A SDI's in Archaeology in the Iberian Peninsula: Achieving interoperability of Cultural Heritage data in INSPIRE

Carlos Fernández Freire
GIS Laboratoty
Center for Social
Sciences and Humanities
(CCHS) Spanish National
Research Council (CSIC)
C/Albasanz, 26-28
Madrid, Spain
carlosf.freire@cchs.csic.
es

Ernesto Salas Tovar
GIS Laboratoty
Center for Social
Sciences and Humanities
(CCHS) Spanish National
Research Council (CSIC)
C/Albasanz, 26-28
Madrid, Spain
e.salas@csic.es

Rocío Gutiérrez
González
GIS Laboratoty
Center for Social
Sciences and Humanities
(CCHS) Spanish National
Research Council (CSIC)
C/Albasanz, 26-28
Madrid, Spain
rocio.gutierrez@cchs.csi
c.es

Juan M. Vicent García
Institute of History (IH)
Spanish National
Research Council (CSIC)
C/Albasanz, 26-28
Madrid, Spain
juan.vicent@cchs.csic.es

Abstract

The Center for Social Sciences and Humanities (CCHS) of the Spanish National Research Council (CSIC) puts forward IDEARQ. It is a platform aimed at the dissemination of collections of georeferenced archaeological data, product of the research carried throughout many years. The aim is to offer a reference node for archaeological research data, arranging an Iberian Archaeology Spatial Data Infrastructure.

It provides OGC-compliant WMS on archaeological thematic data that can be linked to Protected Sites included in the corresponding data specification. IDEARQ gathers data relative to the chronological and typological characterization of archaeological sites, related documents such as rock art pictures, analysis results and related bibliography.
.

Keywords: SDI, INSPIRE, Archaeology, Data Modelling, GIS, OGC Services

1 Introduction

IDEARQ a platform aimed at the dissemination of georeferenced archaeological collections, product of the research carried throughout many years. It has been jointly developed by the GIS Laboratory from the Center for Social Sciences and Humanities (CCHS) and the Landscape Archaeology and Remote Sensing Laboratory (LabTel) of the Spanish National Research Council (CSIC), along with the different research groups responsible for each datasets (Instituto de Historia, CCHS – CSIC).

Nowadays, it comprises three datasets linked to nearly 2,000 archaeological sites spread all along the Iberian Peninsula: Radiocarbon dates, Archaeometallurgical analysis and Levantine Rock Art documentary data – mainly images –. The aim is to offer a reference node for archaeological research data, arranging an Iberian Archaeology Spatial Data Infrastructure.

A specific data model has been proposed as an extension from the Protected Sites Data Specification (PS Data Spec. onwards) aimed to be extendable in the future with any cultural heritage data related to a geographical location.

2 Datasets

IDEARQ comprises datasets that have been elaborated by several research projects within the CCHS (Center for Humanities and Social Science, CSIC). The extendable design of the data model allows for the integration of many datasets, although, up to now it is set up of three datasets:

2.1 Rock Art Images

The "Corpus de Pintura Rupestre Levantina" (Corpus of Levantine Rock Art or CPRL) is a photographical archive generated by the Postpaleolithic Art inventory, directed during 1969-1974 by M. Almagro Basch along the Spanish mediterranean geographic area. It consists of a collection of documents with information on 139 Levantine and Schematic rock art sites, currently included by UNESCO on the World Heritage List

2.2 Radiocarbon Analysis:

More than 4,000 analytic data from Radiocarbon analysis all along the Iberian Peninsula gathered from bibliographic sources (Gilman Archive). Its publication enables access to basic reference data used in prehistoric archaeology studies ranging from Mesolithic (X mil cal. BC) to Iron Age.

2.3 Archaeometallurgy:

Nearly 20,000 archaeometallurgy artifact's elementary analysis, samples and materials related to metallurgic activity, ranging from Chalcolithic to Modern Times. It gathers analysis of chemical composition and metallographic data of artifacts from the Iberian Peninsula and Balearic Islands along with lead isotopes analysis over the artifacts and mineral extraction areas samples to find out the mineral object's place of origin.

3 Interoperability within INSPIRE

Since the INSPIRE directive was published in 2007 its development has been a long way with many efforts and work from its promoters and partners. One of the main challenges of this Directive is to achieve the full spatial data interoperability in the EU. Although the assemble of Cultural Heritage data within the INSPIRE directive does not offer an straightforward solution, Heritage is subjected to protection becoming part of INSPIRE's Annex I Protected Sites (for a discussion see [1,2]) which means that they are considered as reference data. Since Cultural Heritage data are included in the Protected Sites theme, guaranteeing their inclusion in the Directive is priority [1,3].

The main goal of IDEARQ is to enhance the dissemination of archaeological research projects carried out by CSIC researchers that have a spatial dimension focused on archaeological sites that always have some degree of protection. Thus, the linkage of these datasets with INSPIRE is a relevant matter, and we have therefore undertaken the task of elaborating a data model that can be integrated within the Directive.

