Architectural Heritage Semantic Data Managing and Sharing in GIS

Elena Cerutti\(^1\), Francesca Noardo\(^2\) and Antonia Spanò\(^1\)

\(^1\)Politecnico di Torino, DAD, viale Mattioli, 39, 10125 Torino, Italy
\(^2\)Politecnicodi Torino, DIATI, c.so Duca degli Abruzzi, 24, 10129 Torino, Italy

\{elena.cerutti, francesca.noardo, antonia.spano\}@polito.it

Keywords: Semantics, Standard, Images, GIS, Interoperability, PostGIS, Cultural Heritage, Architectural Heritage.

Abstract: GIS can be effective instruments for managing Architectural Heritage data, in order to query the data for preservation purposes and to realize advanced analysis. These capabilities can be improved using some tools developed by the fields of informatics and internet services such as standards, ontologies and object-oriented programming. The official standards (languages and models) permit the encoding of data so that they can be effectively shared and integrated, concurrent with the knowledge and integration of data in Cultural Heritage (CH). Moreover, an even better interoperability of data can be achieved using open-source management software that normally features more standard data formats and can be used by everyone. These tools have been used in the research presented here for managing different kinds of data (spatial, non-spatial, images) on different views, in a unique database respecting the standards codes. In this way some schemas have been defined, and they can be exported to reach effective data interoperability.

1 INTRODUCTION

The well-known abilities regarding multi-format, multi-scale and multi-temporal data management are essential GIS tools in Cultural Heritage digital archiving. A number of projects have been developed in order to prove how GIS tools enhance storage, analysis, and data processing (Apollonio et al., 2012; Petrescu, 2007). They have become important support for any kind of knowledge and planning phase in the field of CH Preservation and Protection. Central and local authorities responsible for Cultural Heritage experienced in creating complex and integrated GIS and/or WEB-GIS have recognized these tools as useful aids in the decision-making phase at different scales (Taboroff, 2000). The success has mainly been rooted in the archaeological field, because of the intrinsic spatial connotation of archaeological data (Wüst, 2004).

Some systems with different purposes have been well-established for at least fifteen years; Djindjian (1998) has detected the main archaeological sectors related to GIS use: archaeological surveys (prevision of sites location), spatial analyses for territorial inquiries, Cultural Resource Management for CH Protection and Preservation, and lastly, intra-site GIS.

In the framework of Architectural Heritage protection, the GIS services requested are very similar to those well-established for intra-site GIS. The similarities reside in the scale of details needed in both systems, and the large amount of archived multifaceted heterogeneous data. Furthermore, the use of GIS is suitable for processing complex and specialized geometrical entities, allowing them to be manageable in a 3D spatial context. These are increasing in number, since in recent years spatial objects have often been derived from LiDAR or photogrammetry methods of points model generation. For all these reasons, specialized semantic values of database objects are needed.

The management of object meanings is being developed by other sectors, such as web technologies. However, some necessary infrastructure in order to easily implement systems able to manage this information is missing: the available ontologies are often incomplete for the overall management of some kind of CH item, and the software currently used is not always the best solution for the implementation of the models.

The convenience of using the building information modelling (BIM) to manage high-scale semantic representations has been tested. In this paper we are going to discuss some issues concerning the ways in which these needs can be addressed in GIS, as listed in the framework of Architectural Heritage.
1.1 Ontologies, Semantic Representations and Interoperability Issues

Semantics is the study of meanings, and focuses on the relation between the signifiers - the symbols used to communicate a concept - and the meaning of the concept itself. An essential element of this discipline is the study of language. For the exigencies of web communications and automatic computing, the study of formal languages for expressing concepts and for relating them to one another has been developed. This development carried over to the advancement of ontologies; these can help to resolve heterogeneities, as they define a unique frame by making the conceptualisation unambiguous (Guarino, 2009).

The superabundance of data and its misinterpretation is a real problem in the CH field. These issues can lead to the risk of carrying out incorrect interventions and, consequently, of losing some valuable CH items (tangible or intangible). The modelling concepts of information systems based on domain ontologies (Guizzardi, 2005) can effectively reduce this risk.

