of working with a spatial DBMS that allows an interaction between the database and a GIS software (both commercial software, such as ESRI ArcGIS, and Open Source software, such as QGIS).

The conceptual model of the database is provided in the form of an Entity Relationship Diagram (ERD) in ***Figure 1***, where the entities, the relationships between them and the cardinality of the relationships are represented. The medieval database ERD provides a peculiar complexity due to the temporal component; in fact, when this information is considered, the number of possible relationships between the entities increases.

Nearly all the relationships defined in the model have 1:N (one-to-many) cardinality, except for the “Fiscal unit – Place”, “Fiscal district – Source” and “Fiscal unit – Source” relationships, which have N:N (many-to-many) cardinality. Particularly, the relationships between “Present place” and “Place” and “Character” and “Officer” have 1:N cardinality due to the presence of the temporal component (a present-day place may have been mentioned with different toponyms in different ages, and an historical person may have exerted different functions during his career).

Lastly, the logical model of the medieval database was designed, which defined the attribute data for each entity in the database. The logical model (drawn using MySQL Workbench) is presented in ***Figure 2*.**

Figure 1 approximately here

Figure 2 approximately here

The primary characteristic of the medieval database model is the centrality of the entity “Source document”; in fact, it is related to most entities in the database. Generally, in a GIS, information on data acquisition and sources are stored in a metadata section. In the case of an historical GIS, the very information on the original source is to be integrated in the database itself; it can either be regarded as an entity in the relational database or as an attribute of one or more entity. The choice of considering the source as an entity arises from the historians' need to store the data in the original source document and maintain the relationship of that same piece of information to other data. The direct effect of this situation is an increase in data redundancy. The presence of redundancies does not always represent a problem; however, an evaluation of the advantages and disadvantages is necessary. In this case, the primary advantage is the reduction in the number of accesses that, in case of complex queries, are necessary to extract the required piece of information. The primary disadvantage is the increase in memory occupation; however, in the current example, it can be considered as a negligible cost (on the order of a few kilobytes).

Along with the database, a data dictionary has been created to store metadata information on all the attributes defined for the entities of the database. The first (prototype) version of the dictionary had been prepared by the history researchers and accompanied the “file based” table structure that had initially been used to store the data. The final version of the dictionary has been completed after the database implementation; for every entity, the attributes have been defined in a more comprehensive way. ***Table 4*** and ***Table 5*** provide extracts of the data dictionary, which describes the attributes characterising the “Source document” and “Historical place” entities. These entities are fundamental to the database: the former because of its centrality in historical research (as already explained) and the latter because it represents mapping the data in the UTM WGS84 reference system. Relationships have been established among the "Historical place" and other entities stored in the database (as represented in the conceptual and logical models of ***Figure 1*** and ***Figure 2***) so that spatial queries can be performed on the data in the GIS and their results visualised in a mapping environment.

From ***Table 4*** and ***Table 5***, it is possible to see the variety of data that had to be archived to preserve full information of the original source documents and understand the criteria based on which the data have been coded.

Table 4 approximately here

Table 5 approximately here

Data analysis and maps based on the medieval GIS

During the project phases, while the system was being developed, it was possible to adapt the database model to the users' requirements through discussions on the aspects of modelling and data implementation. One particularly important activity was the implementation of a prototype GIS to be exploited and tested by historians to check the functionalities of the system. The experimentation initially was conducted under the supervision of a group of GIS experts; at a later state, historians worked on queries and data analysis autonomously.

As an example, an often used operation turned out to be the selection of places belonging to a given tax district (see the definition in ***Table 5***) and their visualisation onto the UTM mapping system (see ***Figure 3***).

Figure 4 presents another example of a layout obtained after a query on the medieval spatial database regarding the amount paid at every historical place for two types of taxations: *focatico* and *tassa milizie*.

Figure 3 approximately here

Figure 4 approximately here

Even simple functions, such as the computation of distances between points or areas of polygons, have been used to assist in outlining a “quantitative” analysis of the historical data. Consequently, this has given rise to a new request, i.e., to implement the possibility of drawing boundary lines between territories based on information derived from source documents. Certainly, such boundaries will have to be superimposed onto a proper mapping system, and presumably a certain degree of accuracy explaining the level of uncertainty connected to the choice of a particular boundary line should be devised.

