

Título: La Gestión de la Información en el Proyecto Arqueológico El Caño (PAEC)

Autores:

Alfredo Fernández-Valmayor. valmayor@ucm.es CIAI y Facultad de Informática UCM.

Ana Fernández-Pampillón. apampi@filol.ucm.es Facultad de Filología UCM.

Julia Mayo. mayo.julia@gmail.com CIAI.

Mercedes Guinea. mguinea@ucm.es Facultad de Geografía e Historia UCM.

Carlos Mayo. mayotorne@gmail.com CIAI.

Miguel Angel Hervás. mangelhervas@yahoo.es CIAI y Baraka Arqueólogos.

Jesus Herrerin. jesus.herrerin@uam.es CIAI y Facultad de Biología UAM.

CIAI: Centro de Investigaciones Arqueológicas del Istmo, Panamá

UCM: Universidad Complutense de Madrid, España

UAM: Universidad Autónoma de Madrid, España

Eje temático: Arqueología

Palabras clave: Arqueología, bases de datos, objetos digitales, modelo de datos, facetas.

Resumen:

En esta ponencia presentamos ODA, el sistema de gestión de la información utilizado en el Proyecto Arqueológico El Caño. Este sistema es una aplicación web, implementada con software abierto y tecnologías estándar. En ella se introduce el concepto de “objeto digital”, definiéndolo como el conjunto de documentos que documentan un “objeto real”. Además, cada “objeto digital” se caracteriza mediante un conjunto de atributos. Así el “objeto digital” y sus atributos se convierten en el elemento organizador de la base de datos documental de ODA. La construcción de este repositorio implica tres tareas: 1) utilizar la noción de “objeto digital” para agrupar los documentos multimedia que describen algún aspecto de los “objetos reales” del proyecto, 2) asignar atributos a los “objetos digitales” para posibilitar su clasificación y análisis, 3) construir de forma incremental el “modelo de datos” del repositorio. Este “modelo de datos” debe ser también un modelo del dominio, o de la investigación realizada sobre el dominio, y es la base para guiar el análisis de la información en el repositorio. En la primera sección de la ponencia se analiza la gestión de la información en un proyecto arqueológico. En la segunda se trata de los requisitos que debe satisfacer ODA y de su diseño. En la tercera sección se presentan las líneas generales de su arquitectura e implementación. Finalmente, en la última sección, se hace una valoración de ODA y sus posibilidades en el contexto de los sistemas de gestión de la información arqueológica.

0.- Introducción

El objetivo de esta ponencia es describir y valorar el sistema de gestión de la información utilizado en el Proyecto Arqueológico El Caño (PAEC). Este sistema es una aplicación web desarrollada utilizando software abierto y tecnologías estándar. Esta aplicación introduce el concepto de “objeto digital” como el conjunto de documentos que informan de un “objeto real”. Además, el “objeto digital” se caracteriza mediante un conjunto de atributos. Así, este “objeto digital” es el principal elemento organizador de la base de datos documental que

constituye el núcleo del sistema de gestión. El sistema recibe el nombre de ODA (Objetos Digitales de Aprendizaje) porque en su origen estuvo vinculado al desarrollo de entornos informatizados para la enseñanza (Fernández-Valmayor et al., 2011). La versión que aquí se presenta, ODA 2.5, ha sido desarrollada por los investigadores del PAEC, en colaboración con el grupo de investigación *Implementation of Language-Driven Software and Applications* (ILSA) de la UCM (Universidad Complutense de Madrid), y con la empresa *Varadero Software Factory*, esta última *spin-off* del grupo de investigación de la UCM.

La gestión de la información de un proyecto arqueológico, utilizando las funciones que proporciona ODA, se basa en la construcción de un repositorio de datos y documentos. Este repositorio es accesible en la web y en él se almacena, clasificada bajo múltiples perspectivas, toda la documentación generada por el proyecto que es considerada relevante. La construcción de este repositorio implica tres tareas: 1) utilizar la noción de “objeto digital” para agrupar los documentos multimedia relativos a cada uno de los “objetos reales” del proyecto, 2) asignar atributos a los “objetos digitales” para posibilitar su caracterización, clasificación y análisis, 3) construir de forma incremental el “modelo de datos” del repositorio. Este “modelo de datos” debe ser también un modelo del dominio, o de la investigación realizada sobre el dominio.

1) Objeto digital.

En ODA, un “objeto digital” representa y describe nuestro conocimiento sobre un “objeto real”. Inicialmente, un “objeto digital” es el conjunto de documentos digitales -archivos multimedia- generados por el estudio del “objeto real”. Si ponemos por caso una vasija arqueológica, el objeto real es el “objeto físico”, la vasija, y el objeto digital el conjunto de informes, fotografías, dibujos técnicos y cualquier otro archivo multimedia que se haya generado a partir del estudio de esa vasija. Cabe señalar que los objetos digitales también pueden representar y describir, y de hecho en el Proyecto Arqueológico El Caño así lo hacen, “objetos conceptuales”. Los objetos conceptuales son otro tipo de objetos reales. Por ejemplo, son objetos conceptuales las unidades estratigráficas identificadas durante el proceso de excavación. En este caso el objeto digital está compuesto por el conjunto de documentos, croquis, fotografías y textos, que identifican y explican cómo es dicha unidad estratigráfica. Otros objetos conceptuales más elaborados pueden ser, un conjunto de referencias bibliográficas, los episodios del ritual funerario o el propio ritual funerario.

En una primera aproximación se puede decir que ODA es un repositorio de objetos digitales, o lo que es lo mismo, una base de datos documental en la que la abstracción “objeto digital” sirve como elemento organizador primario.

2) Atributos de los objetos digitales.

Para completar el registro de los objetos en el repositorio ODA, cada objeto digital se describe y caracteriza asignándole un conjunto de datos o pares “atributo: valor”. A través de los atributos se ingresan en el sistema las características más relevantes de los objetos. El conjunto de valores de estos atributos podemos verlos como el conjunto de coordenadas que sitúan al objeto en el espacio físico y conceptual de nuestra investigación. La posterior navegación en este espacio, las rutas del conocimiento dependerán de la forma en que se relacionen entre sí estos atributos, o lo que es lo mismo, dependerán del “esquema de clasificación” o “modelo de datos” del repositorio, que veremos en el apartado siguiente.

Poniendo por caso las cerámicas de El Caño, para situar una vasija en el espacio físico

podemos usar estos tres pares atributo: valor, 1) “alto mm: 180”, 2) “diámetro mm: 520”, 3) “cota del hallazgo m: -3.25”, y para situarla en un espacio conceptual (tipología cerámica) podemos usar los siguientes cuatro pares: “estilo cerámico: conte temprano”, “vajilla: conte rojo”, “forma: cuenco”, “variedad: hemisférico”.

Debemos señalar que asignar a un objeto los pares atributo: valor, necesarios para describir sus características físicas suele ser inmediato, pero asignar los pares atributo: valor, necesarios para describir adecuadamente los objetos con características conceptuales, como pueden ser, por poner un ejemplo, los episodios de los ritos funerarios, requiere hacer un análisis y un esfuerzo de abstracción importante.

3) “Modelo de datos”, o “esquema de clasificación”, construido de forma incremental durante el desarrollo del proyecto.

ODA, como la mayoría de los sistemas avanzados de gestión de la información, tiene en su base un formalismo de representación de la información que permite construir un modelo de datos del repositorio basado en las características conceptuales del dominio (Rodríguez-Castro et al., 2010; Hearst, 2006a, 2006b). Utilizando ODA podemos formalizar mediante atributos las características más importantes de los objetos digitales y construir con ellos una estructura que modele el dominio de conocimiento al que pertenecen los objetos. Esta estructura se construye tomando como base los atributos formalizados, las relaciones entre ellos, las propiedades de estas relaciones y finalmente los vocabularios controlados, es decir, la lista de valores que puede tomar cada atributo.

Es importante señalar que, con estos elementos, construir el modelo de datos de un dominio de conocimiento moderadamente complejo no es en absoluto una tarea sencilla (Giess et al., 2007). No obstante, el esfuerzo es rentable ya que un modelo de datos apropiado es la base para hacer intuitiva la navegación, el acceso y el análisis de la información almacenada en el repositorio. Además, los usuarios pueden aprender sobre los temas y conceptos del dominio de conocimiento, mientras navegan por él, lo que hace del repositorio una herramienta para la educación muy importante.

ODA, un modelo de datos cuya estructura se puede reconfigurar. En los sistemas de gestión de la información, las tareas de diseño y construcción del modelo de datos preceden a las tareas de crear y poblar con datos el repositorio (Schwarzbach et al., 2014). Introducir posteriormente cambios en el modelo no suele estar previsto, o no es fácil

“Of course, it is possible to change the database as the excavation progresses, but even quite minor changes can be time consuming and naturally restrictive. It is unlikely that the field archaeologists will want to contact the database technician every time they want to add a new criterion or field to the database”. (Eve & Hunt, 2008).

Con ODA la situación es diferente, en cualquier etapa de la construcción de un repositorio ODA es posible modificar y reconfigurar el modelo de datos. Es decir, a medida que la excavación progresa, se pueden ensayar distintos atributos y distintas relaciones entre atributos –distintas configuraciones del modelo- para adecuarlo mejor a nuestra investigación sobre el sitio arqueológico. Dicho de forma más general, para adecuarlo mejor al dominio de conocimiento.

El PAEC se ha beneficiado de estas características de ODA. La primera etapa de la

construcción del repositorio del PAEC comenzó con un modelo de datos muy sencillo con unos pocos atributos relativos a las características físicas de los hallazgos. Después, como puede comprobarse accediendo al repositorio, el modelo ha evolucionado a un modelo más completo y complejo, y está previsto que esta evolución continúe para el modelo pueda mostrar nuevos aspectos de la cultura Coclé (Mayo, J. & Mayo, C., 2017).

El resto de esta ponencia está organizada de la siguiente forma. En la primera sección se describen brevemente las peculiaridades que tiene la gestión de la información de un proyecto arqueológico como el PAEC. En la segunda se concreta lo que podríamos describir como un amplio consenso de los investigadores sobre los requisitos que debe satisfacer una aplicación informática para la gestión de la información de un proyecto arqueológico, y a continuación, y tomando como base el diagrama UML, se describe el diseño de ODA. En la tercera sección se presentan, en forma esquemática, las líneas generales de la arquitectura e implementación de ODA. Finalmente, en la última sección y a modo de conclusión, se hace primero una valoración de ODA en el contexto de los sistemas de gestión de la información comparables con él, se trata después lo que ha supuesto y supone para el PAEC utilizar ODA y por último se hace una previsión de posibles mejoras de ODA.

1.- La gestión de la información en proyectos arqueológicos

En los proyectos arqueológicos como el PAEC, las operaciones de campo producen una gran cantidad de material y por ende de datos y documentos. El imprescindible trabajo de laboratorio multiplica el número de documentos y de datos que los investigadores deben considerar en su trabajo post-procesual que a su vez genera nuevos documentos. En el caso del trabajo de campo, esta documentación consiste en el registro de artefactos, restos humanos y contextos. Las herramientas usadas para recoger esta información son fichas-protocolo de unidades estratigráficas, fichas-protocolo de restos óseos humanos e inventarios, una videocámara y cámara fotográfica y una estación total. En el caso del trabajo de laboratorio, lo más habitual es que los datos se registren en hojas de cálculo, bases de datos e informes. Por lo tanto, en la actualidad en un proyecto arqueológico toda la documentación generada es digital, o digitalizable, lo que en principio la debería hacer que todo el material fuese manejable y accesible.

