

"SILEX: La Infraestructura de Datos Espaciales del yacimiento arqueológico de Casa Montero"(*).

A. Fraguas¹, A. Menchero², A. Uriarte¹, J. Vicent¹, S. Consuegra¹, P. Díaz-del-Río¹, N. Castañeda¹, C. Criado¹, E. Capdevila¹ y M. Capote¹

¹Grupo de Investigación "Prehistoria Social y Económica"
Instituto de Historia, CCHS - CSIC
Albasanz 26-28, 28037 Madrid (Spain)

¹{alfonso.fraguas, antonio.uriarte, juan.vicent, susana.consuegra, pedro.diazdelrio, nuria.castanyeda, cristina.criado, enrique.capdevila, marta.capote}@cchs.csic.es,

²antonio.menchero.fernandez@gmail.com

RESUMEN

El objetivo es presentar la experiencia de utilización del Software Libre y de Código Abierto en el desarrollo e implantación del Sistema de Información (SI) del yacimiento arqueológico de Casa Montero. Se muestra el potencial de combinar las Infraestructuras de Datos Espaciales y las bases de datos orientadas a grafos en la gestión, investigación y difusión del Patrimonio Arqueológico (PA). Se ejemplifica la propuesta tecnológica a través del ejemplo de SI del yacimiento arqueológico de Casa Montero (Madrid), SILEX (Sistema de Información Locacional en XML). La implementación de sistemas de información para la gestión de PA colabora fuertemente en tres actividades básicas que han de ser realizadas sobre él. En primer lugar, la gestión, con el objetivo de protegerlo y conservarlo; en segundo término, la investigación, de cara a elaborar conocimiento histórico mediante el análisis arqueológico de los restos materiales del pasado; y en tercer lugar, la difusión, con la pretensión de divulgar el PA y el conocimiento histórico generado a partir del mismo entre la sociedad.

Palabras clave: Patrimonio Arqueológico, IDE, SOA, ROA.

ABSTRACT

The aim is to present the experience of using Free Software and Open Source in the development and implementation of the information system (IS) of the archaeological site of Casa Montero. One shows the potential of combining spatial data infrastructures and graph-oriented databases in the management, research and dissemination of the Archaeological Heritage (AH).

One exemplifies the technological proposal by means of the IS of the archaeological site of Casa Montero (Madrid), SILEX (Locational Information System in XML). The implementation of information systems for the management of AH is working hard on three basic activities that must be carried out. First, management, in order to protect and preserve it; secondly, research, in order to develop historical knowledge through archaeological analysis of material remains of the past; and thirdly, diffusion, with the purpose of disseminating AH and historical knowledge generated from it among society.

Key words: Archaeological Heritage, SDI, SOA, ROA.

() SILEX ha sido desarrollado como parte del "Proyecto de Investigación Arqueológica en el yacimiento de Casa Montero (Madrid). Producción y circulación de sílex en el neolítico de la Meseta", financiado por Autopista Madrid Sur C.E.S.A. en el marco del Convenio de Colaboración entre la Consejería de Cultura y Deportes de la Comunidad de Madrid, el Consejo Superior de Investigaciones Científicas (CSIC) y Autopista Madrid Sur Concesionaria Española S.A. para la investigación, conservación y difusión del yacimiento arqueológico de Casa Montero (Madrid).*

0. INTRODUCCIÓN

El proyecto de Tecnologías de la Información y la Comunicación (TIC) para la sistematización digital del material arqueológico recuperado en la zona de Vicálvaro (Madrid, España) se ha desarrollado en su totalidad utilizando software libre y de código abierto.

En el año 2006, cuando comienza el diseño y ejecución de la parte tecnológica del proyecto Casa Montero, confluyen diferentes factores que llevan a la propuesta y, finalmente, al empleo de este tipo de licencias libres en detrimento de otras privativas. Entre esos factores se encuentran la historia del grupo de investigación que lo gesta dentro del Centro de Ciencias Humanas y Sociales del CSIC (GIPSE: Grupo de Investigación en Prehistoria Social y Económica) desde la década de 1990 [1], el prototipo desarrollado para una tesis doctoral leída dentro de GIPSE en abril de 2006 sobre arte rupestre de África nororiental (ARANO [2]), y la firma por parte de la institución a la que pertenece GIPSE del acuerdo *Open Access* en enero de 2006 [3] por la que se compromete a

“su disponibilidad libre en la Internet pública, permitiendo a cualquier usuario leer, descargar, copiar, distribuir, imprimir, buscar, o enlazar a los textos completos de estos artículos, capturarlos (crawling) para indización, utilizarlos como datos para software, o usarlos para cualquier otro propósito legal, sin otras barreras financieras, legales o técnicas que aquellas inseparables del acceso a la propia Internet. La única limitación a la reproducción y distribución, y la única función del copyright en este dominio, debería ser dar control a los autores sobre la integridad de su obra y el derecho a ser adecuadamente reconocidos y citados”.