The use of semantics theory to achieve interoperability and data sharing has been examined in recent years by the developers of the Semantic Web, the evolution of the World Wide Web, in which the meaning of data is managed through its semantic contents. This development is headed by W3C (World Wide Web Consortium, www.w3c.org), an organization that publishes explicit standards. W3C defined some useful languages for representing information that are both human and machine-readable. Markup languages (such as HTML and XML) allow one to write content and provide information about which role that content plays. In particular, XML (www.w3.org/XML/) is a metalanguage for markup: it provides a uniform framework, and tools for the interchange of data and metadata among applications. XML does not provide any means of talking about the semantics (meaning) of data. Many software applications use XML for exporting and exchanging files, but these files cannot always be read correctly by different software programs. Even so, XML is the base for several formal languages that are able to define and to express semantics in a machine-readable format: RDF (Resource Description Framework), and OWL (Ontology Web Language). These can structure the semantics of data effectively, and can be queried using query languages like SPARQL.

In spite of these advantages, these languages do not consider the spatial dimension of data. The management of spatial information is instead the primary objective of the OGC (Open Geospatial Consortium), which has spearheaded several efforts at defining standards for reaching interoperability solutions that “geo-enable” the Web. The Mapping sector also requires more and more data integration and exchange. The international directives, such as INSPIRE (INfrastructure for SPatial InfoRMation in Europe) (http://inspire.ec.europa.eu/), promote the development of spatial data infrastructures (SDI) that can rely on the availability of spatial data standards.

The OGC defined standards explicitly for this kind of data. One of the basic OGC standards is the GML (Geographic Markup Language), which is similar to the XML in structure is intended to express geographical objects. This is used for the definition of the standard CityGML, which is a model for the representation of city objects (http://www.opengeospatial.org/standards/citygml).

The same language is used for the INSPIRE UML model, aimed at the harmonisation of digital maps in Europe. Another important standard language is the OGC geoSPARQL, which is useful for managing spatial data (http://www.opengeospatial.org/standards/geosparql).

Some projects have been developed for the integration of spatial information, in OWL ontologies using the OGC geoSPARQL. See as an example for Cultural Heritage the project CRMeo (Doerr and Hiebel, 2013).

The successes of standards adoption and sharing arose thanks to ICT (Information and Communication Technology) support. For further enhancing interoperability, open source tools are nowadays increasingly of interest in different environments, including CH. The Open Source Geospatial Foundation collects the most popular Open Source GIS projects, and they refer to standards of the Open Geospatial Consortium (OGC). Examples of GIS tools employed in the CH framework are: GRASS-GIS (Geographic Resources Analysis Support System), QuantumGIS, with a user-friendly graphical interface, SAGA GIS, System for Automated Geoscientific Analyses, MapWindow for modelling and analysis, and ILWIS GIS (Integrated Land and Water Information System) with image analysis and photogrammetric functions.

1.2 Architectural Framework Needs and Standards Availability

The Historical Architectural Heritage is subject to continuous use, and, over time, to maintenance, repair and restoration. As such, it is essential to be able to document and record dynamic...
translating in order to allow for continuous updating and monitoring. A multiplicity of heterogeneous information must be stored according to the relation of specific parts of cultural artefacts. Concurrent with these archiving requirements, the need for data meanings management arises; meanings are needed in order to unambiguously interpret, share and correctly exchange information. It is obvious that the implementation of accessible systems using a supported standard for the management of CH information should be encouraged.

Some studies about the development of some semantic GIS have been performed, beginning in the mid-1990s (Mennis, 2003; Fonseca et al., 2002). In these studies, an object-oriented approach was used as an effective solution for expressing and storing the data meanings (Scholl, 1992). In this way, even more powerful systems could be built with significant data interoperability and a reduction of any potential ambiguity.

The international CH institutions, including UNESCO, ICOMOS, ICOM-CIDOC, CIPA and the Getty Conservation Institute, as well as the ISCR (Istituto Superiore per la Conservazione e il Restauro) in Italy, have made significant efforts in developing guidelines, recommendations, ontologies and structured vocabularies to construct this complex framework. It is possible to identify some effective standards applied to CH that have been tested with good results. For many subsectors of the Cultural Heritage field, such as museum archives and photo inventories, some standard data models have been defined in spite of continuing uncertainty regarding a more general framework. For example, MIDAS Heritage, developed by the Historic Buildings and Monuments Commission for England, is a free data standard available for recording information about CH. It adopts INSCRIPTION, a collection of wordlists for monument classification. Category for the Description of Works of Arts (CDWA), defined by the Getty Research Institute, describes the content and format for the records of art databases, including architecture. In the Italian framework, some efforts have been made by the Commissione NorMaL (NORMalizzazione Materiali Lapidei) in the field of monument restoration to define unified methodologies and specifications for materials preservation. These guidelines are to become standard UNI (Italian) and aim to be recognized in Europe (NorMaL, 2006).

