In this paper, the Polytechnic University of Madrid and the Complutense University of Madrid present EPIARQ. Under this acronym an epigraphic data model SDI (Spatial Data Infrastructure) is described as a first approach to designing and implementing a standard data model to catalogue epigraphic evidence as an aid to research in this scientific field under current European legislation and, in particular, within the framework of the European Directive INSPIRE, OGC standards and ISO 19100 standards. Monumental and architectural epigraphy from Late Antiquity in Hispania (5th to 7th centuries) was chosen for the experimental validation of the new data model database (DB). Once the entire structure of the model has been built, this physical database will be implemented with the aim of analyzing the improvements and innovations introduced in both access to information and maintenance processes, as well as in the scientific analysis of the epigraphic information it contains. In this research both the lexical variables of the pieces and their geographic location are addressed in answer to questions about types, materials, intrinsic elements, people / deities mentioned, etc. The new relational data model will cover the entire lifecycle of an epigraph, including data associated with the epigraphic text itself, the context of its discovery, its physical condition, its location and any associated literature or documents published in relation to it.

Key words: INSPIRE, SDI, Database, Hispania, epigraphy, Late Antiquity.

1. INTRODUCTION

Scientific methods have been used to study epigraphy for several centuries. This field has a long tradition of creating a body of knowledge concerning Greek and Roman epigraphs, and there are a number of specialist journals that publish periodic reviews. There is therefore an immense amount of information available, and even though a number of data bases are in existence to store this information and make it available to researchers, there is as yet no generally accepted standard documentation or cataloguing software. There is thus no standard data model in existence for transferring these methods to a digital data base in an orderly way, or one that can deal with all types of epigraphic finds in a single interoperable format.

On the other hand, Geographical Information Systems (GIS) have been applied to the study of epigraphs for some time. They are now used, for example, to build 3D models of inscriptions to obtain digital replicas of new finds, as an aid in maintaining existing pieces or to be put on view to the public in exhibitions. However, it is unusual to find a GIS that allows spatio-temporal studies to be made with this data.

EPIARQ can now be considered as a possible standardized method of epigraphic cataloguing and spatial analysis, in accordance with the European INSPIRE directive, and the OGC and ISO 19100 standards. This system would allow the creation of a Spatial Data Infrastructure (SDI) to study how, when, why and where inscriptions were used at any time in history, and as an aid to interoperable cataloguing of the information associated with these epigraphs between the different parts of a study.

The thesis project from which sprang this idea was the “Method for the cataloguing and analysis of monumental and architectural epigraphy of the late Ancient Hispania Period (5th, 6th and 7th centuries), in accordance with the INSPIRE directive and the OGC and ISO 19100 standards”, by doctoral student Rocío Gutiérrez, in the Geographical Engineering Doctorate Program at the Polytectnic University of Madrid (PUM). The project was approved firstly by Dr. Mercedes Farjas, in charge of the area of research on Patrimony and Archaeology in the Department of Topographical and Cartographical Engineering of the UPM, and by Dr. Isabel Velázquez Head of Latin Philology at the Complutense University of Madrid, who are also the present directors of the thesis. The thesis also forms part of the projects Sacra Tempora and Certae Aedes, epigraphical and textual sources dealing with religion in Hispania in the high Middle Ages (5th to 10th centuries), directed by Isabel Velázquez and Corpus Architecturae Religiosae Europae (saec. IV-X) (CARE-Hispania) directed by Gisela Ripoll. Both these projects give technical support in the form of a prototype web page that gives access to the implemented EPIARQ data model database.

One of the main objectives of the model, if not the most important of all, is to be able to integrate the data bases
that exist at present either on paper or in digital form into a common interoperable framework that would allow the evolution of epigraphs to be studied by means of a comparative historical analysis with the aid of GIS geovisualization tools. It would also for the first time offer researchers the opportunity to visually validate their theoretical models by a simple process.

Some of the responses that can be expected include:

- Analysis of the literary and geographical evolution of epigraphs through time.
- Identification of geographical areas by different influences such as, gods, political posts, etc.
- Identification of possible networks associated with family relationships.
- Comparison of customs in different geographical areas.