Furthermore, the spatial representation of historical places in a mapping environment helped to detect a few rough errors (including positioning errors) that had not been highlighted during the data entry. These types of errors can be avoided by implementing consistency tests to check a variety of situations. This activity is on-going at the moment; a routine for checking consistency has been implemented and will be integrated in the DBMS. The possibility of defining the positions of medieval settlements recorded in the *Quaternus Declaracionum* in a modern reference system has provided the opportunity to produce digital medieval maps, and visualize time and spatial components in the data. Typically, the spatial component is kept separated from the time component in mapping applications.

Lastly, fiscal data have easily allowed valuable information to be derived on which activities, such as an inn or tavern (*taberna)*, were common at given locations. It was also possible to draw maps of locations in which a certain official was present in different epochs or locations in which different officials were present in a given epoch. In other words, historical figures could be "traced", and their spatial distribution could be visualised on maps.

Discussion and conclusions

In this paper, a GIS for historical research has been presented. Beginning with medieval data extracted from documents dating back to the XV century (stored in Italian National Archives), census and fiscal data had been initially recorded in a "file-based" digital structure, which was not suitable for spatial or complex queries. Therefore, a relational spatial database structure has been designed and implemented to help historians analyse and visualize data from a mapping point of view. Accordingly, positions in a modern mapping and reference system have been associated to the entity “Historical place”, thus allowing not only such places (found in the source documents) to be mapped but also various types of analyses and queries to be performed based on related entities from a spatial point of view. The positions of the historical place names were defined based on the corresponding present-day names that could be associated to medieval names (Pizzuto 2009). The spatial accuracy of the derived coordinates, lacking any other criterion to define it, was determined specifying an ad-hoc ordinal scale of values ranging from 1 (very low accuracy) to 5 (very high accuracy).

Tests were performed on the implemented spatial database and the GIS. The goal of the testing activities is twofold. On one hand, the GIS represents a new environment for the historians’ research, and they need to be acquainted with the new technology; on the other hand, their feedback on the usability of the implemented mapping and query tools is fundamental alongside with their comments on the structure of the database and definition of the attributes. This will greatly help improve the GIS and integrate new functionalities if needed.

A few advantages of the implemented system are already evident. First, the dynamic nature of a GIS is a great improvement with respect to traditional maps. Particularly, for information dating back to the medieval period, historians are extremely careful in assigning geographic positions to historical places and even more careful in assigning geographic boundaries (Cengarle and Somaini 2009). The GIS provides the possibility to display different hypotheses at the same time, e.g., with different layers of data, keeping track of different theories and their authors.

The implemented GIS allows demonstration of the spatial distribution of the places cited in the sources and performance of quantitative analyses (e.g., see ***Figure 4***). Moreover, in the framework of the cooperation with the historians, other sources (e.g., *Liber Focorum Regni Neapolis*, Berio Library in Genoa, produced by the financial officers of Alfonso V of Aragon in the 1440s) are being considered and integrated in the DBMS, thus allowing the visualisation and querying of multiple sources simultaneously.

The choice of the base map for a medieval GIS is not trivial. As previously mentioned in the paper a modern map (official IGM Italian cartography produced during XX Century) has been considered as a reference to locate medieval places. Nevertheless, the GIS allows considering other base maps of different epochs. Some work has already been performed considering older maps, even a medieval map, and interesting results have been obtained in terms of geo-referencing possibilities (Carrion et al. 2008).

Some important issues are still under discussion and must be defined in the near future. In the implemented database only point positions have been defined: is there any other sensible way (from the historian's point of view) to define polygon features based on information extracted from the source documents? This will inevitably imply the need of introducing a suitable index for the uncertainty in the definition of polygon boundaries.

In addition to dealing with such methodological questions, studies are being conducted to ingest data from other source documents, implement tools for consistency testing of the data and provide historians with a WebGIS so that different research groups can share data and maps through the Internet.