Toda la información generada es importante para la investigación y potencialmente muy útil para la educación y la difusión del patrimonio cultural, por lo que cabe preguntarse, como lo hicieron en su momento Hochstetter et al. (2011), por qué en general los datos y documentos de los proyectos arqueológicos no están disponibles en sistemas de información públicamente accesibles, o no lo están en la medida que actualmente sería posible, ya que al estar en soporte digital su difusión debe de ser relativamente sencilla. Estos autores señalaron con acierto tres factores que frenaban la difusión de los resultados de los proyectos arqueológicos, pero también señalaron otros factores que podían hacer cambiar la situación:

" First factor [...] advocates of keeping data confidential cite issues related to the preservation of the archaeological record [...] One alternative strategy is to make information about the record more public since knowledge and awareness contributes to stewardship "

" The second factor [...] follows from the tradition of intellectual property [...] but open access to primary information do not preclude appropriate citation [...] Creative Commons provide an excellent means of releasing data that can be shared as long as proper attributions are made. "

" The third consideration about the sharing of data is the technological problems associated with information dissemination [...] but, a number of recent developments have made the sharing of archeological information vastly simpler. [...] Technologically, we are now in an excellent position to make the archaeological record freely available for study to researches Worldwide, and to the general public." (Hochstetter et al., 2011)

Ciertamente, estos factores reflejan lo que ha sido la realidad de muchos proyectos arqueológicos y el cambio paulatino que se está produciendo (Eve & Hunt, 2008; Schulze et al., 2012; Geser, & Gonzalez-Perez, 2017), aunque también debemos señalar que en la actualidad todavía persiste una relativa escasez de bases de datos con información arqueológica públicamente accesibles, o lo que sería mas exacto, una relativa escasez de sitios web que den acceso a bases de datos en correcto estado de funcionamiento. Bases de datos que estén bien mantenidas, sin enlaces rotos, que ofrezcan una navegación intuitiva e información actualizada de los proyectos arqueológicos pasados o en curso. En nuestra opinión, la situación actual puede deberse a motivos que van más allá de los enumerados por Hochstetter et al. en 2011. Estos motivos estarían más bien relacionados con la dificultad del trabajo de gestión y difusión de la información arqueológica en si mismo. En efecto, los proyectos arqueológicos son proyectos complejos y multidisciplinarios. Proyectos en los que la dificultad que implica gestionar información heterogénea, y mostrarla en un sitio web bien organizado y puntualmente actualizado, podría estar siendo subestimada. Esta situación lleva a que los proyectos arqueológicos no destinen para este fin los recursos necesarios y como consecuencia la calidad de la información accesible en la web puede estar quedando lejos de lo que hoy se podría esperar.

Entre los motivos que hacen que la gestión de la información en un proyecto arqueológico sea un tema complejo podríamos citar las siguientes:

- Muchos datos y documentos de diferentes especialidades confluyen en el estudio del mismo área o sitio arqueológico. Sin embargo, no siempre está claro, dentro del equipo de investigación, quienes, cómo, y con que medios informáticos, deben integrar y hacer accesible toda esa masa de información heterogénea.
- Muchos datos se toman manualmente en el campo: cuadernos de registro, notas, etc. y aunque después se digitalizan nunca llegan a ser debidamente evaluados y clasificados y por lo tanto al final resultan difícilmente accesibles para la comunidad científica, y a veces hasta para los propios miembros del proyecto.
- Datos e informes pueden estar dispersos en diferentes herramientas informáticas y computadoras. Estos documentos pueden ser difíciles y laboriosos de organizar o integrar en una estructura común, o en un relato, coherente y compartido.
- Gestionar bien la información requiere seleccionar la información relevante, analizarla y clasificarla desde diferentes puntos de vista, o taxonomías, que no siempre son fáciles de anticipar, especialmente, si en los proyectos no se ha programado el tiempo y los recursos humanos y materiales necesarios para ello.

La solución a estos problemas iría en la dirección de planificar mínimamente las tareas de gestión de la información. Tareas que debe realizar de forma consensuada el equipo de investigación que además debe contar con el adecuado soporte informático. Ejemplo de estas

tareas podrían ser las siguientes:

1. Seleccionar los datos y documentos más relevantes para los objetivos de la investigación. Digitalizarlos. Hacerlos accesibles. Establecer una política de gestión de los permisos de acceso, es decir, controlar quién puede acceder a que información y cuando.
2. Organizar y estructurar la información de acuerdo con las principales líneas de investigación del proyecto y así facilitar el acceso a la misma y su comprensión.
3. Prueba/error y consenso en el grupo de investigación para decidir cómo organizar la información, es decir, cómo será el modelo de datos o esquema de clasificación que utilizaremos como referencia.
4. Uso de taxonomías ya consolidadas o propuesta de modificaciones/extensiones a las mismas.
5. Definición de nuevas taxonomías y vocabularios controlados para clasificar la información de forma coherente con el desarrollo del proyecto.
6. Para facilitar este trabajo, es importante disponer de las herramientas informáticas apropiadas que faciliten estas tareas.

El objetivo de este esfuerzo es obtener un corpus de información clasificada de forma consensuada y que sirva de base para ejecutar y dar a conocer las directrices y resultados del proyecto de investigación. Este corpus debe ser accesible públicamente y, a medida que avancen las investigaciones, actualizable por todos los miembros del proyecto que tengan atribuciones para ello. Así mismo, se debe disponer de medios informáticos que permitan seleccionar que parte de la información se hace accesible en la web a otros grupos de investigadores y al público en general. Finalmente, es importante señalar que no hay que confundir los objetivos de un sistema de gestión de la información y publicación *online* de la evolución inmediata de un proyecto, con las publicaciones científicas en revistas especializadas y congresos. Éstas últimas tienen una escala de tiempo completamente diferente y también un contenido y un objetivo distintos.

2.- Requisitos y diseño de la aplicación ODA

Antes de presentar el diseño de la aplicación ODA vamos a hacer una revisión de los requisitos más relevantes exigidos a las aplicaciones informáticas para que puedan dar el adecuado soporte a un proyecto arqueológico. Para hacer esta síntesis hemos considerado los requisitos identificados en el Proyecto Arqueológico El Caño. Además, hemos revisado las propuestas formuladas en los siguientes proyectos: EISP, *Easter Island Statue Project* (Hochstetter et al. 2011); ARK, *Archaeological Recording Kit* (Eve y Hunt, 2008); OpenInfRA, *Open Information System for Research in Archeology* (Schulze et al., 2012) y ARIADNE, *Advanced Research Infrastructure for Archaeological Dataset Networking in Europe* (Geser, & Gonzalez-Perez, 2017).

Es importante destacar que en general los requisitos identificados en estos proyectos coinciden esencialmente en lo que se refiere a la recogida de datos en el campo, en el laboratorio y a la gestión post-procesual de la información arqueológica, pero, en temas de interoperabilidad, integración aplicaciones y servicios web especializados, las exigencias pueden ser muy distintas. Respecto a estos últimos requisitos puede variar entre proyectos como ARIADNE, que centran todo su esfuerzo en la interoperabilidad y extracción de la información en sistemas de bases de datos previamente existentes, y proyectos como ARK que explícitamente deja estos

temas casi fuera de sus objetivos, para centrarse en la recogida y clasificación coherente de la información arqueológica: datos de campo, datos de laboratorio y documentos post-procesuales. En nuestra opinión, esta variabilidad depende en gran medida de los recursos materiales disponibles en cada proyecto. Siendo la principal razón para excluir todos, o algunos, de los objetivos con exigencias más tecnológicas su elevado coste de desarrollo. Este coste estaría ligado a las complejas arquitecturas y estándares y al mayor esfuerzo de desarrollo requerido para lograr la interoperabilidad e integración efectiva de aplicaciones (Henze et al. 2013).

Agrupamos los requisitos identificados en cuatro bloques, siendo el último bloque en el que más diferencias entre proyectos se observan:

a) Almacenamiento y organización de la información

- Repositorio (base de datos) accesible mediante una aplicación web.
- Soporte para almacenar los datos primarios de una excavación (sondeo o exploración) de forma eficiente, bien estructurada y consistente.
- Soporte para todo tipo de archivos multimedia.
- Soporte para recogida de datos de un amplio rango de métodos de trabajo de campo, datos cuantitativos y cualitativos.
- Soporte para adaptar cualquier sistema de toma de datos que estuviese basado en papel.
- Repositorio de carácter acumulativo. Soporte para que los investigadores puedan contribuir en cualquier etapa del proyecto con nuevos datos y análisis, añadiendo/modificando datos y resultados anteriores.

b) Modelo de datos/esquema de clasificación

- Homogéneo y flexible. Módulos predefinidos pero adaptables a las necesidades específicas de cada investigación.
- Permite integrar los documentos más relevantes utilizados y/o producidos por la investigación.
- Permite organizar la información de acuerdo con la metodología y directrices de la investigación.
- Permite definir escalas discretas mediante vocabularios controlados.
- Permite remodelar el esquema de clasificación, incorporando nuevos campos, nuevas relaciones y nuevos ítems a los vocabularios especializados.

c) Control de acceso y servicios a investigadores y público.

- Gestión de permisos y derechos de autor.
- Diseminación temprana de resultados de investigación
- Difusión cultural y usos educativos. La información debe ser accesible a las comunidades locales y a todo el público interesado.
- Museo virtual y complemento digital a un posible museo real.

d) Base tecnológica

- Desarrollo basado en tecnologías estándar y software libre y abierto
- Versiones *online* y *offline* para poder trabajar sin conexión a internet
- Procedimientos de sincronización de versiones
- Soporte para diversos idiomas

- Integración de aplicaciones especializadas p.e. WebGIS, 2D, 3D, web services, ...
- Posibilidad de indexar y situar espacialmente la información sobre los hallazgos.
- Exportación e importación de datos basada en XML
- Soporte para mapear el modelo de datos a SKOS (atributos con vocabularios controlados) y a otros estándares.
- Soporte específico para mapear el modelo de datos a CIDOC-CRM (*International Committee for Documentation - Conceptual Reference Model*). “*The CIDOC Conceptual Reference Model provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation*” (CIDOC-CRM, 2017).
- Métodos de búsqueda y recuperación de información avanzados

Para desarrollar una herramienta que cumpliera al menos con el núcleo más importante de estos requisitos (básicamente requisitos de los tres primeros bloques), el PAEC llegó a un acuerdo con el grupo de investigación ILSA de la Universidad Complutense de Madrid (UCM) y *Varadero Software Factory*, y el resultado de este acuerdo ha sido la aplicación ODA 2.5. Para hacer este desarrollo inicialmente se dejaron al margen los requisitos del cuarto bloque salvo el desarrollo con software libre, tecnologías estándar y versiones *online* y *offline*. Posteriormente se han probado módulos para exportar contenidos en diversos formatos y dar soporte a varias lenguas.

Para construir esta herramienta se utilizó como punto de partida “Chasqui” (Guinea, 2004), una herramienta basada también en tecnologías estándar y software libre, y que fue desarrollada por el mismo equipo para informatizar varios museos docentes de la UCM con el objetivo de potenciar su utilización en la enseñanza en conexión con el campus virtual de la UCM (Sierra et al. 2006; Guinea et al. 2009).