From the beginning, the aim is to create the most generally possible data model to include all types of epigraphs and all the information relating to them throughout their existence, from their discovery until their possible disappearance, with references to any publications and documents issued concerning them. However, we are aware that it must also be open to modifications, for example, being able to add new fields for new uses or variations through time of existing uses.

EPIARQ was scheduled to be developed as follows:

- A study of existing data sets chosen as references for the implementation of the first prototype.
- Creation of the first data model with the information obtained from the first stage and after interviewing epigraphic specialists from the two associated projects (Sacra Tempora and CARE-Hispania).
- Implementation and optimization of the first Data Base.
- Implementation of the web application to be used as the access interface and data modification for the new Data Base.
- Georeferencing of data in the WGS84 reference system.
- Input of data in the Database (DB).
- Implementation of a viewer to facilitate space-time analysis.

Generation of web services under OGC standards with the most representative thematic maps.

Creation of metadata of data and services in accordance with the indications of the Spanish Metadata Nucleus (NEM) and expanding it as required.

Implementation and publication of an Spatial Data Infrastructure (SDI) geoportal to contain the previously created metadata, services, and GIS viewer.

2. LEGAL FRAMEWORK

The national heritage consists of inherited historical items considered to be of cultural interest. This value does not usually change much, so that the definition of the heritage can be considered as closed, although the set of items of which it is composed are subject to progressive modifications and new incorporations. The same interpretation can be applied to epigraphy, so that the epigraphical heritage can be considered as a sub-set of the historical heritage.

The historical heritage is protected by numerous figures. The architectural heritage is included in several of these figures and incorporates inscriptions and epigraphs found in churches and non-ecclesiastical buildings. In Spain, the care of the historical heritage is the responsibility of the different Autonomous Communities, as laid down in the Law 16/1982 dealing with the Spanish historical heritage. Almost all the items in this heritage can be geo-referenced by a system of space-time coordinates. Both GIS and CAD (Computer-Aided Design) systems have been added as conservation and restoration supports (e.g. 3D surveys by laser scanner) to new techniques applied to geo-referencing and modelling in 2 and 3 dimensions. This implies an increase in problems of interoperability among the different data capture and analysis systems.

2.1. The ISO 19100 Family

After several preliminary initiatives by AENOR and CEN in standardizing most aspects of the information and communication technologies and their application to geographical Information (GI), ISO/TC211 defined a wide set of standards on all aspects of geographical information. Today, there exists a complete set of 33 approved documents and reports and 20 more are in preparation.

2.2. INSPIRE

Considering the need to regulate the data models used in geo-referencing, the INSPIRE Directive (2007/2/CE) laid down general rules for creating spatial data infrastructures in the European Community for environmental purposes and based on the SDIs of the member states. Its philosophy is to share and make accessible to citizens all the available official geographical information, while avoiding redundancies as far as possible. Heritage data should be incorporated into the set of data on Protected Sites since all heritage elements are considered to be protected by definition. This data set is incorporated into the First Inspire Appendix and is defined in INSPIRE Data Specification on Protected Sites – Guidelines.

To construct the model and ensure its interoperability, we must resort to the conceptual scheme defined in ISO 19011:2002, the common framework for maintaining interoperability among different data sets added to the INSPIRE model and the standards laid down by ISO19100, which indicate the use of classes, stereotypes and relationships, etc. Of special importance are ISO 19101:2002, 19103:2005, 19019 y 19131.
2.3. LISIGE

Law 14/2010 (5 July) on infrastructures and geographical information services in Spain (LISIGE) incorporates the Directive 2007/2/CE (INSPIRE) and guarantees its compliance, including the establishment of the Spanish GI Infrastructures, which integrates all geographical infrastructures under the responsibility of Spanish Public Administrations. In order to adapt INSPIRE to the laws of individual countries, different working groups were formed to develop the appendices to the directive. One of these working subgroups of the GT IDEE has the mission of harmonizing and integrating heritage data available to the public administrations within the INSPIRE Directive (2007/2/CE) and under Spanish Law 16/1982, the laws created by the Autonomous Communities and the recommendations and standards of the Higher Geographical Council (CSG). As a result, a data model has been created for archaeological items by extending the INSPIRE protected sites model, but so far no publication has been issued for modelling epigraphs, which has now become the objective of the present study.