Moreover, it can be studied if positions may be better associated with entities defined in the GIS database (i.e., how to associate spatial information to the data). When working in an archaeological project, locations (e.g., of buildings, artefacts, etc.) can be easily collected by GPS or other types of surveying instruments. When addressing information from historical documents, positions cannot be “measured”; they must be “deduced” starting from the information recorded in the document and based on hypotheses involving an “interpretation” of the content of the historical document. In this stage of research, positions for medieval census and fiscal data have been defined by locating present-day place names corresponding to historical place names; for the future steps better solutions could be envisaged.

Acknowledgments

The project has been partly funded by the Italian Ministry of Education, University and Research (MIUR Prin 2009 project 2009YCXKZH_003).

The authors wish to thank the History researchers with whom they have cooperated throughout the project, particularly Prof. Francesco Somaini, Dr. Federica Cengarle, Dr. Francesca De Pinto, Dr. Simona Pizzuto.

References

Ardissone, P., and Rinaudo, F. 2005. A GIS for the Management of Historical and Archaeological Data. In *Proceeding of the CIPA 2005 XX International Symposium*: Torino.

A.S.N. *Diversi della Sommaria*. II numerazione, reg. 248.

Baiocchi, V., and Lelo, K. 2002. Confronto di cartografie storiche con cartografie attuali per l'area del centro storico di Roma. In *Atti della VI Conferenza Nazionale ASITA*: Perugia.

Balletti, C. 2000. Analytical and Quantitative Methods for the Analysis of the Geometrical Content of Historical Cartography. In *International Archives of Photogrammetry and Remote Sensing, vol. XXXIII, part B5*: Amsterdam .

Balletti, C. 2006. Digital elaborations for cartographic reconstruction: the territorial transformations of Venice Harbours in historical maps. *e-Perimetron, Autumn 2006, 1, 4*: 274-286.

Balletti, C., Guerra, F., and Monti, C. 2000. Analytical Methods and New Technologies for Geometrical Analysis and Georeferenced Visualization of Historical Maps. In I*nternational Archives of Photogrammetry and Remote Sensing, vol. XXXII, part 6W8/1, ISPRS WG VI/3 and IV/3, Bridging the gap*: Lubiana.

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.

Brovelli, M. A., Minghini, M., Giori, G., and 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., Cengarle, F. & Somaini, F. 2008. Un approccio alla georeferenziazione di una carta antica, *Atti della 12a conferenza nazionale asita*, L'Aquila, Italy 21 - 24 Ottobre 2008, 655-660

Cengarle, F. and Somaini, F. 2009. Mappe informatiche e storia. In *I domini del principe di Taranto in età orsiniana (1399-1463)*, Somaini F and Vetere B eds, Congedo Editore: 3-35.

Codd, E. F. 1970. A relational model of data for large shared data banks. *Comm. of the ACM, Vol. 13 (6)*: 377-387.

De Moor, M., and Wiedemann, T. 2001. Reconstructing territorial units and hierarchies: a belgian example. *History & Computing 13*: 71-97.

Fitch, C. A., and Ruggles, S. 2003. Building the National Historical Geographic Information System. *Historical Methods: A Journal of Quantitative and Interdisciplinary History 36*: 41-51.

Grafarend, E. W., and Krumm F. W. 2006. *Map projections. Cartographic Information Systems*: Springer, Berlin-Heidelberg-New York.

Gregory, I. N. 2002. *A Place in History: A Guide to Using GIS in Historical Research*: Oxbow, Oxford

Gregory, I., and Healey, R. 2007. Historical GIS: Structuring, mapping and analysing geographies of the past. *Progress in Human Geography 31*: 638-653.

Henderson, M., and Berman, L. M. 2003. Methods for Space-Time Analysis: Examples form CHGIS. In *ESRI Users Conference*: San Diego.

Liber Focorum Regni Neapolis, Berio Library in Genoa

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.

Schlichting, K. 2008. Historical GIS: New Ways of Doing History. *Historical Methods: A Journal of Quantitative and Interdisciplinary History 41*: 191-196.