El diseño de ODA (Objetos Digitales de Aprendizaje)

La decisión más importante del diseño del sistema ODA es la de no modelar una base de datos específica, si no, una base de datos genérica en la que se almacenan tanto los datos de la excavación como el modelo de datos del repositorio (podríamos decir que el modelo de datos es en si mismo un resultado del proyecto). Además, se utiliza el concepto de “objeto digital” como metáfora básica para organizar y agrupar los documentos del proyecto. Es decir, la actividad básica trata de identificar y asignar la documentación que le corresponde a cada uno de los objetos físicos, o conceptuales, que son objetivo del proyecto. En el diagrama UML (Universal Modelling Language) puede verse como entorno a esta abstracción central se desarrollan cuatro componentes:

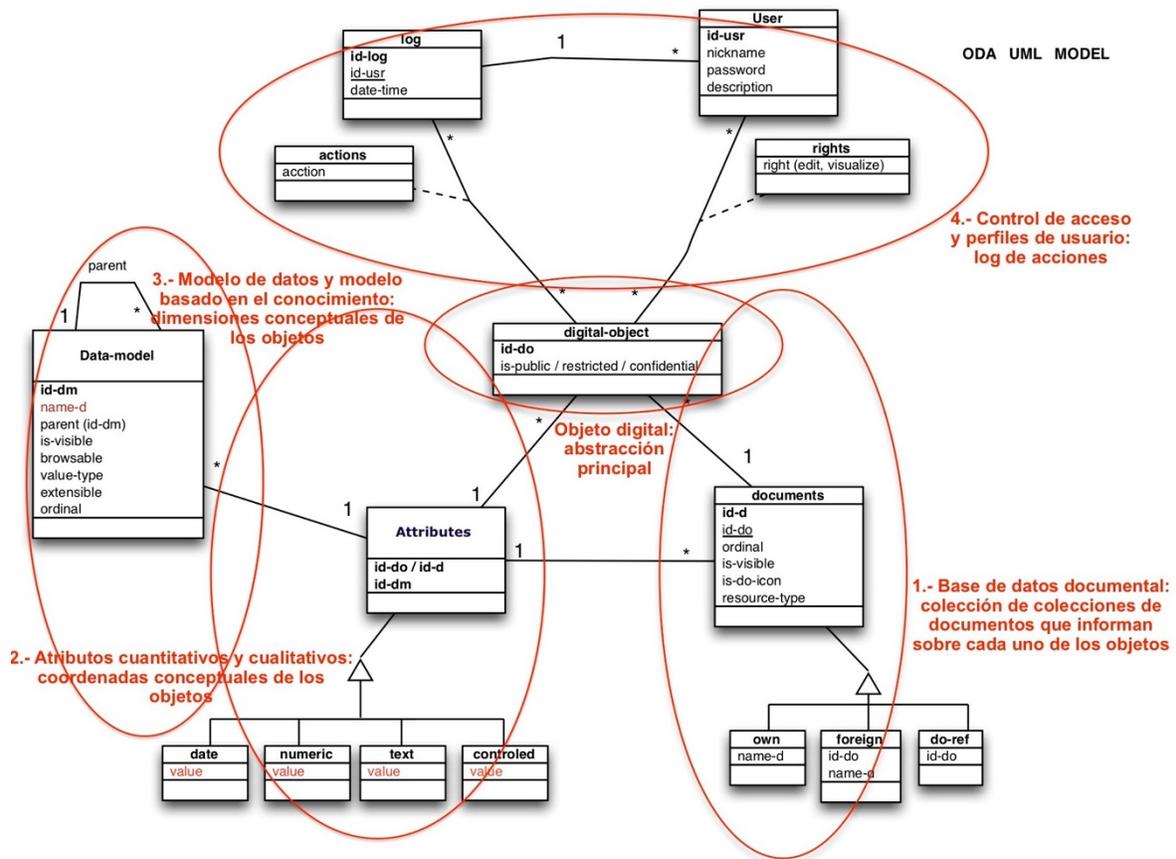


Fig. 1: Diseño de ODA en UML. (Universal Modelling Language)

1.- Base de datos documental.

En una primera aproximación, ODA es simplemente un repositorio de “objetos digitales” y cada objeto digital es un “contenedor” con todos los documentos multimedia: texto, fotos, gráficos, video o audio, que informan sobre un objeto real previamente identificado. Este contenedor, el objeto metafórico, puede implementarse como una carpeta del Sistema de Archivos (FS) del Sistema Operativo (OS) del servidor o computador anfitrión (ver Fig. 6). Esta carpeta contendrá todos los archivos digitales que informan del objeto real o referencias válidas a los mismos.

Cada uno de los archivos digitales que contiene un objeto digital se considera un “documento-propio” del objeto, siempre y cuando éste se refiera directamente al objeto, es decir, que el archivo digital sea una foto, un video, un audio, un dibujo, o un informe del objeto registrado. Por decirlo así, el objeto digital es propietario de estos documentos, razón por la cual están en su carpeta y no en otra. Pero los documentos que son parte de un objeto digital también pueden ser “una referencia” a otra documentación no propiamente del objeto pero que, por alguna razón, el investigador considera importante incorporar al objeto digital. Estas referencias deben considerarse tanto en el plano semántico y como en el plano sintáctico. En el plano conceptual, y dependiendo de los objetos y documentos relacionados, pueden ser referencias *part-of*, *is-a*, *instance-of*, etc., según la semántica que queramos atribuir al modelo de datos del repositorio. De otro lado y desde un punto de vista meramente sintáctico estas

referencias pueden ser de tres tipos: a) a un “documento-propio” de otro objeto del repositorio; b) a otro objeto digital del repositorio (el contenedor y todos sus documentos) y c) al contenido de una *url* externa a ODA. En los dos primeros casos hablamos de “documento-referencia-a-documento-existente”, y de “documento-referencia-a-objeto-digital-existente”. En el tercer caso hablamos simplemente de “documento-referencia-externa”. En este último caso la relación entre documentos no es parte de la estructura del modelo de datos que informa el repositorio (por ello no está representada en el esquema UML de la Fig.1), debe considerarse una relación meramente informativa y el sistema no garantiza su permanencia.

2.- Atributos.

Además de ser un conjunto de documentos, un objeto digital se caracteriza mediante un conjunto de datos o pares atributo: valor. Conceptualmente estos pares atributo: valor pueden referirse tanto a datos cuantitativos como a datos cualitativos del objeto. Como se representa en la Fig.1, en ODA los valores de los atributos pueden ser de cuatro tipos sintácticos: 1) texto libre, 2) fecha, 3) numérico y 4) controlado.

1) El tipo texto libre permite describir informalmente las características de un objeto y cualquier otro tipo de circunstancia especial del mismo. Disponer de este tipo de atributo en ODA nos permite construir modelos de datos híbridos, es decir, modelos que mezclan descripciones informales en lenguaje natural con datos formalizados.

2) y 3) Los tipos numérico y fecha son datos objetivos y cuantitativos del objeto. Normalmente estos atributos son inmediatos de identificar y tienen una interpretación unívoca.

4) El tipo controlado. En ODA, la construcción de un modelo de datos formal se basa principalmente en este tipo de atributos. Los valores que pueden tomar estos atributos se eligen entre los elementos de un “vocabulario controlado”, normalmente una lista de términos, o coordenadas cualitativas, que permiten situar al objeto con respecto a una dimensión (una propiedad significativa) de nuestro dominio de conocimiento. En palabras de la Oficina Nacional de Estándares de la Información USA:

“... controlled vocabularies that are used for the representation of content objects in knowledge organization systems including lists, synonym rings, taxonomies, and thesauri. The primary purpose of vocabulary control is to achieve consistency in the description of content objects and to facilitate retrieval” (ANSI/NISO Z39.19-2005, R2010).

El conjunto de atributos cuyo valor es de tipo controlado, y sus correspondientes vocabularios, son la base para definir subconjuntos de atributos del objeto digital que describan sus principales categorías semánticas, o ejes conceptuales, del dominio de conocimiento. En la literatura estos atributos reciben el nombre de “facetras” (y sus valores “términos”) si son mutuamente exclusivos y exhaustivos respecto al conjunto de objetos que queremos registrar en el repositorio (Denton, 2003; Vickery, 2008). Conceptualmente podemos decir que cada objeto digital ocupará una “posición” en nuestro espacio conceptual y tendrá como coordenadas los términos, o valores, de las facetras tomados de sus respectivos vocabularios controlados.

Hay que señalar que ODA también permite construir sistemas híbridos en el sentido de que no todos los atributos están formalizados ni todos los objetos del repositorio se caracterizan utilizando el mismo conjunto de atributos. En general, puede haber objetos que pertenecen a

distintos subespacios conceptuales y que por tanto se caracterizan mediante distintos subconjuntos de atributos. Por ejemplo, en el PAEC la gestión de los hallazgos arqueológicos y la gestión de la documentación del proyecto se describen mediante subconjuntos de atributos casi disjuntos.

No obstante, para unificar la gestión de un repositorio ODA, todos los objetos comparten por defecto dos atributos básicos: un atributo “descripción”, con la descripción informal del objeto en texto libre, y un atributo “tipo”, cuyo valor es un ítem de un “vocabulario controlado” con la lista de todas las categorías de objetos presentes en el repositorio.

Finalmente, ODA particiona el conjunto de atributos de un objeto digital en dos clases:

1. “Datos”, es decir, aquellos atributos que definen una propiedad del objeto, como pueden ser sus medidas, su forma y el material de que está hecho.
2. “Metadatos” (datos sobre los datos), son los atributos que expresan las circunstancias en las que se obtienen los datos, como puede ser el nombre del investigador responsable de tomar las medidas, o la fecha en que el objeto fue registrado y por quien.

Hay que señalar que puede haber atributos para los que la decisión entre si se consideran datos o metadatos, sea dependiente del punto de vista adoptado por el investigador. Por lo tanto, no siempre se puede establecer de forma objetiva la distinción datos/metadatos. Por ejemplo, puede depender de una consideración subjetiva del investigador establecer si la caracterización tipológica de una vasija cerámica debe registrarse como un dato, o como un metadato, de la misma.

3.- Modelo de datos

El modelo de datos, o esquema de clasificación, es la estructura que sirve de base para organizar los datos en un repositorio y que, por reproducir en cierta forma de la estructura del dominio, debe facilitar una gestión mas intuitiva y eficiente de la información almacenada en él (Giess et al., 2007). Normalmente, el diseño de un repositorio de datos comienza por la construcción del modelo de datos (Schwarzbach et al., 2014). Por ejemplo, cuando se diseña una base de datos siguiendo el modelo “entidad-relación”, todas las tablas, sus campos, tipos y las relaciones entre tablas que componen el modelo, deben estar definidos antes de comenzar a utilizar la base de datos. Esta estructura de tablas relacionadas entre sí, sus campos y tipos de datos, constituyen un modelo de la estructura del dominio para el cual se está construyendo la base de datos. Utilizando ODA el planteamiento es distinto. La construcción del modelo de datos puede realizarse de forma incremental. Partiendo de un modelo inicial sencillo, éste se puede ir ampliando y adecuando al dominio en etapas posteriores, es decir, al tiempo que la investigación avanza y ampliamos nuestro conocimiento del dominio.

En el esquema UML de la Fig. 1 puede verse que esto es posible porque ODA es una base de datos genérica (su dominio no es el de las excavaciones arqueológicas si no el de las bases de datos arqueológicas) y por tanto el modelo de datos se trata como cualquier otro conjunto de datos de la investigación. ODA utiliza un modelo relacional para modelar un sistema genérico en el que los investigadores pueden añadir elementos o cambiar la estructura del modelo de datos del dominio igual que añaden o cambian los datos arqueológicos. La metodología propuesta por ODA permite al investigador modificar un modelo inicial hasta convertirlo en

un modelo con la complejidad suficiente para representar las principales líneas de desarrollo de la investigación. Así pueden obtenerse, en una aproximación por etapas, resultados similares a los diseños de modelos facetados basados en el conocimiento propuestos por Broughton (2013), o las *faceted lightweight ontologies* propuestas por Giunchiglia et al. (2009).