1. CASA MONTERO COMO PATRIMONIO ARQUEOLÓGICO

Las vetas de sílex que los pobladores del valle del Henares (en el suroeste de la actual Comunidad de Madrid) (Fig. 1) utilizaban hace 7.000 años fueron explotadas mediante lo que se denomina técnicas de minería prehistórica [4].

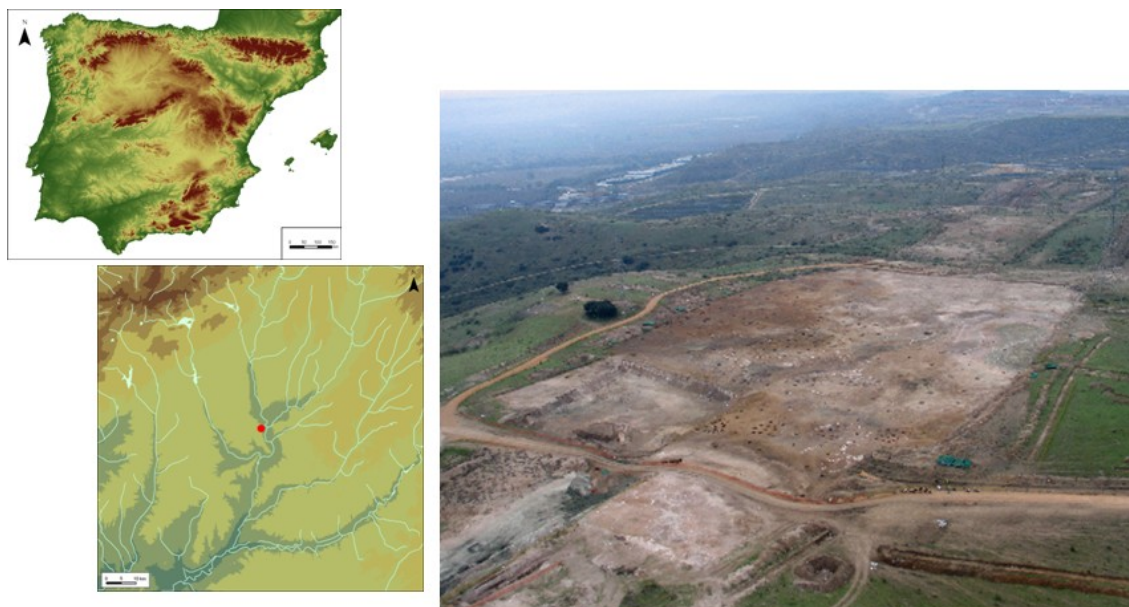


Figura 1: Localización del yacimiento de Casa Montero

El yacimiento arqueológico de Casa Montero fue excavado parcialmente entre septiembre de 2006 y agosto de 2008 generando gran volumen de material. Las 4 hectáreas ocupadas por el yacimiento contienen 4.000 pozos mineros de los que se han excavado 324. En el interior de estos pozos había 65.885 kilos de material repartidos en 3.375 estratos que los rellenaban. En el interior de estos estratos se han localizando miles de piezas, de las que se han clasificado 190.491 (Fig. 2).



Figura 2: Materiales del yacimiento de Casa Montero

El Patrimonio Arqueológico forma parte del Patrimonio Histórico y del Patrimonio Cultural, siendo objeto de protección en la legislación española (Ley 16/1985). En esta ley se contempla la necesidad de realizar trabajos arqueológicos previos a cualquier remoción de tierra. En los trabajos previos a la construcción de la autopista de circunvalación a Madrid M-50 se localizaron importantes restos de extracción de sílex.

Mediante la implementación de Sistemas de Información para gestionar el Patrimonio Arqueológico se aportan los medios para abordar su protección y conservación (la gestión), obtener conocimiento histórico a través del análisis arqueológico (la investigación), y difundir en la sociedad el conocimiento histórico obtenido tras el análisis del Patrimonio Arqueológico (la difusión).

2. ARQUITECTURA Y SOFTWARE EMPLEADO

Para gestionar y difundir el ingente material arqueológico recuperado en yacimiento de Casa Montero se han planteado servicios Web de información temática y geográfica y una interfaz de usuario Web que es cliente de dichos servicios y los integra (Fig. 3).