Southall, H. 2014. Rebuilding the Great Britain Historical GIS, Part 3: Integrating Qualitative Content for a Sense of Place. *Historical Methods: A Journal of Quantitative and Interdisciplinary History 47*: 31-44.

Web sites

Italian National Geoportal, “Geoportale Nazionale”. WWW document, http://www.pcn.minambiente.it/GN/

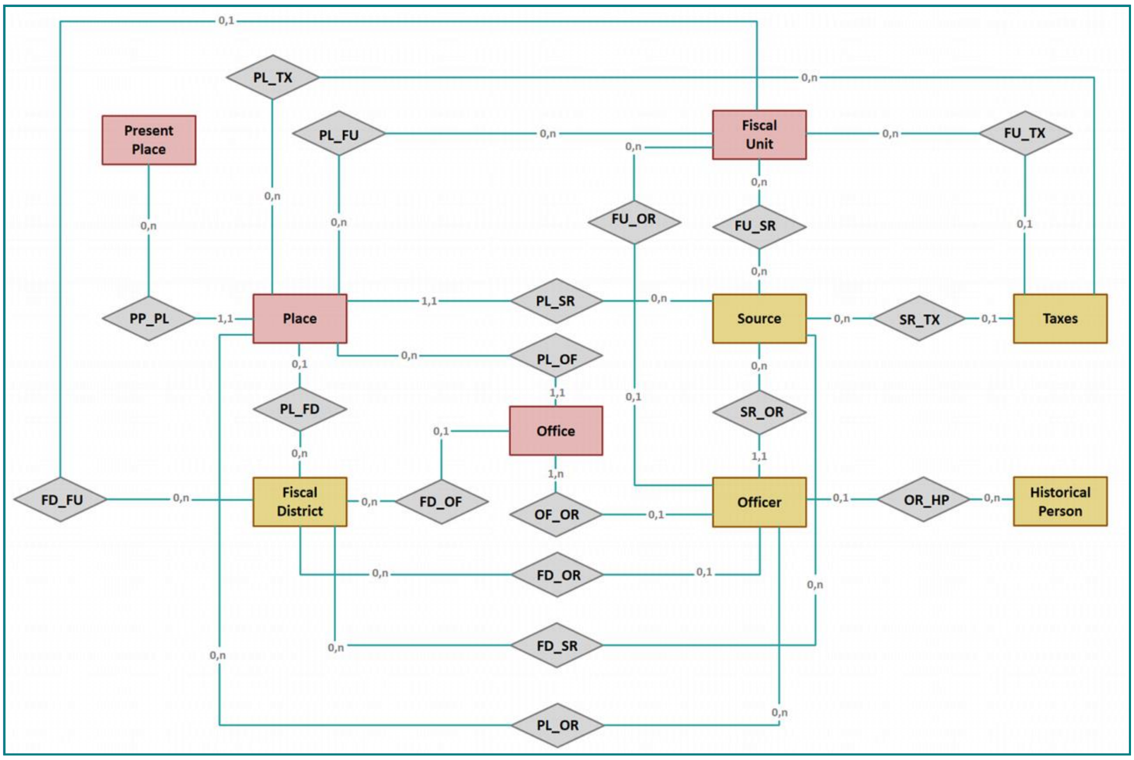

Figure 1: *ERD of the Medieval database; spatial entities are shown in red; non-spatial entities are shown in yellow; and junction tables (to manage the many-to-many relationships between entities) are shown in grey*