Moreover, the CIDOC (International Committee for Documentation) of the ICOM (International Council of Monuments) has defined what is considered the core ontology for Cultural Heritage: CIDOC – CRM (Conceptual Reference Model) (Doerr et al., 2007), which became the standard ISO 21127. It is defined using OWL, and is a formal ontology for exchanging cultural heritage information and enabling semantic inferences. An enhancement of CIDOC-CRM is MONDIS (Monument Damage Information System) (Blaško et al., 2012; Cacciotti et al., 2013), the ontology developed for specializing the CIDOC CRM in the field of preservation, restoration and intervention. It also uses OWL for the encoding of the schema.

In recent years, the institutions cited, as along with the World Monument Fund, have risen to the challenge of filling in many “CH inventories” - essentially, Cultural Resource Management - in order to improve the effectiveness of Heritage Protection. Such inventories are valuable for public administration action plans, for specialized research, for tourist attractions, and generally to promote cultural awareness. An example of substantial GIS support aimed to effectively exploit heritage inventories is the ARCHES project, which invites specialized users to upload data regarding any assets across the globe (http://archesproject.org, Myers et al., 2013).

However, the assets’ data are located on maps using satellite or small-scale vector web maps; these kinds of representation are useful for representing partial or whole regions. These maps are usually not adequate for representing the assets with the appropriate level of detail, or from the necessary point of view. Moreover, these standard models currently lack integration between spatial and attribute data for many fields of application.

1.3 Proposal Aims

In this research effort, we modelled a GIS structure by integrating different spatial and thematic standard data-models in order to effectively represent the information in a chosen case study.

The main purposes of our work include the representation of different aspects and points of view, the use of various levels of detail concerning built structures, and the addition of relations with the landscape context. These aims have been achieved through the use of different spatial maps, so that digital regional maps and orthophotos representing the building fronts, or vector graphic drawings representing features of facades, can be visualized separately while remaining stored in a unique geodatabase.

Our intention is to provide a system in which information concerning the architectural elements and their measurable morphology, material decay,
eventual repair interventions and non-routine maintenance works can be related with the purpose of better coordinating the planning phases of restoration and monitoring activities.

Moreover, the ontological models used follow an object-oriented structure, which offer several advantages in representation (such as the possibility of managing inheritance or polymorphism, essential characteristics for a multi-scale approach). This is the reason why we choose the open-source software PostgreSQL, an object-relational database management system (ORDMBS), with a spatial extension (PostGIS).

2 SYSTEMS ENHANCING CH KNOWLEDGE AND PRESERVATION PLANS

The interconnection and cooperation of experts and authorities responsible for CH is a key feature of CH preservation. For this reason, the increase in availability of systems able to advance stakeholders’ interactions is significant (Rodríguez et al., 2014). The goal is the sharing of effort and resources to address planned actions and financing decisions, so as to encourage stakeholders to become actively engaged in preservation processes.

When dealing with CH items one must face problems of heterogeneity of data from multiple points of view. The relevant fields of study are numerous; the formats of data may be different (vector data, raster data, alphanumeric data tables, text documents, and so on), and different sources can provide very different data. At different times in the history of CH, situations can change. Moreover, it is important to represent these objects at different levels of detail for a multi-scale approach; similarly, different surfaces of the same object that are not necessarily coplanar must be depicted.

To fulfil these requirements successfully, a conceptual model derived from multiple self-integrated standard data models has been chosen. Particular care has been taken to choose the software in which this model is implemented.

2.1 The Spatial Object-relational Database Modelled on an Integrated Standard

Existing standards for data models correspond to specific fields of application, and it is unusual to find a comprehensive model appropriate for all the features of multidisciplinary and multifaceted fields like Cultural Heritage. For this reason it is necessary to integrate different data model standards in a unique conceptual model that suitably represents the object of study in an exhaustive way from the perspective of the interested party.

In the example presented, we extracted most of our spatial entities from CityGML. We then integrate them with some entities derived from CIDOC-CRM, which is essential for representing CH entities (even if the spatial contents must often be integrated). Moreover, it can be useful for expressing details at a higher representation scale. For example E26_Physical_Feature can be effectively used for mapping the physically damaged areas on a surface. Doing so is not necessarily straightforward - “This class comprises identifiable features that are physically attached to particular physical objects, but there are no natural borders that separate them completely in an objective way from the carrier objects” (Le Boeuf et al., 2013). As these kinds of details are inherently spatial features, they have been regarded as such by archiving them in the database with a geometry attribute. The third key application field concerning the study object is the preservation field, modelled by the MONDIS ontology. The entities extracted from this last one and related to CIDOC-CRM “E26 Phisical Feature” are “Material”, “Manifestation of Damage” and the related “Intervention”. The conceptual model obtained includes only few entities (Fig. 1), but they are useful for showing a new approach in modelling and implementing the GIS.