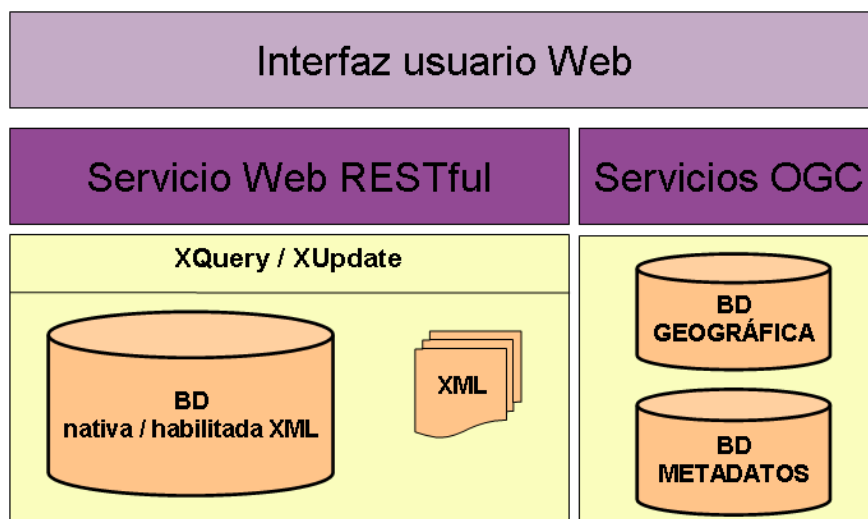


Figura 2: Arquitectura multicapas de SILEX

2.1 Bases de datos

Para el almacenamiento de datos temáticos se ha optado por la base de datos open source eXist y para la información espacial por la combinación de PostgreSQL y PostGis para la información vectorial y el sistema de archivos del sistema operativo (Linux Ubuntu 8.04) para la información ráster.

eXist (<http://exist.sourceforge.net/>) es una base de datos nativa XML *open source* suficientemente madura tras 10 años de desarrollo. Entre sus valores añadidos conviene destacar que potencia el uso de XQuery como lenguaje de servidor -no sólo de consulta-, incorpora el procesador de formularios XForms, llamado XSLTForms, y

dispone de interfaces RESTful, XMLRPC y WebDAV. La versión utilizada en este proyecto es la 1.4.0, aunque ya está disponible la versión 1.5.

2.2 Servicios Web

Para el caso de la información espacial, se han utilizado los definidos por el *Open Geospatial Consortium* (OGC), implementando una *Service Object Architecture* (SOA). Se ha utilizado Geonetwork (<http://geonetwork-opensource.org/>) para dar respuesta al servicio *Catalog Web Service* (CWS) del OGC y Geoserver (<http://geoserver.org/>) para implementar los servicios OGC *Web Map Service* (WMS), *Web Feature Service* (WFS) y *Web Coverage Service* (WCS).

Según la descripción que se realiza en *Sourceforge*, GeoNetwork es un sistema para la gestión de un catalogo de metadatos basado en web desarrollado por la *Food and Agriculture Organization of the United Nations* (FAO-UN), el *World Food Programme* (WFP) y el *United Nations Environment Programme* (UNEP). El sistema implementa las normas ISO19115/19139 de metadatos geográficos y CSW 2.0 entre otros estándares. Durante la ejecución del proyecto tecnológico de Casa Montero se han utilizado dos versiones del software Geonetwork: la v2.2.0 del 4 de abril de 2008 y, actualmente, la v2.4.2 de 12 de octubre de 2009. El movimiento de las listas de distribución donde se da soporte al programa ha sido contenido (Fig. 4). La respuesta a los mensajes incluidos en la lista de correo de usuarios ha sido muy lenta.

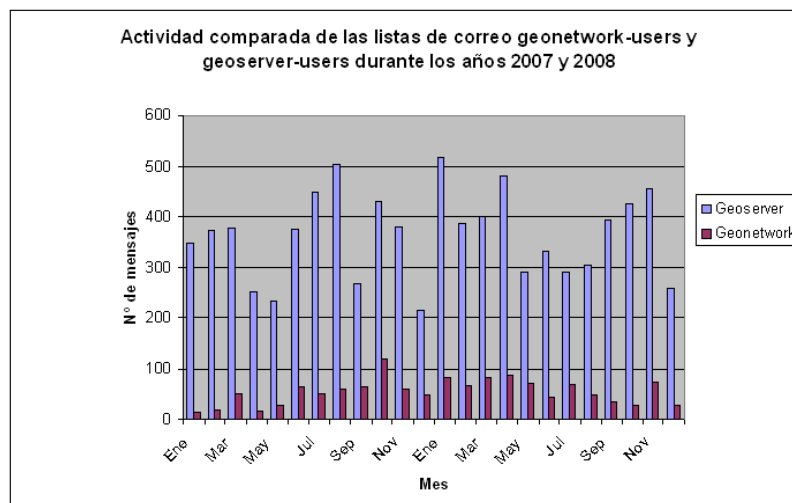


Figura 4: Actividad de las listas de correo geonetwork-users y geoserver-users en 2007 y 2008