La lista siguiente resume las acciones de ODA que tienen como objetivo construir o modificar el modelo de datos:

1. Identificar (asignar atributos) a los documentos (archivos y referencias a archivos) que forman los objetos digitales que representan los objetos reales (físicos y conceptuales) de nuestro dominio.
2. Definir y/o redefinir el conjunto de pares atributo-valor necesarios para describir los datos y metadatos de los objetos digitales del repositorio.
3. Definir y/o depurar los vocabularios controlados que constituyen el conjunto de valores que los atributos de tipo controlado pueden tomar.
4. Establecer y/o reconfigurar las relaciones jerárquicas entre los pares atributo: valor (relaciones padre-hijo), para agruparlos en una o más estructuras acordes con la estructura de nuestro dominio.
5. Establecer relaciones entre los objetos del repositorio a través de sus documentos. Los documentos de un objeto pueden ser documentos propios o referencias. Estas referencias nos permiten establecer relaciones no jerárquicas del tipo documento-documento o del tipo documento-objeto.
6. Decidir sobre la navegabilidad y accesibilidad, total o parcial, del modelo construido.

El tipo de modelo a construir con estas acciones es una decisión de los investigadores del proyecto. La aproximación seguida en el PAEC combina el modelo propuesto por Denton, basado en facetas, con el propuesto por Hearst que incorpora a las facetas múltiples jerarquías.

Según Denton (2003) los modelos de clasificación facetada se pueden entender como:

“... a set of mutually exclusive and jointly exhaustive categories, each made by isolating one perspective on the items (a facet), that combine to completely describe all the objects in question, and which users can use, by searching and browsing, to find what they need”

Hearst et al. (2006) propusieron ampliar estos modelos de clasificación facetada definiendo múltiples jerarquías de categorías “*Hierarchical Faceted Categories*” (HFC), cada una de las cuales desarrolla una faceta en sucesivas subfacetes. En esta aproximación podemos considerar cada jerarquía como una taxonomía que describe una perspectiva, o subespacio conceptual, dentro del espacio total.

“Rather than creating one large category hierarchy, build a set of category hierarchies each of which corresponds to a different facet (dimension or feature type) relevant to the collection to be navigated.”

Giunchiglia et al. (2009) proponen construir ontologías descriptivas estableciendo relaciones entre facetas y términos:

“facets and terms have real world semantics, namely, they are descriptive ontologies which are formed using part-of, is-a and instance-of relations”.

Hasta el momento, el modelo de datos construido para el repositorio del PAEC se ha desarrollado siguiendo la propuesta HFC en tres etapas sucesivas. Añadiendo en cada etapa nuevas jerarquías-taxonomías representando nuevos enfoques o perspectivas de la investigación. En la Fig. 2 se puede seguir la idea general de este desarrollo por etapas.

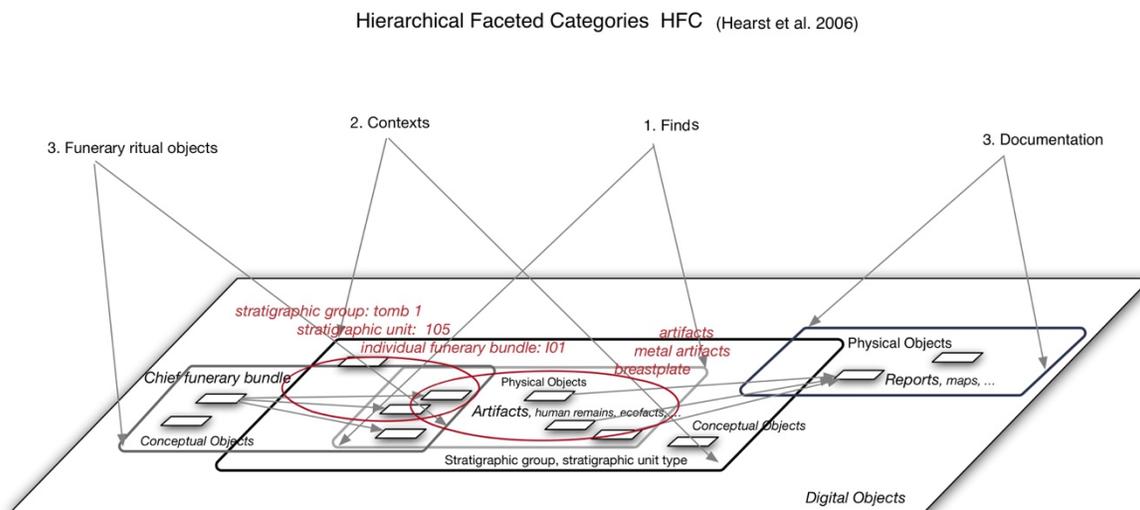


Fig. 2 : Jerarquías desarrolladas hasta el momento en el repositorio del PAEC

Primera etapa. Los objetos físicos recuperados en la excavación, “hallazgos”, se indexaron utilizando una taxonomía que incluía todas las categorías de objetos encontrados (en la Fig.2 “1. Finds”). Esta faceta se desarrolla en las subfacetas: artefactos, restos humanos, ecofactos y muestras. Dentro de éstas nuevas subfacetas como, por ejemplo, artefactos de metal, de piedra, de concha, ... ; sexo y edad de los restos humanos, ... Los detalles de esta taxonomía se pueden seguir en el propio repositorio (Mayo, J. & Mayo, C., 2017).

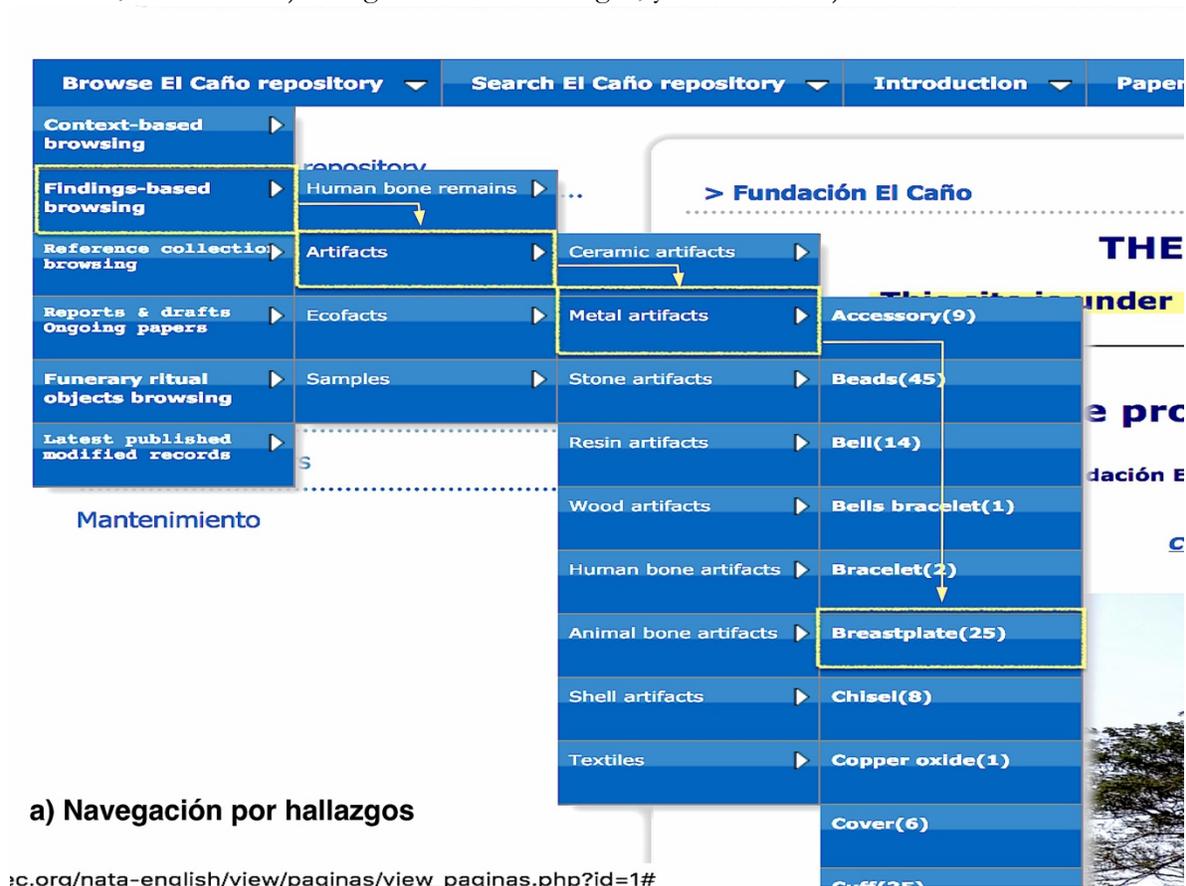
En esta etapa se definieron otros atributos básicos, tales como, descripción en texto libre del objeto, tipo de objeto, dimensiones del objeto físico, etc. Con estos atributos se completó la información sobre cualquier detalle de interés de los objetos de la excavación que no hubiese sido considerado previamente.

Segunda etapa. Se procedió a incorporar al modelo el análisis estratigráfico de las tumbas excavadas, identificando los objetos conceptuales necesarios para describir los contextos en que fueron recuperados los objetos físicos. Estos objetos conceptuales se organizan en una taxonomía de contextos (en la Fig.2 “2. Contexts”). Esta taxonomía nos permite navegar hasta los hallazgos registrados en la primera etapa, pero organizándolos y agrupándolos desde esta nueva faceta. Se utiliza la jerarquía de subfacetas: grupos estratigráficos, unidades estratigráficas y sub-unidades estratigráficas (enterramientos de individuos). Además, se utiliza un registro de tipos de grupos estratigráficos y otro de tipos de unidades estratigráficas. En esta segunda etapa se añaden también al modelo nuevos atributos con metadatos e información complementaria. También se refinan y amplían las jerarquías-taxonomías de hallazgos establecidas en la primera etapa. Por ejemplo, nuevos atributos describiendo los trabajos de conservación y una tipología más completa para la cerámica.

En la Fig. 2 se puede ver gráficamente que el conjunto de objetos digitales indexados desde “2.

Contexts” subsume al conjunto indexado por “1. *Finds*”, y esquemáticamente, como un mismo objeto, por ejemplo un pectoral de oro, se encuentra en nuestro espacio conceptual en la intersección de dos conjuntos, definidos, el primero siguiendo la taxonomía de hallazgos (un pectoral es un elemento de una clase de hallazgos) y el segundo siguiendo la taxonomía de contextos (un pectoral es parte del ajuar funerario de un individuo).