2.4. OGC Standards

Similarly, and in parallel to previous developments, the OGC (Open Geospatial Consortium) has laid down a series of standards for geo-spatial services and services based on localization, and has defined geo-spatial information access interfaces and data coding specifications, consultations and viewing modes to make them freely available to all who want to use them.

3. CONCEPTS OF SDI, GEOPORTAL AND INTEROPERABILITY

3.1. SDI

The proposed epigraphic data model will belong to the Spatial Data Infrastructure. According to LISIGE Article 3, Point 1a, a spatial data infrastructure is a virtual structure in a network formed by:

- Georeferenced data and its description in metadata.
- Interoperable GI services distributed throughout different information services accessible by Internet with the minimum of protocols and standard specifications.
- Services should include technologies of search and access to the data and the standards for their production, management and diffusion.
- Agreements on common information, access and use among producers and between the latter and other users.
- Mechanisms, processes and coordination and follow-up procedures established and managed in conformity with the LISIGE law.

3.2. Geoportal

The term Geoportal refers to an Internet page or equivalent that provides access to interoperable GI services from different organizations, organisms, entities or public administrations and incorporates at least one service that allows the search for and access to geographical data and services.

3.3. Interoperability

In defining a data model, within European standards, one of the fundamental features is that the data should be interoperable. The interoperability of GI services refers to its ability to combine geographical data sets and to facilitate the interaction of GI services without the need for repeated manual intervention, providing consistent results and adding to the value of the geographical data and services.

4. SOURCE DATA SETS

In order to build the prototype of the new Epigraphic Data Base, and therefore the first specific data model of the Late Antiquity in Hispania, we considered data sets relating to official buildings and public works. Not only are the commemorative inscriptions of buildings, monuments, works and churches included, but also other internal inscriptions relating to their functions, builders, promoters, and in general any inscriptions that might give information on any aspect of the item under study.

Although a detailed analysis of the categories in the model will be given below, the first step is a general classification of the epigraphic data sets according to their origins into four main groups:

- Material inscriptions, or those whose location and exact reference are known. In this group, a distinction is made between those conserved in situ or in museums and collections but whose origin is known, and others known and conserved but whose origin and the circumstances of their discovery are unknown.
- Inscriptions that have been transmitted in writing, whose geo-references may or may not be known. These latter may not always be available, even though a good description is in existence. The inscriptions may have been transmitted by means of epigraphic manuscripts or other compilations.
- Traditional literary inscriptions: these are not physically conserved but exist in references to literary sources, which may also include their content. As in the preceding case, there is often no geographical reference and so they cannot be placed on a map. They can however be catalogued in the data base and in certain cases they may be assigned to a town or province.

---

* IDEE, Documento de constitución del Grupo de Trabajo Temático de Patrimonio Histórico (GTT-PAH) en las IDE. Consejo Superior Geográfico, Madrid, 28-May-2010.
* http://www.opengeospatial.org/
* http://www.idee.es/faq
* http://www.idee.es/faq
• Inscriptions whose text may have been composed but it is not known whether the actual inscription ever existed.

When the origin is unsure, for example in cases in which it is not known whether the inscription ever really existed or whether it was merely a literary composition, any indications of the location offered by the content in the form of mentioning persons or places is considered in assigning a geographical location.

The data sets to be studied belonging to the above classification will be obtained from the online library Hispania Epigraphica, from the AEHTAM data base (both from one designed for the Sacra Tempora project and from another for the CARE-Hispania project), and from the digital and non-digital archives of the Department of Latin Philology of the Complutense University of Madrid. These will be transformed into data sets according to the applicable INSPIRE standards and a check will be made as to whether they correspond with existing models in the INSPIRE and other data specifications.