3.1 The Cultural Heritage Application Schema and IDEARQ Data Model:

Our starting point is the Cultural Heritage Application Schema (CHAS onwards), that offers a linkage between the PS Data Spec and Cultural Heritage [1,3]. It is an extendable, generic and interoperable Application Schema to achieve an implementation basis for heritage spatial data with a known location, boundary and area, as stated in the PS Data Spec. [4].

Cultural Heritage is commonly considered as the features, tangible or intangible that a community reckons as their own, reinforcing a common identity throughout history. Therefore this kinship accepts that these elements deserve to be protected to recall their own achievements [3]. The location of Cultural Heritage items make them into Spatial Data that can be found in a place, or linked to a geographical location. A relationship becomes apparent in the case of immovable features such as buildings, archaeological sites, historical landscapes, etc. Nonetheless other items which aren't tied to a place, like folklore, also have a spatial dimension.

This Schema had to respond to the gap left in the PS Data Spec. which is addressed to the legal aspects of protection but ignores the real entity of the protected site [1]. In this context it is aimed at three elements: the protection legal framework, the protected cultural entities and the related documentation.

IDEARQ is modeled as an extension of the protected cultural entities and documentation parts of the CHAS in an effort to enhance future interoperability with other datasets.

Cultural Entities are the protected heritage objects tangible or intangible, with several subtypes according to the CIDOC-CRM Schema: Human Made Object, Human Made Features and Natural Features provides enough variability to describe any tangible cultural item. Many different real – world objects can be cultural entities, in this case archaeological sites, Rock Art Stations, metal pieces and diagnosis material.

The application of this schema to IDEARQ leads to two main types definition: On one hand the archaeological sites all along the Iberian Peninsula (including Balearic Islands) as buildings, cities, villages, necropolis, etc. On the other hand the CPRL Stations, the places where the rock art is placed (caves, shelters, etc.). Both are subtypes of the main IDEARQ class: Cultural Entities.

The bibliographic research carried out by professor Antonio Gilman has yielded a vast corpus of Radiocarbon dates obtained from archaeological sites all along the Iberian Peninsula. It is a dataset of great significance because it had a great impact in Prehistoric research. The publication of these radiocarbon dates forced researches to reconsider previous hypothesis [5]. IDEARQ gathers information on the kind of material that was dated, its archaeological context, technical information of the dating process such as the age, standard deviation, isotopic fractionation [6], the methodology used, the responsible laboratory, and information on the calibration, should there be.

The CPRL Stations, as we mentioned, are the places of rock art paintings located all along the Levantine coast. The photographical corpus attached to them is an outstanding compilation of colour photographs and contextual information on three quarters of the paintings known in Mediterranean Spain. Iconographic motives and their description (Human or animal shapes, symbols, war or hunting scenes, etc.) are collected in the database, along with information on their degree of conservation and bibliographic notes to identify the motives captured in the photographs. Since all the photographs were scanned for a previous project that led to the publication of a web page (http://www.prehistoria.ih.csic.es/AAR/menu), a link to the scanned picture grants its visualization.

The archaeometallurgic corpus joins two different datasets: the main one gathers information on elementary analysis of metal artefacts found in archaeological contexts. This includes the results of elemental analysis -for an explanation of these techniques see [7]- (Cu, Au, Ag, Sn, Pb...), which gives an indication on ancient metallurgic techniques, along with information on the artefact subject to analysis: chronology and typology, description, and the museum where it is kept, as well as the archaeological site of reference. The second dataset gathers isotopic analysis of archaeological metals, a key information in provenance and trade/exchange studies [8]. $207Pb/206Pb$, $208Pb/206$, $208Pb/204$ and $206Pb/204Pb$ ratios for archaeological items and ore samples can help us to link these pieces to the mines where they were extracted. Another geographical layer showing the location of the samples is added for this dataset.

4 Project Development

The model has been implemented in a relational database accessible to all participants for data analysis, maintenance and data input. It has been deployed using PostgreSQL and the PostGIS extension to handle spatial data. Everything is stored in a common server that also constitutes the core of the web services and online cartographic viewer.

The complexity of the data model led to a multipart database design that will enable future integration of new datasets.

4.1 OGC Services

IDEARQ fulfils the basic requirements of an SDI offering a Web Map Service (WMS) and service and dataset metadata (ISO 19115 compliant) available through a service catalog that fosters metadata of all the CSIC – CCHS projects.

The WMS serves three point geometry layers, based on database views, that grant a thorough access to the data stored in IDEARQ through the GetFeatureInfo method:

4.1.1 Layers

"idearq" summarizes the information of all datasets offering generic information about the archaeological sites, their description, chronology, typology (human settlement, cemetery, industrial activity, etc.) and the number or radiocarbon datings and Rock Art pictures.