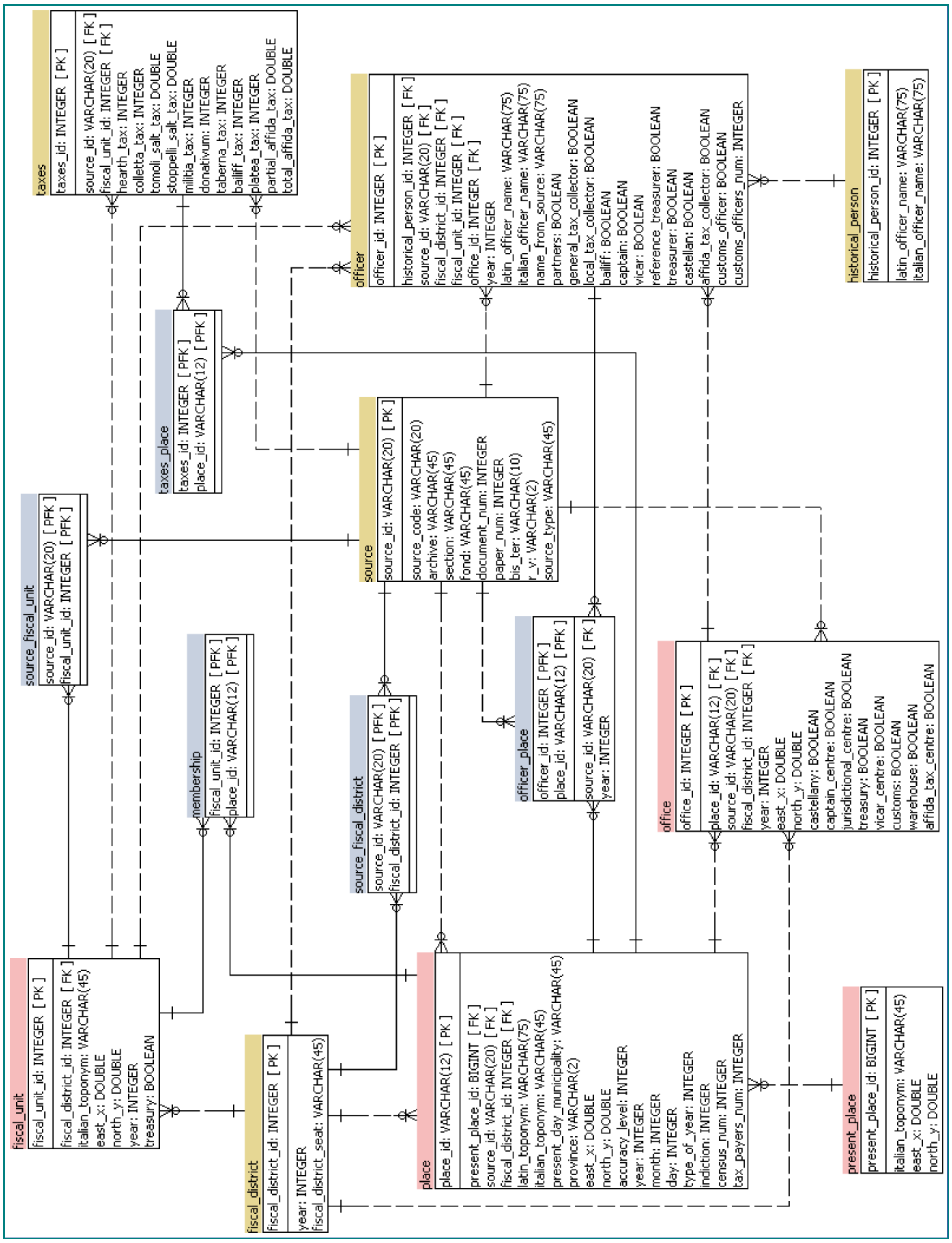

Figure 2: *Logical model of the medieval database for the data extracted from the Quaternus Declaracionum; spatial entities are shown in red; non-spatial*

entities are shown in yellow; and junction tables (to manage the many-to-many relationships between entities) are shown in grey