5. DATA MODEL

The EPIARQ model was written in UML (Unified Modeling Language) and consists of three main blocks:
• A block defined in the INSPIRE specifications for Protected Sites and therefore unmovable.
• A block that will act as an adapting bridge between the INSPIRE specifications group and the model required to represent any inscription.
• A block that defines the classes necessary for the implementation of information associated with an epigraph.

The first problem we meet with is whether or not to include the class known as Building in the INSPIRE specifications to which building inscriptions might be added. To reach a conclusion we simply have to ask ourselves whether a building inscription has always been a part of a building throughout its entire life cycle. In other words, if the building disappears, does the epigraph still exist? To answer this question, we can consult Point Nº4 – SET OF SOURCE DATA – of this document, which indicates a preliminary classification according to the origin of the data. Data can come from three sources: material inscriptions, from manuscripts (and may or may not have a material form), and inscriptions from literary tradition. Therefore, information will not always be available on the building in which the epigraph was inscribed at some point in history and may also be lacking on its geographical location.

However, this topic does not finish here, since when deciding whether or not to include Building in the new epigraph data base we must also ask; Are not all epigraphs at some time associated with a building? The straight answer is No, only building inscriptions, since they may also be associated with other types of object. Therefore, if we want to make a model that can deal with any type of epigraph, even though the test prototype is a building inscription from the late Hispania Period, having Building as the only container of an epigraph would add an error to the model. We therefore concluded that we should omit the Building class from the EPIARQ data model.

5.1. INSPIRE data models

The first block of the EPIARQ model consists of the INSPIRE data model, and can be studied in the INSPIRE Protected Sites data specification.

If we look at the diagram in Figure 1, it is formed by a series of “featureType”, “dataType” and “enumeration” classes. Special importance must be given to the “protectedSite” class, since all the other models cited below depend on it.

5.2. Bridging data model

As the bridging data model, we accepted the one recommended in GTT-PAH, as it was proposed by the group of specialists designated by the IDEE, and also its variant in the IDEARQ project. The authors had already worked with these models in the design and implementation phases of IDEARQ, which became the first known IDE prototype.16

15 GTT-PAH, Borrador del modelo de datos de Patrimonio Histórico como Lugares Protegidos de INSPIRE. Consejo Superior Geográfico, Madrid, 28-Feb-2011.
Figure 2. Data model in UML.
As can be seen in Figure 2, when implementing a data model for the historical heritage the most important class is the featureType “Protected Sites” from the INSPIRE data model. This featureType has been amplified and redefined, as it was originally designed for protected environmental sites. New attributes and relationships are included in the new model to meet the requirements of the Spanish heritage modeling (yellow-coloured classes in Figure 2).

The importance of the featureType “CulturalEntity” should be pointed out here, as it will be the link between the new model described in the following section and is also the nucleus of the EPIARQ study.

Among the Cultural Entities, we are interested in Material Entities, especially those created by men and classified as “HumanMadeFeature”, since epigraphs are made by men but are not always inscribed on man-made objects.

5.3. The EPIARQ data model

The third block consists of EPIAQ’s own data model, which is in turn sub-divided into three different classes:
• Classes belonging to the main model, coloured white. These are described in greater detail in this section.
• Enumeration classes or closed lists, coloured pink. These are used to list elements in general types, and will be used as data types for the attributes of the principal classes.
• CodeList classes or open lists, coloured yellow. These will list possible values of attributes, will be expandable in the future and may contain more elements.

The main class, on which all other classes depend or are related, will be Epigraph (see Figure 3), and to it will belong all the attributes pertaining to individual epigraphs, such as for example:
• The material on which it is written
• The support or object on which the inscription was made
• Type and form of the object
• Technique used for the inscription
• Number of inscriptions contained by the epigraphic element
• Type of epigraph considered
• Dimensions of object and inscription
• Number of lines
• Minimum and maximum height of letters
• State of conservation both archaeologically and epigraphically
• Persons responsible for the piece
• If a donation, by whom
• Any accessory elements
• The corpus, volume and entry in which it is inscribed
• Where it is conserved and its inventory number

The place in which an epigraph was discovered is included in “Location” and details of its surroundings in “ArchaeologicalContext”.