En la Fig. 3 se puede ver el mismo ejemplo, pero en este caso se navega utilizando el interfaz de ODA, en el caso 3a) navegando desde hallazgos, y en el caso 3b) desde contextos.



a) Navegación por hallazgos

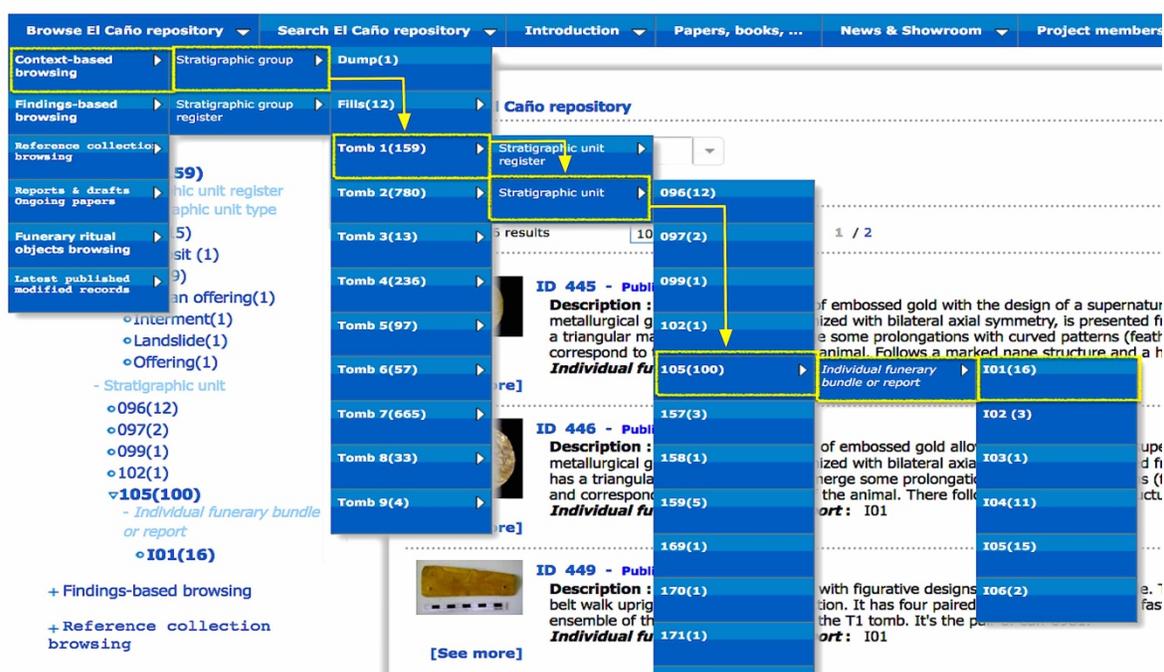
Fig. 3a: Navegación hasta un pectoral, a) desde hallazgos, como un elemento del conjunto de 25 pectorales registrados en el repositorio (hallazgos> artefactos> artefactos de metal> pectorales)

Los resultados de estas dos primeras etapas son accesibles públicamente en el repositorio de “El Caño” (Mayo, J. & Mayo, C., 2017). Accediendo a él se puede comprobar que el interfaz de ODA utiliza las dos facetas “hallazgos” y “contextos” como vértices de dos jerarquías de atributos que permiten navegar de forma sistemática y desde dos perspectivas distintas a todos los objetos arqueológicos recuperados en la excavación. Así como conocer la estructura de las tumbas excavadas.

Tercera etapa. Esta etapa está actualmente en fase de desarrollo y todavía no es pública. En ella además de extender las taxonomías ya iniciadas en etapas anteriores, por ejemplo, extender los artefactos de piedra con una tipología para líticos, se introducen dos nuevas taxonomías. La primera para organizar la documentación utilizada y generada por el proyecto y que es casi

independiente de las anteriores. La segunda para estudiar los objetos y su contexto desde la perspectiva de la serie de episodios que componen el ritual funerario (en la Fig.2 “3. Documentation” y “4. Funerary ritual objects”).

En la Fig. 2 también se puede seguir esquemáticamente como estas dos nuevas jerarquías introducen relaciones entre objetos basadas en documentos referencia. Por ejemplo, en la documentación tenemos un informe forense con los resultados de analizar los restos de un grupo de individuos. Este informe puede incluirse como documento-referencia en cada uno de los objetos que representan los restos de cada individuo del grupo. Otro ejemplo sería cuando un objeto del ritual funerario es un conjunto de objetos, por ejemplo, una ofrenda, y estos objetos ya han sido registrados individualmente. En este caso desde los documentos de la ofrenda se pueden incluir referencias a cada uno de los objetos que la componen.



b) Navegación por contextos

Fig. 3b: Navegación hasta un pectoral, b) desde contextos. El pectoral es un objeto del conjunto de 16 objetos del ajuar funerario del individuo I01 (contextos> grupo estratigráfico> tumba 1> unidad estratigráfica> 105>ajuar funerario individuo> I01).

Finalmente, y aunque todavía no se han comenzado a desarrollar, está previsto en etapas futuras, construir jerarquías-taxonomías que potencien el uso del repositorio en la difusión del conocimiento sobre las culturas indígenas, sensibilizando e implicando a la población local especialmente a las escuelas locales. Para conseguir esta implicación es importante potenciar el valor económico de la actividad turística en el entorno del Río Grande de Coclé. El parque arqueológico y su museo tendrían aquí un papel muy importante. Para que el repositorio apoyase esta labor, se ha pensado inicialmente en incluir en él dos nuevas jerarquías. Una primera que organice “recorridos educativos” del repositorio y muestre los trabajos escolares más representativos realizados por alumnos de las escuelas locales. La segunda jerarquía con recorridos tipo “museo virtual” que sirvan de apoyo al museo real del parque arqueológico “El Caño”.

4.- Control de acceso y perfiles de usuario

ODA permite establecer diversos niveles de control en el acceso a objetos, documentos y atributos. En el diseño UML de la Fig.1 puede distinguirse el registro de usuarios y sus derechos de acceso, por un lado, y por otro, la bitácora que registra las acciones realizadas por los usuarios con permiso para crear/modificar/borrar objetos (autoría). Así mismo se indican los tres niveles de acceso de los objetos: públicos, restringidos y confidenciales. Estos niveles de acceso dan lugar a cuatro perfiles de usuario: 1) Visitante o usuario-final, que puede acceder a los objetos públicos y dentro de éstos a sus documentos visibles. Bien porque han sido calificados como visibles, o bien, porque se ha facilitado una referencia de acceso al documento, es decir, puede haber documentos que solo son visibles para los visitantes que sigan un recorrido determinado. 2) Usuario-registrado. Los usuarios registrados en el sistema pueden acceder a un conjunto adicional de objetos que no es accesible a los visitantes. Dentro de los objetos pueden acceder sin limitaciones a todos sus documentos. 3) Administradores. Tienen los mismos privilegios de acceso que los usuarios registrados, pero además pueden crear nuevos objetos y editarlos. También pueden editar objetos que no han sido creados por ellos pero que hayan sido autorizados a hacerlo por el superadministrador. 4) Superadministradores. Pueden autorizar o dar de baja a cualquier tipo de usuario del sistema. Pueden crear y editar cualquier objeto del repositorio incluyendo un tipo de objeto privado que solo es accesible para los superadministradores. También pueden crear y acceder a los atributos de tipo confidencial que solo son accesibles a los superadministradores.

Por encima de estos cuatro niveles de usuario estaría el superusuario del servidor que tendría acceso directo, y por tanto todo el control, de la base de datos de ODA (ver Fig.6).

Resumiendo, ODA tiene los siguientes niveles de acceso a la información:

- Tres niveles de visibilidad para los objetos digitales: objetos visitables, objetos restringidos y objetos privados o confidenciales.
- Dos niveles de visibilidad para los atributos: atributos visibles y atributos confidenciales.
- Dos niveles de visibilidad para los documentos: documentos visibles y documentos restringidos (accesibles solo por referencia).
- Dos niveles para la autoría-edición de objetos: creación-edición limitada y sin restricciones.

3.- Arquitectura e implementación

La arquitectura de ODA tiene como núcleo una base de datos relacional estándar con la que materializamos casi todo el diseño UML visto en el apartado anterior. Básicamente, cada elemento del diseño UML se corresponde con una tabla en el modelo relacional (Fig 4). Para obtener el modelo relacional, se simplifican algunos elementos del diseño UML, como los tipos de documentos, que se incluyen en la tabla “documents” y “data” cuya información se incluye en el resto de las tablas de datos.

En la Fig.5 se detalla el diseño E-R y las restricciones que las claves externas imponen sobre las tablas. En la implementación del modelo, no se impone la integridad referencial en las

relaciones para facilitar el mantenimiento de las tablas. La programación del interfaz verifica las simplificaciones introducidas. Este núcleo está dentro de una aplicación web que sirve de interfaz con la base de datos y que no se detalla en este documento.

Es interesante recordar que el modelo entidad-relación utilizado en la arquitectura de ODA no es un modelo del dominio en el que vamos a utilizar ODA. En el caso concreto del PAEC, no es un modelo de la investigación llevada a cabo en El Caño. El modelo entidad-relación que estamos describiendo aquí (tablas, campos, tipos y relaciones entre tablas) es el modelo de un repositorio de datos genérico que incluye, como un dato más, los datos sobre la organización de sus propios datos. Lo que es lo mismo que decir que es un metamodelo de una base de datos jerárquica que mediante una instancia concreta, permite modelar un dominio concreto.

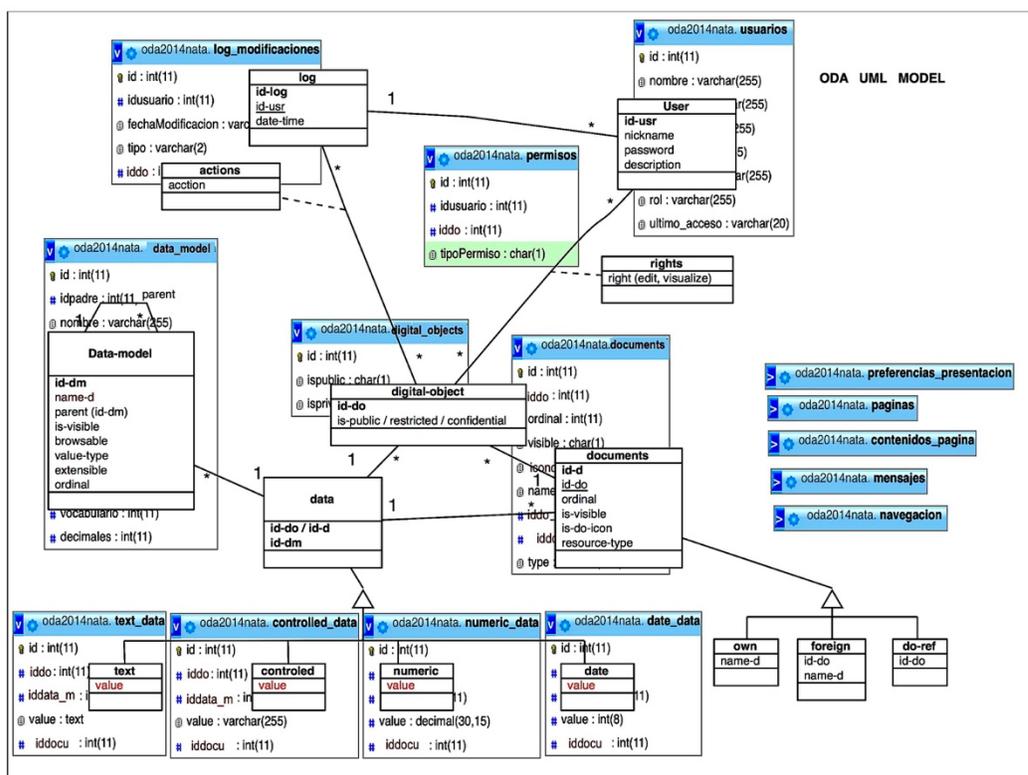


Fig. 4. Del modelo UML al modelo Entidad-Relación (E-R)

Implementación

La implementación de ODA tiene tres componentes principales (Fig.6)

- Una base de datos relacional con los datos del dominio y los datos del modelo de datos (Fig.5). Esta base de datos se gestiona utilizando el interfaz ODA, pero es importante destacar que, como se señala en la figura 6, el super administrador del servidor puede gestionarla directamente y sin restricciones utilizando el servidor de base de datos (en nuestro caso MySQL).
- Una carpeta del sistema de archivos (FS) del sistema operativo (OS) con toda la base de datos documental, es decir, todos los documentos digitales con información relevante para el proyecto. Agrupándolos en carpetas cuyo identificador coincide con el identificador del objeto digital que representan. Como en el caso anterior, los objetos

se gestionan utilizando ODA, pero el super administrador del servidor puede acceder a ellos directamente a través del FS del OS del servidor.

- Una carpeta del OS con todos los documentos digitales de la aplicación web que sirve como interfaz del sistema y que estrictamente hablando no forma parte de ODA.

