

El proceso de desarrollo de una biblioteca digital y su apertura e integración en el recolector *Hispana*: el caso de la *Biblioteca Digital Ovidiana Abierta*¹

PALOMA CENTENERA CENTENERA
Universidad Pontificia de Salamanca en Madrid
paloma.centenera@upsam.es

MARÍA LUISA DÍEZ PLATAS
Universidad Pontificia de Salamanca en Madrid
mluisa.diez@upsam.es

LA BIBLIOTECA DIGITAL OVIDIANA

Unos requisitos basados en la calidad

Desde un punto de vista técnico la biblioteca ha sido diseñada para ser una herramienta eficaz para usuarios e investigadores. Para ello, además de proporcionar un marco estético que acompañe la belleza de las ilustraciones de Ovidio que en ella se muestran, debe permitir a los usuarios encontrar fácilmente la información que necesitan.

Una biblioteca digital, como cualquier aplicación web, tiene unas características propias que diferencian su desarrollo del de otros sistemas informáticos. Estas características diferenciadas hacen que los atributos con los que se mide su calidad tengan también singularidades. Olsina *et al.* (Olsina, 1999) identifican los siguientes requerimientos de la calidad en aplicaciones web:

- Facilidad de uso. Coloquialmente este término hace referencia a la facilidad para utilizar una herramienta. La Oficina Internacional para la estandarización (ISO), en una referencia muy adecuada para nuestro fin, define este atributo como “la efectividad, eficiencia y satisfacción con la que un producto permite alcanzar objetivos específicos a usuarios específicos en un contexto de uso específico” (ISO/IEC 9241-11).
- Funcionalidad. Esta característica está relacionada con el valor añadido de los servicios proporcionados por el sitio web. Estaríamos hablando básicamente de capacidad de búsqueda y recuperación de la información.

¹ Este artículo forma parte de los resultados del proyecto de investigación “BIBLIOTECA DIGITAL OVIDIANA: Las ediciones ilustradas de Ovidio. Siglos XV al XIX (III): Las bibliotecas de Castilla y León” (HAR2011-25853), financiado por el ministerio de Economía y Competitividad.

- **Confiabilidad.** Podríamos definir este término como ausencia de errores. La confiabilidad de un sitio web está relacionada tanto con la correcta ejecución de la funcionalidad proporcionada por el sitio como con el correcto procesamiento de los vínculos.
- **Eficiencia.** Hablamos del uso adecuado de los recursos. Puede verse en el tiempo de respuesta y la velocidad de generación de las páginas y gráficos así como en el buen uso del espacio de almacenamiento en el servidor.
- **Facilidad para recibir mantenimiento.** Está asociada con la fácil corrección, adaptación y extensibilidad del sitio web.

Todas estas características, junto con otras mencionadas por Offutt (Offutt, 2006) como es la escalabilidad, han sido tenidas en cuenta a la hora de enunciar los requisitos a satisfacer por la biblioteca digital ovidiana.

La facilidad para el uso está asociada con características como sencilla comprensión global del sitio web y la estética de la interfaz pero también, en el contexto de la biblioteca digital, con la agilidad para acceder a la información que se desea consultar. Estos elementos se han cuidado especialmente al diseñar la página, buscando un justo equilibrio entre estética y funcionalidad que compatibilice la usabilidad con la potencia de las funciones de búsqueda y consulta de la información. Para este fin se han diseñado búsquedas básicas por los principales criterios, siglos y comunidades autónomas, así como una búsqueda avanzada que permite localizar una subcadena en cualquiera de los campos. También se ha implementado un listado de ediciones por siglo a partir del que se obtienen todos los ejemplares de cada edición. La interfaz, basada en la ficha de los ejemplares, resulta muy homogénea e intuitiva. Esta sencillez facilita la consulta del contenido ya que, suministrando funciones para la navegación ágil entre páginas y funciones, no distrae del objetivo último de mostrar de una forma comprensible y bella la información de los ejemplares y sus ilustraciones.

La confiabilidad se ha cuidado mediante la realización de pruebas y también de una forma preventiva. Para ello se ha llevado a cabo un diseño de calidad y se han empleado los estándares propuestos por la World Wide Web Consortium (W3C) en la implementación; HTML5, estilos CSS adaptativos, y tecnologías para las aplicaciones web. Estas características del software lo hacen también más fácil de mantener, una necesidad de este tipo de sistemas cuyo desarrollo se extiende de forma continuada en el tiempo.