The “AssociatedDocumentation” class is linked to “ArchaeologicalContext” and may provide ideas on location, position of the object, other surrounding objects and their position, etc.

“AssociatedDocumentation” is also related to the epigraph itself, since there may be rubbings, photos, 3D images, or videos about it.

The “Text” class complements the information in “Epigraphy” and refers to the epigraphic text and where information about it can be obtained, such as:
• text
• translation of the text if available
• any critical reviews
• comments by a reviewer
• any known variants
• type of writing and direction
• punctuation
• number of texts it is divided into
• language in which it is written
• signs used, metrology,17
• abbreviations and links
• special signs in the text
• dating of the script

Finally, the “Bibliography” class is described, which indicates:
• the page of the book or article/paper in which reference is made to the epigraph
• comments about the epigraph in the book or paper
• if it has been transmitted in manuscript form or physically exists
• any other books or articles related to it, either by editio princeps, text editions or complementary bibliography

According to information from the Hispania Epigraphic Archives, at the present time there are an estimated 28,000 inscriptions in existence from the Pre-Roman, Greek, and Roman Ages up to the 7th century, and medieval inscriptions are now beginning to be incorporated. This is precisely the data base that for the first time will systematically register the immense number of new medieval Latin and Romance inscriptions that to date have only been published in a limited number of volumes and have never previously been included in a data base.

17 These attributes are included to enable new types of inscriptions to be included in the Data Base in the future, in order to make the model as generic as possible, with epigraphs of whatever content, from any period and even from other languages.
Figure 3: EPIARQ Data Model.
6. TOOLS ASSOCIATED WITH THE PROJECT

Besides the tools described above, in the research project described in this paper, a data input interface, a map viewer and an IDE web page (described above) will also be developed for the BDD. These will be activated by different keys that incorporate the signs of the epigraphs themselves and the diacritical signs used to represent them, implemented in php programming language to facilitate data input from the online interface.

7. REFERENCES

7.1. Bibliographic references

Authors’ Note: The bibliography relating to the epigraphic corpora is a selection based on the inscriptions being selected. This selection is of corpora only, since there is an abundant bibliography on specific pieces and publications in relation to Hispanic epigraphs of all types. This work is fundamentally done by the library of the Hispanic Epigraphic Archives and their digitalized data bases, and by the continuous bibliographic revisions carried out by Hispania Epigraphica.


ALMACRO BASCH, M., Segobriga II. Inscripciones ibéricas, latinas paganas y latinas cristianas (EAE 127), Madrid, 1994.


DIARIO OFICIAL DE LA UNIÓN EUROPEA, “REGLAMENTO (UE) No 1089/2010 DE LA COMISIÓN de 23 de noviembre de 2010 por el que se aplica la Directiva 2007/2/CE del Parlamento Europeo y del Consejo en lo que se refiere a la interoperabilidad de los conjuntos y los servicios de datos espaciales que se implementan en infraestructuras de datos espaciales. [BERLÍN]”.

DIARIO OFICIAL DE LA UNIÓN EUROPEA, “REGLAMENTO (UE) No 1089/2010 DE LA COMISIÓN de 23 de noviembre de 2010 por el que se aplica la Directiva 2007/2/CE del Parlamento Europeo y del Consejo en lo que se refiere a la interoperabilidad de los conjuntos y los servicios de datos espaciales”. 12-Aug-2010.

DIEGO SANTOS, F., Epigrafía romana de Asturias, Oviedo, 1959.

DIEGO SANTOS, F., Inscripciones romanas de la provincia de León, León, 1986.


ESCRIBANO PAÑO, Mª. V., coord., La investigación sobre la Antigüedad Tardía en España: estado de los estudios y nuevas perspectivas, Málaga, 2009 (Mainake 31: número monográfico en homenaje a Rafael Puertas Tricas).


7.2. Web’s references