The standards schemas and conceptual models must be expressly software-independent, and there are no systems that manage them completely.

However, different logic data models are useful for different data management requirements. The standard schemas could take advantages from the characteristics of object-oriented database management systems (OODBMS), which manage properties typical of object-oriented programming (inheritance, polymorphism, identity).
These could be very meaningful constructs for managing some aspects of both CH information and cartographic objects (Worboys, 2004). Some OODBMS (Object-Oriented DataBase Management System) exist (e.g. EyeDB www.eyedb.org), but they are not specifically designed to manage spatial data. Most of DBMS and GIS management software systems follow the relational logic model, so that they cannot manage the useful characteristics of Object-Oriented systems. Some OOGIS were implemented in the past (e.g. the project GODOT, Gaede et al., 1994), or O2, (Scholl et al., 1992), but today the most widespread GIS management software packages (neither commercial nor open-source) offer these functions. The more advanced systems use a hybrid object-relational model (ORDBMS) - a relational model that includes some functionalities of the OODBMS, such as inheritance and polymorphism. The most widely spread programs are the commercial software Oracle and the open source software PostgreSQL. Both enable spatial feature management (through the applications “Oracle Spatial” and “PostGIS”, respectively).

For the choice of software we considered the scenario offered by the use of open source tools to store and retrieve information, which is closely connected to standard issues. Commercial software packages often use their own formats for storing and exchanging data, which is a limitation. Standard application and the storage of CH data in open source tools certainly foster interoperability and the exchange of knowledge between different specialists involved in CH preservation. For this reason, we chose PostgreSQL – PostGIS for our case study. It is based on Structured Query Language (SQL) and presents advantages, handling large volumes of data and having effective spatial support. Moreover, the application is based on a client-server system; as such, the data can be managed in a centralized way (server) and are accessible to multiple users (clients).

The case study has thus been represented on the basis of the model built, using PostgreSQL as the main repository for the system. Since this system does not have its own graphical interfaces, pgAdminIII has been used for the management of the data table; the open source software QGIS has been employed for the editing and visualisation of spatial data. (Spanò et al., 2014)

At first, the data tables corresponding to each entity identified were created; subsequently, useful attributes and the settings of mutual relations as foreign keys constraints, topological constraints or inheritance (including multiple inheritance) among classes, have been added.

The spatial extension of PostGIS permits the addition of attributes with data type “geometry”, enabling the recording of geometric information. An advantage of this system (impossible in the management of spatial entities with relational GIS, such as ESRI ArcGIS) is the possibility of adding more than one geometric column. This permits the representation of entities with characteristics of polymorphism. Indeed, the same entities can be used in different formats (for example, as polygons and as lines), enabling a multi-scale representation. (For example, a building can be represented as a point or as a polygon in a minor or a major representation scale, respectively.) Another case is the representation of the same object according to different data sources (like data detected from different historical cadastral maps).

In all these cases, the object never loses its identity, by remaining a single record of a table with different representations. Furthermore, the same object can be represented from different points of view, in different reference systems. This necessity arose as the GIS are very useful for managing mainly 2,5D data, and for mappings on surfaces. This is a very valuable tool for many fields of application: functionality can be effectively exploited by maintaining the unity of an object, and more links to different reference systems can be achieved.

In the example we constructed, a geographically referenced map represents the building investigated in its blueprint projection (Fig. 2). Here, the GML entities “LandUse” and “AbstractBuilding” are stored; they are useful for embedding the building studied in its territorial context. For choosing the values of some classification attributes we always prefer affirmed taxonomies. For example, in this case, the values of “LandUse” have been chosen according to the HILUCS classification (Hierarchical INSPIRE Land Use Classification System) given by the INSPIRE European Directive (inspire.ec.europa.eu/codelist/HILUCSValue/).

On this map, a line (in red in the figure) distinguishes the projection of one façade (part of the same entity, “AbstractBuilding”). This line can be directly queried and, through a clearly defined action, another QGIS interface visualising the façade map can be opened. This could be done with multiple views of internal rooms or external façades, so that the surfaces of an architectural object could be mapped on the whole object using different reference systems; these would treat each surface as an independent plane while obtaining all the similar data stored in the same central table and making the
entire structure searchable. For example, it is possible to have statistics about the whole area (on all the surfaces of the building) that requires some kind of preservation intervention, in order to compute the overall cost of necessary preservation actions. These structures remain completely interoperable with the PostgreSQL system at all times. This interoperability could allow for queries on objects viewed in both the interfaces at the same time.