En nuestra implementación utilizamos un servidor de base de datos relacional (MySQL) para crear la base de datos, un servidor web (Apache) con extensión PHP para crear el interfaz ODA con la base de datos relacional y con el sistema de archivos del OS (UNIX).

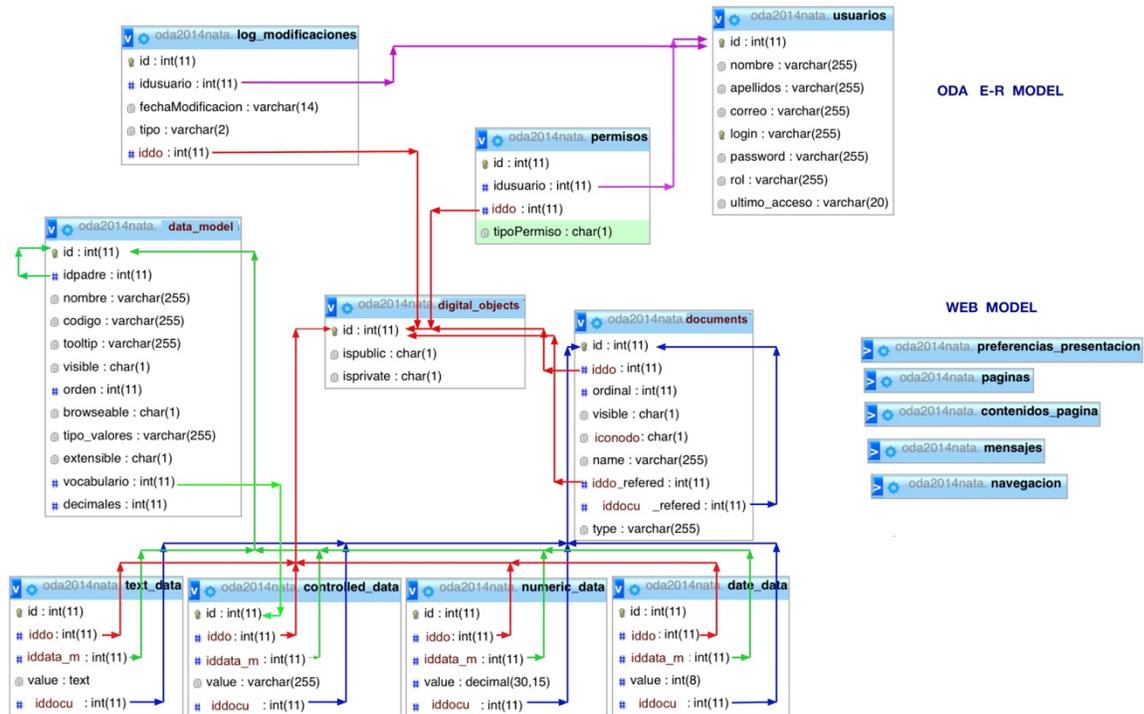


Fig. 5 El modelo E-R de ODA

Las funciones principales de este interfaz son:

- Visualizar los objetos digitales. Presentar al usuario los objetos individuales reconstruyéndolos a partir de la información en las tablas y los documentos almacenados en el OS. Debe comprobar si el usuario tiene permiso para acceder al objeto y a la información que contiene.
- Visualizar el modelo de datos. El modelo de datos del dominio es una variable más del sistema y para conocer su estado en un instante dado el interfaz debe reconstruirlo a partir del contenido de la base de datos, fundamentalmente de la tabla "data_model".
- Navegar. Permitir que el usuario pueda navegar por las estructuras del modelo de datos marcadas como navegables por los investigadores y como accesibles de acuerdo con los permisos de usuario.

d) Buscar información. Visualizar los conjuntos de objetos que satisfacen las condiciones impuestas a las estructuras del modelo de datos en cada punto del espacio de conocimiento.

e) También se programan en PHP resto de las funciones de gestión y mantenimiento del repositorio, de los objetos digitales y del modelo de datos.

f) Otros servicios (en nivel de pruebas mediante java plug-in) para exportar/importar la información en otros formatos (Gayoso-Cabada et al., 2016).

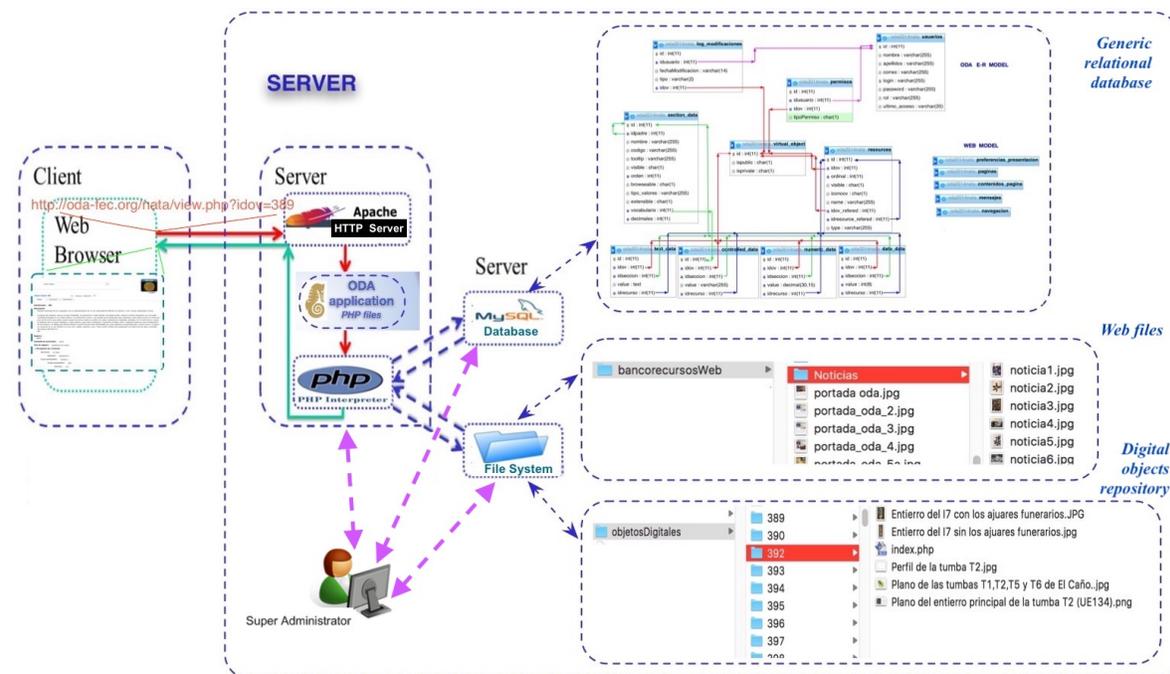


Fig. 6 Esquema de la arquitectura e implementación de ODA

El estado actual del ODA (nov. 2017) del Proyecto Arqueológico El Caño:

- 2.246 N° total de Objetos Digitales (ODs)
 - 1.819 N° de ODs públicos
 - 19 N° de ODs solo usuarios registrados
 - 408 N° de ODs confidenciales
- 3.722 N° total de documentos digitales multimedia
- 4,2 MB espacio consumido por la base de datos (sentencias sql)
- 4.310 MB espacio consumido por los documentos multimedia del repositorio
- 370,9 MB espacio consumido por los documentos multimedia de la web

4.- Trabajos relacionados. ODA en el PAEC. Trabajo futuro.

4.1 Trabajos relacionados

Para valorar a ODA en el contexto de los sistemas de gestión de la información arqueológica similares a él, empezamos explorando dos tipos de sitios web:

- a) museos e instituciones con colecciones arqueológicas importantes,
- b) proyectos arqueológicos con excavaciones activas.

Con respecto a los primeros, hay que tener en cuenta que el legado arqueológico depositado en museos e instituciones no solo está compuesto por colecciones de objetos, si no también, por material documental heterogéneo: informes de campo, de laboratorio y otros trabajos post-procesuales. La tarea de digitalizar todo ese material es costosa en tiempo y recursos, por lo que no es de extrañar que en instituciones con grandes fondos parte de este material esté todavía en proceso de digitalización, y sus sitios web en proceso de actualización. Dos ejemplos significativos de esta situación son: el *Penn Museum, University of Pennsylvania Museum of Archaeology and Anthropology* (accedido en mayo 2018), donde, de los 4.852 objetos arqueológicos de la colección de Sitio Conte, solo 671 tienen imagen digital en la web. El segundo ejemplo son las colecciones de la *Smithsonian Institution del National Museum of the American Indian* (accedido en mayo 2018), aunque en este caso no se dan cifras, la situación parece ser similar.

Aun así, se puede comprobar que los sistemas de gestión de la información de algunos de los sitios web analizados utilizan estrategias de organización de la información similares a las utilizadas en ODA. Mas concretamente, los sitios que tienen aplicaciones web con interfaces mas elaboradas utilizan, como ODA, jerarquías de facetas para hacer mas intuitiva la búsqueda de información y la navegación. Facilitando así, que el visitante pueda explorar la colección siguiendo la línea conceptual marcada por las jerarquías de facetas (Hearst, 2006a). Por ejemplo, esta estrategia se utiliza en las colecciones *online* del *Penn Museum*, del *American Museum of Natural History*, o del *Metropolitan Museum of Art* (accedidos en mayo 2018).

Con respecto a los sitios web de los proyectos con excavaciones activas (en algunos casos asociados a la web del museo del sitio), para poder comparar, necesitamos que el visitante pueda acceder a la base de datos donde están registrados los objetos recuperados y en su caso a la documentación del proyecto. El caso mas interesante de cuantos hemos analizado ha sido el de las excavaciones organizadas por la empresa *DigVentures* del Reino Unido (muchas financiadas mediante *crowdfundig*). En estas excavaciones se utiliza el sistema *Digital Dig Team* que permite poner *online* de forma inmediata los datos de una excavación en curso. Como resultado el visitante puede navegar por los registros de los hallazgos y otros elementos de la excavación y comprobar que la información en ellos se adapta al estado de la excavación, con una flexibilidad comparable a la que tenemos en ODA. El motor responsable de esta funcionalidad es el sistema ARK, que luego analizaremos con mas detalle, y que actúa como *back-end* del sistema *Digital Dig Team*. Un ejemplo es el proyecto *Borrowed time* (accedido en mayo 2018) en el que se está excavando un cementerio de la Edad del Bronce en el Noroeste de Inglaterra.

Sin embargo, esto no parece ser lo habitual en los sitios web de proyectos arqueológicos mas importantes, en los que, o no hay acceso a la base de datos, o solo lo hay para investigadores registrados. Por ejemplo, no hemos encontrado acceso a la base de datos en la web del museo del sitio de la Dama de Cao en el Complejo Arqueológico El Brujo (accedido en abril 2018) ni en la web del Museo Tumbas Reales del Señor de Sipan (accedido en abril 2018). La misma

situación podemos observar en España en los yacimientos de Atapuerca .ATENCIÓN Revisar la situación con Google online (accedido en mayo 2018).