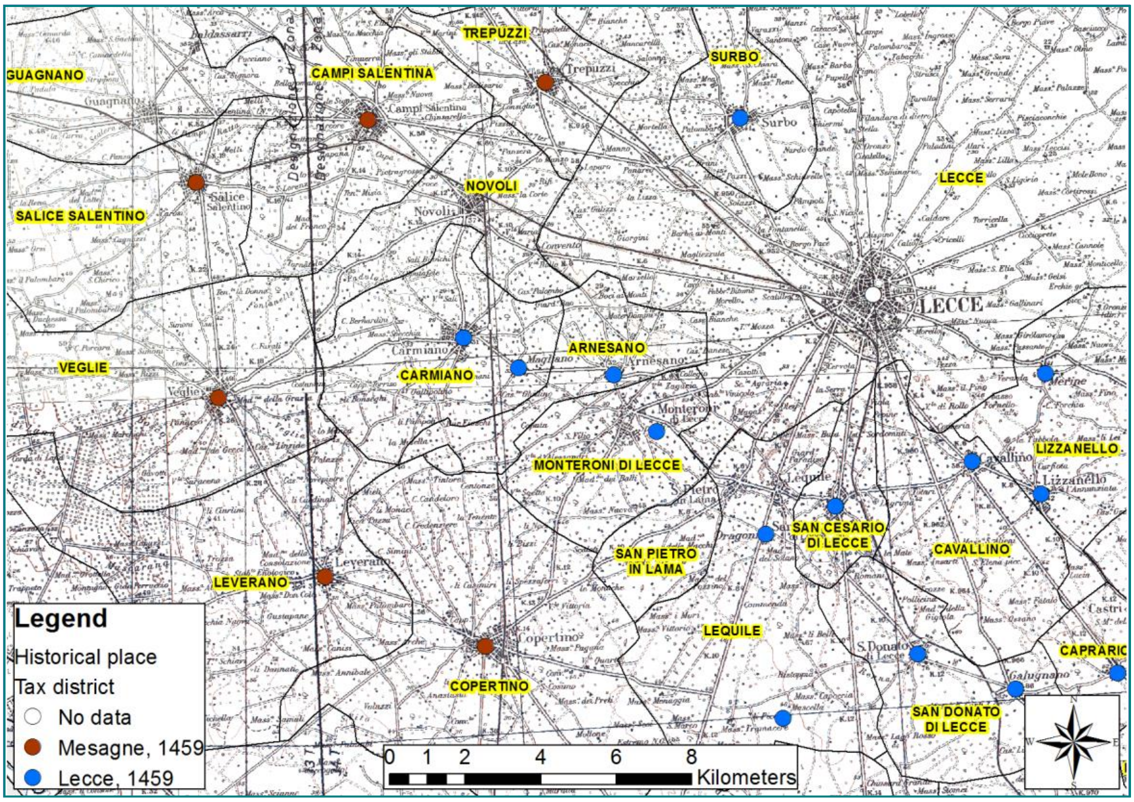

Figure 3: *Selection of places belonging to a given tax district and their mapping on the UTM system*