Según la descripción que se realiza en *Sourceforge*, Geoserver es un servidor escrito en Java que permite compartir y editar datos geoespaciales. Durante la ejecución del proyecto tecnológico de Casa Montero se han utilizado cinco versiones del software: la v1.4.1 del 16 de febrero de 2007, la v1.6.3 de 8 de agosto de 2008, la v1.7.0a del 27 de octubre de 2008, la v1.7.5 del 10 de junio de 2009, y, actualmente, la 1.7.7 del 14 de octubre de 2009. El movimiento de las listas de distribución donde se da soporte al programa ha sido generoso (Fig. 4). La respuesta a los mensajes incluidos en la lista de correo de usuarios ha sido rápida en la mayoría de los casos.

Para distribuir la información temática se ha programado un servicio Web basado en ROA (*Resource Oriented Architecture* [6]), que es una guía de buenas prácticas para construir servicios Web conforme a la arquitectura de sistemas distribuidos hipermedia REST [5]. En este sentido el servicio Web de información temática es un servicio *RESTful*.

Los clientes se comunican con este servicio Web mediante una interfaz estándar HTTP (*HyperText Transfer Protocol*) que les da acceso al conjunto de recursos expuestos, identificados por URIs (*Uniform Resource Identifiers*), intercambiando representaciones de los mismos en formato XML (*eXtensible Markup Language*) e incluso en formatos CSV y PDF.

Las propiedades principales de los servicios Web *RESTful* son las siguientes: el servicio Web expone al menos un URI para cada recurso de información (*addressability*). La representación solicitada incluye URIs de otros recursos, de forma que el cliente pueda descubrir nueva información sin más que seguir los enlaces (*connectedness*). Cada petición al servicio Web ocurre aislada de otras; cuando el cliente hace una petición, ésta incluye toda la información necesaria para poder generar una respuesta (*statelessness*). Además las operaciones que pueden realizarse son las mismas para cualquier recurso (*uniform interface*) y corresponden a los métodos HTTP.

Este servicio Web ofrece acceso público a toda la información temática del proyecto Casa Montero. Basta conocer el URI de un recurso (por ejemplo el del propio yacimiento), recuperar su representación XML con una simple petición HTTP GET y seguir los URI contenidos en esta representación para obtener del mismo modo otros recursos y a partir de estos otros y así sucesivamente. En definitiva es posible recuperar siguiendo los enlaces toda la información del proyecto.

Aquellos usuarios que estén autorizados pueden incluso actualizar la información o crear nuevos recursos mediante peticiones HTTP DELETE, PUT y POST.

Para implementar este servicio Web se ha utilizado la versión 2.1.10 del entorno *open source* Apache Cocoon (<http://cocoon.apache.org/>). Se consideró muy apropiada la flexibilidad de este entorno para definir secuencias de procesos (*pipelines*) asociadas a URIs. Estas secuencias de procesos describen cómo generar documentos XML a partir de una o varias fuentes, validarlos y transformarlos para ofrecer al usuario la respuesta solicitada. A pesar de su largo recorrido, sus orígenes se remontan a 1998. Cocoon es un proyecto paralizado desde mediados de 2008. También se ha utilizado el motor de transformación XSLT Saxon (<http://saxon.sourceforge.net/>), en su versión *open source*, que es sin duda el motor XSLT más completo y mejor valorado.

El servicio Web evalúa la petición del usuario y envía consultas XQuery (algunas se componen dinámicamente) a la base de datos eXist, para obtener la información deseada. A la respuesta de eXist, el servicio Web añade nueva información y el documento agregado se transforma para generar la respuesta final. Como el diálogo entre el servicio Web y la base de datos se define en XQuery sobre HTTP, es posible acoplar el servicio Web a otras bases de datos.

2.2 Interfaz de usuario Web (WUI, *Web User Interface*)

La interfaz de usuario es una interfaz Web rica definida en HTML, JavaScript y AJAX.

La visualización de las entidades geográficas del sistema (como los pozos mineros) ha sido desarrollada utilizando la librería para JavaScript OpenLayers, con la que se ha compuesto un visor básico de mapas.