Desde el punto de vista teórico, la información más completa sobre sistemas de gestión comparables a ODA la encontramos en las publicaciones que describen proyectos informáticos cuyo objetivo es desarrollar sistemas de gestión arqueológicos. Si limitamos los resultados de nuestra búsqueda al periodo más reciente, cuando las tecnologías web aplicadas a las humanidades han alcanzado su madurez, encontramos dos proyectos que son una referencia importante para valorar las posibilidades de las tecnologías web aplicadas a los sistemas de gestión arqueológicos. Estos dos proyectos son:

- *Archaeological Recording Kit* (ARK, accedido en abril 2018), desarrollado por L-P: Archaeology, una empresa de iniciativa privada del Reino Unido pero que colabora con universidades e instituciones públicas (Eve & Hunt, 2008).
- *Open Information System for Research in Archeology* (OpenInfRA, accedido en abril 2018) desarrollado por un consorcio dirigido por el *Deutsches Archäologisches Institut* (DAI) y en el que participan las universidades: Brandenburg University of Technology Cottbus y Dresden University of Applied Sciences (Schulze et al., 2012; Henze et al., 2013; Schwarzbach et al., 2014)

Los sistemas de gestión de la información desarrollados en estos dos proyectos, aunque son proyectos de mayor dimensión que ODA, nos permiten situar la aportación de ODA en el contexto de los sistemas de gestión aplicados a la arqueología. En efecto, aunque cada proyecto aborda la construcción de su sistema bajo un enfoque y con unos objetivos específicos distintos, los tres sistemas optan por desarrollar la aplicación en torno a un núcleo similar: una base de datos relacional con un modelo, o esquema de datos, genérico.

Para los autores de ARK la motivación para construir su sistema es casi ideológica: conseguir poner la informática al servicio del arqueólogo de campo y no al revés. Como consecuencia, su primer objetivo es diseñar un sistema de registro y recogida de datos arqueológicos, que el arqueólogo de campo pueda configurar de acuerdo con las necesidades específicas de su proyecto, y ello sin tener que recurrir al experto en informática (otros objetivos, como exportación e importación de datos en formatos estándar, e interoperabilidad con otras aplicaciones, quedan en segundo lugar o para futuras versiones). Para satisfacer este primer objetivo, los creadores de ARK recurren al concepto de hipertexto, y diseñan una base de datos con un modelo hipertextual genérico: nodos y enlaces. En ARK los nodos pueden ser de dos tipos: “ítems” y “fragmentos”. Los “fragmentos” “enlazan” con un “ítem” para completar la información en el mismo (podemos pensar en los “fragmentos” como algún tipo de atributo del “ítem”). Los “ítems” se “enlazan” entre ellos formando una red no jerárquica, la red hipertextual, con la información arqueológica. Los “ítems” se pueden agrupar para formar nuevos “ítems”. Por ejemplo, como consecuencia de una búsqueda obtenemos un grupo de “ítems”, este grupo puede a su vez formar un nuevo “ítem”. Los elementos básicos del hipertexto pueden ser de diferentes tipos, por ejemplo, los “ítems” pueden ser de tipo: “contexto”, “foto”, “muestra”, etc. A su vez los “fragmentos” pueden ser de tipo: “fecha”, “acción”, “texto-descripción”, “texto-interpretación”, etc.

La consideración final, que hacen los autores, es que ARK debe verse no tanto como un repositorio de datos, si no como un medio para la creación textos arqueológicos (los informes de la excavación).

Como en ODA, el núcleo de ARK es una base de datos relacional (en ambos casos en un servidor MySQL), y lo que es más interesante, con esquemas genéricos entre los que es fácil encontrar equivalencias. En cierto sentido, un “ítem” de ARK es como un “objeto digital” de ODA y el conjunto de “fragmentos” enlazados con un “ítem” es como el conjunto de pares “atributo: valor” que caracterizan un “objeto digital”. Hay más características de ARK con una equivalencia clara en ODA, aunque otras no tanto, por ejemplo, los documentos multimedia que en ODA quedan perfectamente empaquetados como parte constituyente de los objetos digitales, en ARK quedan un tanto dispersos en la red hipertextual. Esto tiene su lógica ya que ambos sistemas tienen un enfoque distinto. Para ARK la información que se obtiene como resultado de una excavación tiene la forma no jerarquizada de un hipertexto (aspectos concretos de la excavación serán fragmentos del hipertexto). Mientras que, para ODA, la información sobre la excavación está organizada jerárquicamente según el modelo del dominio construido por los investigadores. En la actualidad L-P: Archaeology está trabajando en la versión 2 de ARK. En ella se está refactorizando la base de datos para mejorar su rendimiento y facilitar su mantenimiento. También se trabaja en el desarrollo de una suite de herramientas externas a ARK que provean diferentes servicios. Por ejemplo, integrar aplicaciones geográficas y exportar contenidos en formatos estándar como CIDOC-CRM.

El segundo proyecto de referencia es OpenInfRA. En este caso, se trata de un proyecto de mayores dimensiones aun que ARK, basado en dos sistemas anteriores, iDAI.field y CISAR, también desarrollados por el DAI. El primero centrado en la arqueología de campo y el segundo en el estudio y restauración de diferentes tipos de construcciones. El foco de OpenInfRA, según sus autores, es la investigación arqueológica y con el nuevo proyecto se trata de superar ciertos aspectos de los dos sistemas anteriores que se habían quedado obsoletos (por ejemplo, iDAI.field utilizaba una base de datos propietaria, FileMaker12, que daba algunos problemas). El objetivo global del nuevo sistema del DAI es ayudar a elaborar y documentar una investigación arqueológica, a publicarla y a difundirla de forma efectiva. En sus objetivos incluye no solo las excavaciones y sus hallazgos, si no también, trabajos de arquitectura y restauración, sondeos arqueológicos, estudios geológicos, etc. La consecuencia de este planteamiento es que OpenInfRA no solo está interesado en integrar los datos de una excavación actual utilizando el modelo de datos que los arqueólogos de campo consideren adecuado y que puedan configurar fácilmente, si no también (y esto es muy importante en arqueología clásica y en una institución con los fondos del DAI), en recuperar los datos y la documentación de todas las excavaciones anteriores que puedan aportar algo a la investigación en curso. En términos técnicos el foco está en las interfaces abiertas, para publicar y compartir su contenido en un escenario de reusabilidad e interoperabilidad.

Para integrar toda esta información, necesariamente heterogénea, y para difundirla después en el formato solicitado, normalmente un estándar, OpenInfRA también utiliza una base de datos con un modelo genérico. Lo que se pretende modelar desde un punto de vista genérico es la información. Por tanto, se trata de identificar los elementos atómicos necesarios y suficientes para componer, o descomponer, cualquier estructura informativa. Para OpenInfRA estos elementos son, “tópico” y “característica del tópico”, por un lado, y “tipo de grupo de atributos” y “tipo de atributo”, por otro. La relación entre “característica del tópico” y “tipo de atributo” tiene información adicional sobre la multiplicidad de la relación y un valor por defecto. Cada proyecto individual almacenado en el sistema instancia los tópicos y los valores de los atributos que le sean necesarios. Por lo que el modelo se extiende con la clase “instancia

de t3pico” y la relaci3n entre “instancia de t3pico” y “tipo de atributo” se etiqueta con el “valor” del atributo (Henze et al., 2013). En nuestra opini3n, el modelo de esta base de datos es m3s abstracto y m3s dif3cil de manejar que el de ARK y el de ODA, aunque, en cierto modo podemos asimilar los pares “atributo: valor”, y los “objetos f3sicos o conceptuales” de ODA a los “atributos” y “t3picos” de OpenInfRA. La heterogeneidad de la informaci3n que se maneja en este sistema y la complejidad de su base de datos hace necesario ampliar el lenguaje de consulta m3s all3 del SQL. Este nuevo lenguaje de consulta es tambi3n una parte importante de los objetivos del proyecto.

En resumen, despu3s de este breve estudio de los sistemas de gesti3n de la informaci3n en arqueolog3a, se puede decir que ODA est3 en l3nea con los proyectos de gesti3n m3s avanzados y que ofrece ventajas derivadas de la sencillez del modelo de datos gen3rico que utiliza y de la capacidad expresiva de las jerarqu3as de facetas que pueden construirse con este modelo. Destacamos los siguientes puntos de este estudio:

- En ODA, la estrategia de organizar la navegaci3n mediante jerarqu3as de facetas es parte del modelo del dominio. Algunos de los museos m3s importantes utilizan tambi3n esta estrategia para exponer sus colecciones de objetos arqueol3gicos, aunque en general, al no tener un modelo del dominio completo, se aplica de forma m3s restrictiva.
- ARK y OpenInfRA son proyectos m3s ambiciosos que ODA, sin embargo, los tres proyectos comparten aspectos fundamentales de dise1o. El dise1o de ODA, utilizando una base de datos gen3rica, est3 en l3nea con los proyectos estudiados, aunque en nuestra opini3n es un modelo m3s sencillo e intuitivo lo que facilita satisfacer de forma m3s inmediata los requisitos de los sistemas de gesti3n arqueol3gicos.
- El dise1o de ODA permite hacer expl3cito el modelo del dominio y utilizarlo para explorar y analizar el contenido del repositorio. Esta aproximaci3n, tambi3n ser3a posible con los otros dos sistemas, aunque resultaria mas complicada dado que los modelos de bases de datos que utilizan son m3s complejos y est3n dise1ados para satisfacer otros objetivos.
- El esquema de base de datos gen3rico de ODA utilizando el concepto de “objeto digital” facilita la gesti3n de documentos y los procesos de intercambio de datos con otras aplicaciones, que en el caso de ODA se pueden hacer directamente utilizando las herramientas est3ndar de una base de datos relacional. No obstante, este ser3a uno de los puntos a tener en cuenta en una nueva versi3n de ODA.

4.2 ODA en el PAEC.

En los siguientes apartados enumeramos tanto de lo que ha supuesto la utilizaci3n de ODA, como de lo que esperamos poder hacer en proyectos inmediatos:

1. Un sistema para la difusi3n del conocimiento. ODA ha sido dise1ada para ser una aplicaci3n abierta al p3blico, de hecho, gran parte del contenido del repositorio del Proyecto Arqueol3gico El Ca1o est3 accesible en abierto desde el mes de enero del a1o 2016. Los permisos de acceso a los objetos digitales, datos y documentos de este repositorio, son diferentes seg3n el perfil de cada usuario.
2. Una herramienta flexible dado que su “modelo de datos” lo es. Su modelo puede adecuarse, mediante la creaci3n de nuevos pares atributo-valor y nuevas relaciones, a los intereses de los diferentes grupos de actores involucrados en el desarrollo del PAEC: investigadores, gestores del patrimonio arqueol3gico, miembros de las instituciones comunitarias locales, docentes, estudiantes, responsables de museos, o simplemente, personas

interesadas en la cultura de sus antepasados y/o visitantes casuales.

3. Un sistema abierto y democrático a través del cual los investigadores comparten sus datos. Para el PAEC compartir y difundir la información es muy importante y ODA permite compartir la información entre los miembros del equipo de investigación, con investigadores externos y con los museos e instituciones docentes.

4. Compartir información entre los diferentes miembros del equipo de investigación es fundamental para tener una visión coherente y completa de los resultados del proyecto y poder tomar decisiones a medida que éste progresa. Compartir información con otros grupos de investigadores (investigadores externos) es importante porque de esta forma éstos pueden seguir de cerca, y hasta donde los investigadores del proyecto lo consideren oportuno, el desarrollo de las investigaciones. Esto mejora la eficiencia de la difusión de resultados entre la comunidad científica, fomenta la discusión y abre nuevas oportunidades de colaboración entre equipos de investigadores y proyectos.

5. La colaboración con los museos es uno de nuestros proyectos inmediatos. Para los museos, poder interactuar a través de ODA facilita la colaboración con los investigadores y la presentación de las piezas en su contexto. El modelo de conocimiento construido facilita una presentación acorde con los resultados de la investigación y con información más detallada. Está previsto que ODA sea también la base para construir visitas virtuales al museo de sitio de El Caño y a otros museos con exhibiciones sobre El Caño.

6. La colaboración con docentes es otro proyecto inmediato. Desde el punto de vista docente, utilizando ODA los investigadores pueden acercar a los estudiantes la metodología científica, dado que, al acceder a la información a través del modelo de conocimiento, acceden a las estrategias utilizadas por los arqueólogos para clasificar e interpretar sus materiales.