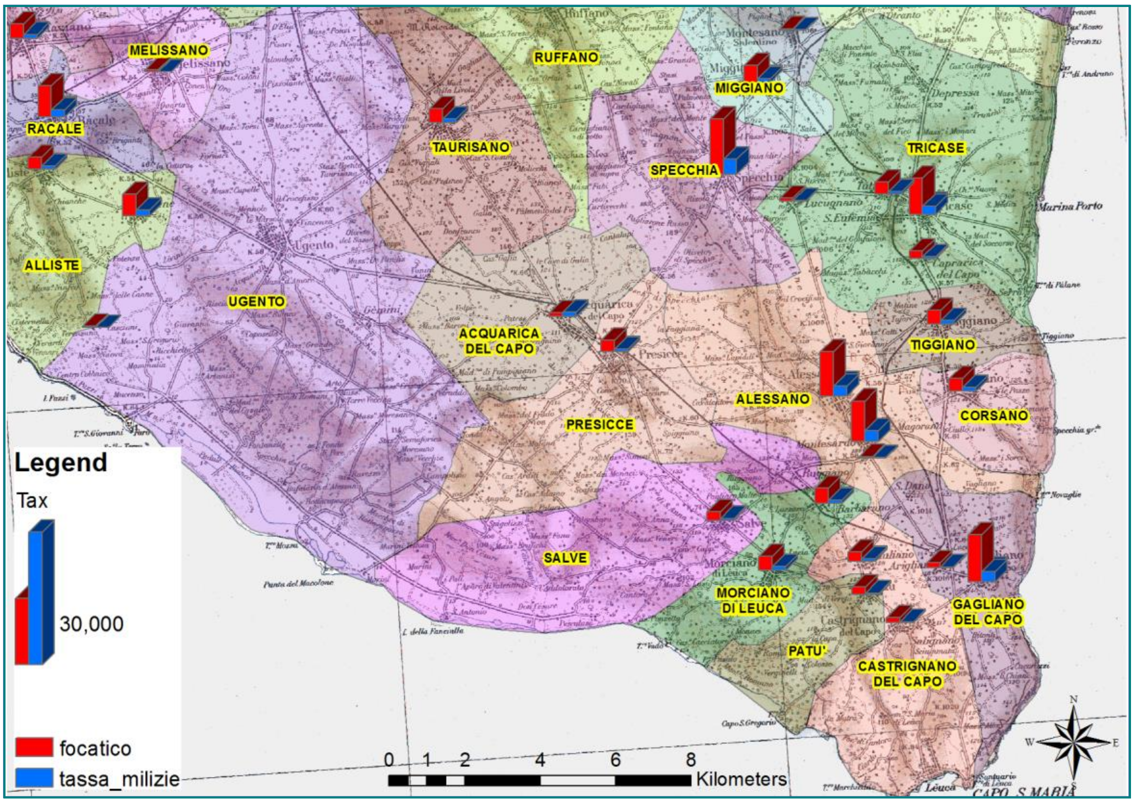

Figure 4: *Amount paid at every historical place for two types of taxations: "focatico" ("fiscal fires-tax") and "tassa milizie" ("militia-tax") based on the data extracted from Quaternus Declaracionum*

Table 1: *List of "spatial" entities in the GIS and their short description*

ENTITY	***DESCRIPTION***
Place	Collects data about historical places as recorded in the source
Present place	Collects data related to the present-day toponym corresponding to the medieval village recorded in the historical document
Fiscal unit	Centre from which a toponym depended from the fiscal point of view
Office	Administrative office that receives taxes

Table 2: *List of "spatial" entities in the GIS and their short description*

ENTITY	***DESCRIPTION***
Fiscal district	A cluster of fiscal units
Source	Historical document from which information is obtained
Taxes	Collects information about typology and amount of taxes
Officer	Officer found in historical documents, exerting one or more functions; a historical person may appear several times in the records of this entity if during his lifetime he has exerted more than one function or if he has exerted the same function for different years.
Historical person	Collects data about confirmed historical persons; this gives the possibility of joining data with other databases that archive similar information.

Table 3: *Ordinal scale of values defining the accuracy level of positioning medieval toponyms*

Location of toponyms	***Accuracy level***	***Index value***
Toponyms with poor information	Very low	1
Uncertain toponyms	Low	2
Towns	Medium	3
Small built-up areas	High	4
Isolated buildings	Very high	5

***Table 4**: Data dictionary entries for the entity "Source document"; primary key in Italic, underlined; the English attribute name has been defined based on http://www2.archivists.org/glossary*

Entity: Source document			
Attribute name (Italian)	***Attribute name (English)***	***Definition***	***Value***
id_fonte	***source_id***	Identification code of the historical source document from which the information has been derived	Alphanumeric code
cod_fonte	**source_code**	Alphanumeric code of the historical source document	Alphanumeric code, composed by: archive, fond, piece number, paper number, recto/verso. As an example: ASN, DS II, 248, 112v ASN = Archivio di Stato di Napoli, DS II = Diversi Sommaria II, 248 = Registro 248, 112 = Carta 112, v = verso
archivio	**archive**	Full name of the archive where the historical source document is preserved	Alphabetic string As an example: Archivio di Stato di Napoli
sezione	**section**	In case it exists, full name of the section of the archive where the historical source document is preserved	Alphabetic string As an example: Sezione di Archivio di Stato di Assisi
fondo	**fond**	Full name of the fond where the historical source document is preserved	Alphabetic string As an example: Diversi Sommaria II numerazione
num_pezzo	**document_num**	Number of the document	Integer number referring to the document piece
num_carta	**paper_num**	Number of the paper or folder where the historical source document is preserved	Integer number
bis_ter	**bis_ter**	Additional information regarding the numbering of the paper	Alphabetic code, such as bis, ter, quater, etc.
r_v	**r_v**	Paper side on which the information is written	Alphabetic code: r = recto v = verso rv = both recto and verso
tipo_fonte	**source_type**	Typology of the historical source of information	Alphabetic string