Con respecto a la información temática, la interfaz permite la exploración de la información navegando por los enlaces que relacionan los recursos, la composición de expresiones de consulta o filtro y, para usuarios autorizados, la edición de la información existente y la creación de nuevos recursos.

Para ello la interfaz de usuario Web hace peticiones al servicio Web y transforma éstas en HTML y JavaScript utilizando hojas de transformación XSLT. Los formularios de edición y de consulta están descritos en el lenguaje estándar W3C XForms y son dinámicamente transformados a HTML y JavaScript con AJAX, de forma que puedan ser interpretados por cualquier navegador sin necesidad de instalar *plugins*. La información es editada en el navegador directamente en formato XML.

La interfaz de usuario Web además ejerce una labor de integración (*mash-up*) de los servicios geográficos OGC (de arquitectura SOA) con el servicio Web temático (de arquitectura ROA). El usuario puede localizar un recurso en el mapa y pulsar sobre él para solicitar su información temática o, inversamente, obtener la posición del recurso en el mapa a partir de la información temática. Esto muestra que, lejos de incidir en la tradicional pugna (casi ideológica en ocasiones) entre ambas arquitecturas, es posible su uso complementario.

Para implementar este servicio Web se ha utilizado también Apache Cocoon y Saxon, además del entorno Orbeon XForms (<http://www.orbeon.com/>), en particular su versión 3.7.1, para convertir dinámicamente los formularios XForms como se ha comentado más arriba.

El uso de XQuery como lenguaje de consulta de fuentes de datos XML, REST como arquitectura para distribuir la información como servicio Web y XForms como lenguaje de definición de formularios ha recibido el acrónimo XRX (<http://en.wikibooks.org/wiki/XRX>). La aplicación Web en su conjunto es un ejemplo de arquitectura XRX.

Tanto el servicio Web como el interfaz de usuario Web son aplicaciones desplegadas sobre un motor de *serv/lets* Apache Tomcat (<http://tomcat.apache.org/>). La versión de Tomcat instalada es la 6.0.18.

3. CONCLUSIÓN

Dos años después del comienzo del proyecto y a la vista de los resultados obtenidos tan sólo podemos afirmar con rotundidad que el resultado de la experiencia con software libre y de código abierto es totalmente satisfactorio. Existía la posibilidad de haber empleado licencias comerciales para desarrollar la implementación de la versión 1.0 de SILEX (Sistema de Información Locacional En XML). Sin embargo su costo hubiera sido muy superior al empleado en la adquisición de las licencias de código abierto y, sobre todo, la dependencia tecnológica con una serie de firmas comerciales hubiera sido absoluta. Probablemente, la utilización de licencias privativas que ocultan su código fuente sea incompatible con el acuerdo *Open Access*.

4. BIBLIOGRAFÍA

- ◆ [1] VICENT GARCÍA, J. M.; CRUZ BERROCAL, M.; TORRE SÁINZ, I. D. L.; FERNÁNDEZ FREIRE, C.; FRAGUAS BRAVO, A.; LÓPEZ ROMERO GONZÁLEZ DE LA ALEJA, E.; URIARTE GONZÁLEZ, A. y WALID SBEINATI, S. (2004): “Laboratorio de Teledetección y Proceso Digital de Imagen”. En (ed.) J. C. Martín de la Cruz y A. M. Lucena Martín: *Informática Aplicada a la Investigación y Gestión Arqueológicas*. Universidad de Córdoba. Córdoba.
- ◆ [2] FRAGUAS BRAVO, A. (2009): *El Arte rupestre prehistórico de África nororiental: nuevas teorías y metodologías*. Consejo Superior de Investigaciones Científicas. Madrid.
- ◆ [3] FERNÁNDEZ DE LABASTIDA, J.M. (2006): *Open Access* [<http://hdl.handle.net/10261/2299>]
- ◆ [4] DÍAZ-DEL-RÍO, P.; CONSUEGRA, S.; CASTAÑEDA, N.; CAPOTE, M.; CRIADO, C.; BUSTILLO, M.A. AND PÉREZ-JIMÉNEZ, J.L. (2006): “The earliest flint mine in Iberia”, *Antiquity* 80 (307), [<http://antiquity.ac.uk/ProjGall/diazdelrio/index.html>]
- ◆ [5] FIELDING, R.T. (2000) *Architectural styles and the design of network-based software architectures*. University of California Press, Irvine. [<http://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm>].
- ◆ [6] RICHARDSON, L. Y RUBY, S. (2007): *RESTful Web Services*. O’Reilly