7. Economizar esfuerzos y maximizar rendimiento. Finalmente se puede destacar la economía de esfuerzo que supone disponer de un entorno de trabajo en el que, una vez que se elabora una información, ésta pueda hacerse disponible para otros colectivos de usuarios sin otro esfuerzo más que el que supone activar los permisos correspondientes.

4.3 Trabajo futuro

Consideramos dos tipos de trabajo futuro:

- Trabajo futuro que afecta a la información que contiene el sistema y a como organizarla. Nos proponemos alcanzar al menos estos cuatro objetivos.

1. Apoyo al museo del sitio. Apoyo a exhibiciones.
2. Crear recorridos docentes.
3. Versión en inglés de ODA. Protocolo de sincronización.
4. Integrar en ODA información de otros yacimientos de la cuenca de Río Grande de Coclé. Especialmente de Sitio Conte y de las colecciones que están en otros museos, tanto en Panamá como fuera de Panamá. Llegar a un acuerdo de colaboración con los museos que tienen esas colecciones.

- Trabajo futuro que afecta a la infraestructura informática del sistema. Nos proponemos también al menos estos cuatro objetivos.

1. Mejorar y simplificar la gestión directa de la BD utilizando herramientas sql estándar.

2. Mejoras en la seguridad, usabilidad y mantenimiento.
3. Nuevas funciones de exportación e importación de datos
4. Integrar servicios de aplicaciones externas

APÉNDICES

1.- Ejemplos de modelado mediante facetas del trabajo de investigación y difusión.

a) Navegación por publicaciones e informes

Informes internos
 Materiales / Tema publicación
 Tipo publicación
 Estatus
 Fecha/estatus

b) Navegación por el ritual funerario

Episodio

[E1-ajuar jefe, E5-ajuar élite, E6-E8-ofrenda colect, E5-ofrenda personal, E4.sacrificio h.]
 Tiempo (ref. entierro principal) [E1-E2 anterior, E3-E5 simultáneo, E6-E8 posterior]
 Lugar [Poblado, centro ceremonial, fosa ...]

Tipo/función

del objeto [insignia clan, símbolo religioso, efigie guardián, utensilio, arma, ...]
 Disposición [relativo al cuerpo/entierro: cabeza, torso, ...]

Iconografía [animal mítico, escena ritual/religiosa,...]

Interpretación [viaje inframundo, símbolo de poder sobre agua/aire/tierra, ...

Soporte físico [placa oro, brazalete oro, resina+oro, arcilla, fundido 3D, ...]

Técnica [martillado en frío, modelado, cera perdida...]

c) Participación escolar

- propuestas de trabajos, temas {}
- visitas virtuales []
- - trabajos de las escuelas, resultados {}

d) Otros ejemplos en <http://oda-fec.org/nata>

2.- Instalación y mantenimiento de ODA

1. Instalar servidores: Apache-PHP web, MySQL base de datos (XAMP).

2. Instalar y mantener en el servidor:

- La base de datos relacional (SQL),
- El banco de recursos de la web (OS anfitrión: macOS),
- El banco de objetos digitales y sus documentos propios (OS anfitrión: macOS).

3. Crear y mantener el sitio web: páginas y navegación
4. Crear y mantener del modelo de datos
5. Gestionar los usuarios y el acceso a la información
6. Controlar acceso a los objetos, a sus atributos, a sus documentos.

Bibliografía

- American Museum of Natural History (2018): Collections. En <https://www.amnh.org/our-research/anthropology/collections>. Último acceso mayo 2018
- ARIADNE (2018): Infrastructure. En <http://www.ariadne-infrastructure.eu> .Último acceso abril 2018
- ARK (2018): An open source solution to project recording. En <https://ark.lparcology.com>. Último acceso abril 2018
- Atapuerca (2018): Exposición online. Google Cultural. En <http://www.museoevolucionhumana.com/es/exposicion-online-google-cultural>. Último acceso abril 2018
- DigVentures (2018): Finds at Borrowed Time. En <https://digventures.com/borrowed-time/ddt/browser.php> . Último acceso mayo 2018
- Broughton, Vanda (2013): Faceted Classification as a General Theory for Knowledge Organization. En: *SRELS Journal of Information Management*, Volume 50, Issue 6, December 2013
- CIDOC- CRM (2018): Conceptual Reference Model. En <http://www.cidoc-crm.org>. Último acceso abril 2018
- El Brujo (2018): Complejo Arqueológico. En <https://www.elbrujo.pe> Último acceso abril 2018
- Denton, W. (2003): *How to make a faceted classification and put it on the web*. En <http://www.miskatonic.org/library/facet-web-howto.html>. Último acceso enero 2018
- Eve, S. and Hunt G. (2008): *ARK: A Development Framework for Archaeological Recording*. En: Posluschny, A., K. Lambers and I. Herzog (eds.), *Layers of Perception*. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, Germany, April 2–6, 2007
- Fernández-Valmayor, Alfredo; Fernández-Pampillón, Ana María; Fernández-Chamizo, Carmen; Navarro, Antonio; Cristóbal, Jesús. (2011): Implantación de un Campus Virtual de Grandes Dimensiones: el Campus Virtual de la UCM. En: *IEEE RITA*, 6(4), pp. 167-174.
- Gayoso-Cabada, Joaquín, Daniel Rodríguez-Cerezo and José Luis Sierra. (2016): *Browsing Digital Collections with Reconfigurable Faceted Thesauri*. En: 25th. International Conference on

Information Systems Development (ISD2016 Poland)

Geser, G., Gonzalez-Perez, C. (2017): *Report on E-Archaeology Frameworks and Experiments*. En <http://ariadne-infrastructure.eu/ger/Media/Files/D17.1-Report-on-e-Archaeology-Framework-and-Experiments>. Último acceso marzo 2018

Giess, M.D., Wild, P.J. and McMahon C.A. (2007): *The Use of Faceted Classification in the Organisation of Engineering Design Documents*. En International Conference on Engineering Design, ICED'07. 28 - 31 August 2007, Cite des Sciences et de L'industrie, Paris, France

Giunchiglia, F., Dutta, B., Maltese, V. (2009): *Faceted lightweight ontologies*. En: Borgida, A., Chaudhri, V.K., Giorgini, P., Yu, E.S.K., (eds.): *Conceptual Modeling: Foundations and Applications*. Volume 5600 of Lecture Notes in Computer Science., Springer (2009) 36–51

Guinea, M. (2004): El proyecto Chasqui. En: Alfredo Fernández-Valmayor Crespo, Ana Fernández-Pampillón Cesteros, Jorge Merino Granizo (eds.). *En apoyo del aprendizaje en la universidad: hacia el espacio europeo de educación superior*. Madrid. Editorial Complutense. ISBN 84-7491-774-3, págs. 228-233

Guinea, M., Avila, D., García, N. (2009): Construcción de objetos de aprendizaje y desarrollo de experiencias docentes. En: Alfredo Fernández-Valmayor Crespo, Amelia Sanz Cabrerizo, Jorge Merino Granizo (eds.). *Buenas prácticas e indicios de calidad*. Madrid. Editorial Complutense. ISBN 978-84-7491-968-4, págs. 8-13

Hearst, Smalley, & Chandler (2006): *Hierarchical Faceted Categories*. En: CHI 2006 85

Hearst, M. A. (2006a): *Design recommendations for hierarchical faceted search interfaces*. En: Proc. SIGIR 2006, Workshop on Faceted Search. pp. 26--30.

Hearst, Marti A. (2006b): Clustering versus Faceted Categories for Information Exploration. En: *Communications of the ACM*, April 2006

Henze, F., Magdalinski N., Schwarzbach, F., Schulze, A., Gerth, Ph., Schäfer, F. (2013): *Concepts and Technologies for a Comprehensive Information System for Historical Research and Heritage Documentation*. En: XXIV International CIPA Symposium, 2 – 6 September 2013, Strasbourg, France.

Hochstetter, F., Haoa, S., Lipo, C., & Hunt, T. (2011): A Public Database of Archaeological Resources on Easter Island (Rapa Nui) Using Google Earth. En: *Latin American Antiquity*, 22(3), 385-397. Society for American Archaeology.

Lawson, A., Eklund, P. W., Goodall, P., Wray, T., Daniel, V. & Van Olffen, M. (2010): *Designing a digital ecosystem for the new museum environment: the Virtual Museum of the Pacific*. En: H. Yeatman (Eds.). *The SInet 2010 eBook* (pp. 227-239). Wollongong: SInet UOW.

Martini, R.G., Araujo, C., Libelotto, G.R., Henriques, P.R. (2016): A Reduced CRM-Compatible Form Ontology for the Virtual Emigration Museum. En: Rocha Á., Correia A., Adeli H., Reis L., Mendonça Teixeira M. (eds.) *New Advances in Information Systems and*

Technologies. Advances in Intelligent Systems and Computing. Vol. 444. Springer, Cham

Mayo, J. & Mayo, C. (2017): Repositorio de datos del Proyecto Arqueológico El Caño. Centro de Investigaciones Arqueológicas del Istmo. En <http://oda-fec.org/nata>. Último acceso noviembre 2017

Metropolitan Museum of Art (2018): En <https://www.metmuseum.org/art/collection/search> . Último acceso abril 2018

Museo Tumbas Reales del Señor de Sipan (2018): En <http://www.museotumbasreales.com> . Último acceso abril 2018

National Museum of the American Indian, Smithsonian Institution (2018): En <http://nmai.si.edu> . Último acceso abril 2018

NISO Z39.19-2005 (R2010) (2010): Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. *National Information Standards Organization NISO (U.S.)* 172nd ed., (Bethesda, Md): NISO Press.

OpenInfRA (2018): En <https://www.b-tu.de/openinfra/> . Último acceso abril 2018
En <https://fallback.dainst.org/projekt/-/project-display/113484> . Último acceso abril 2018

Penn Museum, University of Pennsylvania Museum of Archaeology and Anthropology (2018): <https://www.penn.museum> . Último acceso abril 2018

Rodríguez-Castro, B., Glaser, H. and Carr, L. (2010): *How to reuse a faceted classification and put it on the semantic web*. En: Proceedings of the 9th international semantic web conference on The semantic web. Volume Part I (ISWC'10), Peter F. Patel-Schneider, Yue Pan, Pascal Hitzler, Peter Mika, and Lei Zhang (Eds.), Vol. Part I. Springer-Verlag, Berlin, Heidelberg, 663-678.

Schulze, A., Schäfer, F.F., Gehrt, P., Henze, F. and Schwarzbach, F.: (2012): *OpenInfRA – Storing and Retrieving Information in a Heterogeneous Documentation System*. En: Proceedings of the CAA conference 2012, University of Southampton.

Schwarzbach, F., Schäfer, F., Schulze, A. (2014): *Find the Balance - Modelling Aspects in Archaeological Information Systems*. En: Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology CAA 2014 - 21st Century Archaeology/F. Giligny, F. Djindjian, L. Costa, P. Moscati, S. Robert (eds.)

Sierra, J., Fernández-Valmayor, A., Guinea, M., Hernanz, H. (2006): From Research Resources to Learning Objects: Process Model and Virtualization Experiences. *Journal of Educational Technology & Society*, Vol. 9, Pg. 56-68

Vickery, B. (2008): Faceted classification for the web. *Axiomathes* 18(2) (June 2008) 145–160

Yee, K-P., Swearingen, K., Li, K., and Hearst, M. (2003): *Faceted metadata for image search and browsing*. En: Proceedings of the CHI 2003, Fort Lauderdale, FL, April 2003.

Zhuge, H., Xing, Y., Shi, P. (2008): Resource space model, owl and database: Mapping and integration. *ACM Trans. Internet Technol.* 8(4) (2008) 1–31