Figure 2: GIS layout representing the object studied (the ex-convent building annexed to the “Chiesa del Colletto”) in its context. From this regional map, it is possible to link to another window where another view at a major scale is represented (the projection is the red line).

In the example presented, the second QGIS project shows a façade (the one whose planar projection is the red line in Fig. 2), on which “Materials” and “Manifestation Of Damage” (Fig. 3) have been mapped.

Figure 3: Management and representation in GIS of the “ManifestationOfDamage” values, mapped on the façade analysed, and the related table in PgAdminIII.

These values are connected to the respective tables, in which characteristics and other useful information are also stored. When possible, taxonomy values are used; they are preferably acquired from the NORMAL documents, though in cases of unavailability we used other affirmed bibliographic sources (Carbonara, 2004). The schemas defined can be extracted in an exchangeable format, and can be transposed to other similar projects and reused. The same can be done with the data contained in some significant tables, like the ones containing data extracted from standards, a specialized bibliography could be generated, which could be simply updated by adding information regarding new research and definitions.

2.2 Retrieval of Data Images

This way of managing projects with the use of QGIS connected to PostgreSQL leads to new opportunities in CH documentation strategies. This tool allows for the recording and management of many type of data, including images. The images, stored and managed in PostgreSQL, are visualized by means of the database connection to QGIS. In this way, the images and their attributes describing the state of conservation of building elements as well as heritage management activities can be queried.

The image visualization may be also realized through loading the PostgreSQL database on a web page, starting from a php script connection (Fig. 4). This connection with the Web could be further exploited in future work for effectively publishing and sharing data.

The primary goal of data retrieval here is storing, sharing and updating information about sites and buildings. A georeferenced repository is structured in archives able to manage heterogeneous data: the spatial archive, the images archive and the archive capturing the CH conservation state. These retrievals are queried singularly or by means of an integrated process, according to the aims of the analysis; through the images and attributes, they offer a highly detailed reading of geometric and thematic information in the GIS environment (Fig. 5). Since PostgreSQL is an ORDBMS, the management of dynamic data tables- for example, tables concerning parameters affecting the conservation state - is also straightforward.

Figure 4: The web page connected to PostgreSQL to visualize the photo inventory.
Figure 5: Analysis of materials’ degradation starting from images, and 3D-metric model in the Belmonte Sacro Monte (UNESCO Heritage) building.

The georeferenced images repository collects different images, including the ones acquired to generate dense 3D models with low cost techniques (such as image matching and Structure-from-Motion, Chiabrando et al., 2014). It is possible to query the images of the object acquired by the camera or images of the processed 3D model (Fig. 6).

Figure 6: A phase of image matching technique generating a point cloud model.

The results are different information levels organized according to a multi-scale approach (from a territorial to an architectural scale), with the display of the building surface pathologies directly on the images and on the 3D-metric model. These can eventually be mapped using a similar approach to that described in the previous example.

3 DISCUSSION AND PERSPECTIVES

The management of Architectural Heritage documentation by means of some interoperability tools, such as standards and open-source software, has been tested experimentally in the examples presented above.

In this work, some structures enabling the use of CH semantic information (such as ontologies and formal languages) have been used together with tools that the interoperability of data (such as open-source software). The system could be enhanced through the use of object-oriented DBMS, which could catalyse implementation on ontological models. In any case the management of heterogeneous and multifaceted data has been shown to be possible. Moreover, an actual multi-scale approach has been tested, and the data in the system can be queried on different levels of interpretation an essential capability in the CH framework.

The entire workflow has some limits due to the absence of official integrated standard data models; in particular, there is a lack of standards in the acquisition and plotting phases of historical architectural heritage documentation. This relates to the fault of homogeneous geometric data, and it is resultantly more difficult to translate these data into a unique, integrated and harmonized system.

BIM are recognized as systems suitable for the 3D modelling of historical buildings, and they provide high editing functionalities for managing the object representation comprehensively. On the other hand, GIS fit other CH needs by enabling the management of more complex and irregular surfaces in a 2,5D approach. It is therefore possible to realize analysis and mapping surfaces with semantic thematic information.

Future work will address the enhancement of these systems through the improvement of the mutual integration of models and software implementation. One of the final aims of both systems is the sharing of structured data on the Web, so investigation into the development and implementation of a BIM extension (GeoBIM) on CityGML is needed. (De Laat and Van Berlo, 2011).

REFERENCES


Taboroff, J., 2000. Cultural heritage and natural disasters:


www.w3.org/RDF/ Accessed on 03/12/2014.