"idearq_c14" adds all the information related to every radiocarbon date, the resultant age along with the standard deviation, the kind of material dated (bone, charcoal, etc.), the laboratory responsible and the dating methodology.

"idearq_cprl" holds a description of every picture within the CPRL, specifying the representation (cattle, human figure, etc.), bibliographic information and a link to the picture, possible through a connection to the Martin Almagro Basch Archive.

Two more layers will be available as soon as archaeometallurgy data input is complete: the elementary analysis of archaeological artifacts attached to their site of origin, and led isotopes analytical results of archaeological items and samples outside archaeological sites, useful to trace the spread of mineral sources.

Visualization of this datasets offers some difficulties due to the amount of archaeological sites, the detail of the information conveyed and the necessity to preserve the geolocation information of sites susceptible of pillage.

In order to overcome these difficulties, two different SLD (Styled Layer Descriptor) visualization styles have been generated. Both are scale – limited to 1:200.000 so the exact position of sites is not accessible, in an effort to preserve their integrity; the default style displays all individual sites, being more suitable for small scales, while the alternative style displays a cluster performed by a vector to vector transformation available in Geoserver software (PointStacker). Access to a complete description of every item is granted through the customization of the GetFeatureInfo response, offering a legible reply, enhanced with features such as the link to the CPRL pictures.

5   Map Viewer

Public access to IDEARQ data is granted through OGC web services and through a cartographic viewer that displays all data within the database (Figure 2). It displays only one layer, the archaeological sites that can be filtered out by chronology, typology or dataset. The main documentation of the database is displayed for every site as in the WMS: generic information of the archeological site, date and standard deviation, the responsible laboratory, the kind of material used (charcoal, bone…) and the analytical method for radiocarbon dating; a classification of the motif, the bibliographic note and a link to the picture for Levantine Rock Art.

The viewer relies on the OpenLayers 2.12 API, under javascript technology. It offers two different base layers, MapQuest tiles based on Open Street Maps, and a satellite view composed of the NASA Bluemarble with topography and bathymetry image, and a Landsat images mosaic. It also relies on MapQuest Nominatim service for geolocation.

6   Discussion

Although SDI's where not developed to divulge scientific knowledge we use them as a vehicle to disseminate research projects results. These results belong to many disciplines for which the territory is the common nexus.. Thus we propose a framework where SDI are a powerful and interoperable means to grant public access to research data, accomplishing the objectives of the Berlin Open Access Initiative, signed by the Spanish National Research Center (CSIC) in 2006.

Furthermore, the possibility of an integration of Cultural Heritage data within the INSPIRE Directive might approach it to society to ensure and reinforce its protection. It is a new way to embed Cultural Heritage in community enhancing their perception, helping administration in their effort for coordination and management and helping research institutions both as cultural heritage data consumers and producers [1].

7   Acknowledgements

References

[1] Carlos Fernández Freire, Isabel Del Bosque González, Juan M. Vicent García, Esther Pérez Asensio, Alfonso Fraguas Bravo, Antonio Uriarte González, Pastor Fábrega-Álvarez and Cesar Parcero-Oubiña. A cultural Heritage Application Schema: Achieving Interoperability of Cultural Heritage Data in INSPIRE. *International Journal of Spatial Data Infrastructures Research*, Vol. 8: 74-97, 2013.

[2] P. McKeague. Anthony Corns and Robert Shaw. Developing a Spatial Data Infrastructure for Archaeological and Built Heritage. *International Journal of Spatial Data Infrastructures Research*, Vol. 7: 16-37, 2012.

[3] Carlos Fernández Freire, César Parcero Oubiña and Antonio Uriarte González, editors. *A data model for Cultural Heritage within INSPIRE*. CAPA, 35, Instituto de Ciencias del Patrimonio(CSIC), Santiago de Compostela, 2014.

[4] INSPIRE Thematic Working Group *Protected Sites*. INSPIRE Data Specification on Protected Sites – Guidelines. 2009-09-07.

[5] Antonio Gilman. El impacto del Radiocarbono sobre el estudio de la Prehistoria tardía de la

Península Ibérica: Breves comentarios. *Trabajos de Prehistoria*, 60, N. 2:7-13, 2003.

[6] Ingrid U Olsson. Radiocarbon Dating History: early days, questions and problems met. *Radiocarbon*, Vol. 51, N. 1: 1-43, 2009.

[7] A.M. Pollard, C.M. Batt, B. Stern and S.M.M. Young. .*Analytical Chemistry in Archaeology*. Cambridge University Press, Cambridge, 2007.

[8] Zofia Anna Stos-Gale and Noël H. Gale. Metal provenancing using isotopes and the Oxford archeological lead isotope database (OXALID). *Archaeological and Anthropological Sciences,* Vol. 1, Issue 3:195-213, September 2009.

Figure 1: Cultural Heritage Application Schema extended

Figure 2: IDEARQ Map viewer

Figure 2: IDEARQ Map viewer