***Table 5**: Data dictionary entries for the entity "Historical place"; primary key in Italic, underlined; foreign keys in Italic; the English attribute name has been defined based on http://www2.archivists.org/glossary*

Entity: Historical place			
Attribute name (Italian)	***Attribute name (English)***	***Definition***	***Value***
<u>id_luogo</u>	***<u>place_id</u>***	Identification code of the place corresponding to the latin toponym	Alphanumeric code
id_luogo_attuale	***present_place_id***	Identification code of the present-day place corresponding to the Italian toponym	Numeric code
id_fonte	***source_id***	Identification code of the historical source document from which the information has been derived	Alphanumeric code
cod_fonte	**source_code**	Alphanumeric code of the historical source document	Alphanumeric code, composed by: archive, fond, piece number, paper number, recto/verso. As an example: ASN, DS II, 248, 112v ASN = Archivio di Stato di Napoli, DS II = Diversi Sommaria II, 248 = Registro 248, 112 = Carta 112, v = verso
id_distr_erariale	***tax_district_id***	Code of the tax district to which the place belonged	Five-digit numerical code. The first digit is an integer number between 1 and 4: 1. From Taranto to Lecce (seat of the tax office: MESAGNE); 2. From Lecce to Leuca (seat of the tax office: LECCE); 3. District of Soleto (seat of the tax office: SOLETO); 4. District of Bari (seat of the tax office: BARI). The last four digits represent the year cited in the source document.
toponimo_lat	**latin_toponym**	Latin name of the toponym, as reported in the source document.	Alphabetic string
toponimo_ita	**italian_toponym**	Modern name of the toponym where the medieval place can be located. Coordinates are connected to this attribute.	Alphabetic string
comune_attuale	**present_day_mu**	Name of the Municipality in	Alphabetic string

	nicipality	which the toponym can be located	
provincia	**province**	Acronym of the Province to which the Municipality belongs.	A few upper-case letters
est_x	**East_x**	East coordinate in [m] in the UTM map projection system, WGS84 reference system (zone 33)	Numeric value [m]
nord_y	**North_y**	North coordinate in [m] in the UTM map projection system, WGS84 reference system (zone 33)	Numeric value [m]
liv_affidabilita	**accuracy_level**	Accuracy level which can be associated to the geo-referencing (positioning) of the toponym	Numerical value in an ordinal scale from 1 to 5: 1: Toponyms with very few information: very low degree of accuracy; 2: Uncertain toponyms: low degree of accuracy; 3: Towns: medium degree of accuracy; 4: Small built-up areas: high degree of accuracy; 5 Isolated buildings: very high degree of accuracy.
anno	**year**	Solar year reported in the source document In case of indiction, the solar year is considered, across which the longest portion of the indiction extends	Integer numeric value (four digits)
mese	**month**	Month reported in the source document	Two-digit number
giorno	**day**	Day reported in the source document	Two-digit number
tipo_anno	**type_of_year**	Type of year calculation	One-digit numeric code 0 = Solar year (Jan. 1 – Dec. 31); 1 = Indiction year (Sept. 1 – Aug. 31 of the following solar year). All the different types of solar and indiction years should be considered.
indizione	**indiction**	Indiction number. It defines the specific year within a 15-years indiction period.	Integer numerical value between 1 and 15.
fuochi_fiscali	**census_num**	Number of families corresponding to the population census for the place (the so-called “fuochi fiscali” or “fiscal fires”)	Integer numerical value

fuochi_paganti	**tax_payers_num**	Number of families actually paying a direct tax to the prince (the so-called “fuochi paganti” or “paying fires”)	Integer numerical value