En cuanto a la eficiencia, se ha estudiado especialmente la velocidad de carga de las imágenes. Para ello se han creado archivos con imágenes reducidas (thumbnails) para las consultas en listas y se ha reducido el peso de los archivos de las ilustraciones adaptándolas a la precisión requerida en función del marco en que se van a visualizar.

La escalabilidad es uno de los requisitos básicos de la aplicación al tratarse de un proyecto en evolución. Inicialmente se digitalizaron los ejemplares de en las bibliotecas de Galicia y Cataluña pasando posteriormente a los de Castilla León. En próximas fases se digitalizarán los de las bibliotecas de las comunidades autónomas restantes. Por esto se ha puesto especial cuidado en el diseño de la base de datos en su normalización lo que permitirá, gracias a la inexistencia de información redundante, un mejor aprovechamiento del espacio de almacenamiento y una gestión más ágil y sencilla de la información. Esta ha sido también una de las razones por las que se ha reducido el peso de las imágenes. Sus archivos se han nombrado con una codificación articulada compuesta por el código del

ejemplar y el número (1, 2...) o la posición (p portada, ap anteportada...) de la imagen. Además se han organizado los archivos en carpetas por ejemplar. De esta manera resulta más sencilla la búsqueda de los archivos y su manejo y se facilita el crecimiento de los fondos digitalizados.

Pero ¿por qué no ser más ambiciosos? La biblioteca digital puede convertirse en una herramienta realmente potente de apoyo a la investigación que proporcione servicios de altísimo valor añadido que, haciendo uso de técnicas de inteligencia artificial, guíen al investigador en su búsqueda y le permitan encontrar incluso aquello que el desconoce que existe. ¿Por qué limitar el acceso a sus recursos a los usuarios del sitio web? ¿Por qué no abrir su contenido para hacerlo accesible desde servicios generales de búsqueda o desde aplicaciones que proporcionan servicios basadas en información recogida en muchas fuentes diferentes? Nuestros siguientes pasos deberán estar encaminados a dar respuesta a estas inquietudes.

Un diseño guiado por el contenido

La biblioteca ovidiana es un caso atípico de biblioteca digital. Por regla general una biblioteca digital es una colección de volúmenes digitalizados junto con una breve información descriptiva de los mismos, parte de ella estructurada, como son los campos título, autor, edición, fecha... Nuestro caso es diferente ya que por un lado no se han digitalizado completos los ejemplares sino sólo sus ilustraciones. Se ha realizado un estudio pormenorizado de cada uno de los ejemplares y de su contenido iconográfico. Este estudio da como resultado gran cantidad de información estructurada junto con las imágenes por lo que nos encontramos frente a un sistema con una compleja estructura de información. En esta información radica la clave del éxito. Por ello el trabajo de diseño del sitio web ha girado en torno al contenido y su arquitectura.

La base para el modelado de los datos fue la ficha bibliográfica elaborada por las expertas. A partir de ella, y en función de los criterios de las búsquedas, se elaboró el modelo conceptual de datos, modelo Entidad-Relación², que se representa en la figura 1 (Modelo conceptual de datos). Este tipo de modelos resultan muy útiles para el diseño, dado su naturalidad para el modelado, al representar el mundo en términos de

- entidades, objetos del mundo real sobre los que nuestro sistema guarda información,
- atributos, o propiedades de estas entidades, y
- relaciones, vínculos entre las entidades inherentes a la semántica de los datos en el mundo real.

Para la creación de la base de datos el modelo conceptual ha de traducirse en un modelo de representación³. En la implementación del sitio web <ovidius-pictus.com> se ha empleado un modelo lógico relacional de datos cuyo esquema puede verse en la figura 2 (Modelo lógico relacional). El esquema muestra una

² El modelo Entidad-Relación propuesto por Chen (1976) es la herramienta más extendida para el modelado conceptual de datos dada su naturalidad para diseñar modelos.

³ Los modelos de datos de representación muestran un mayor número de detalles que los modelos conceptuales sin llegar con ello a mostrar detalles físicos sobre cómo los datos son almacenados ni sobre los caminos de acceso a los mismos.

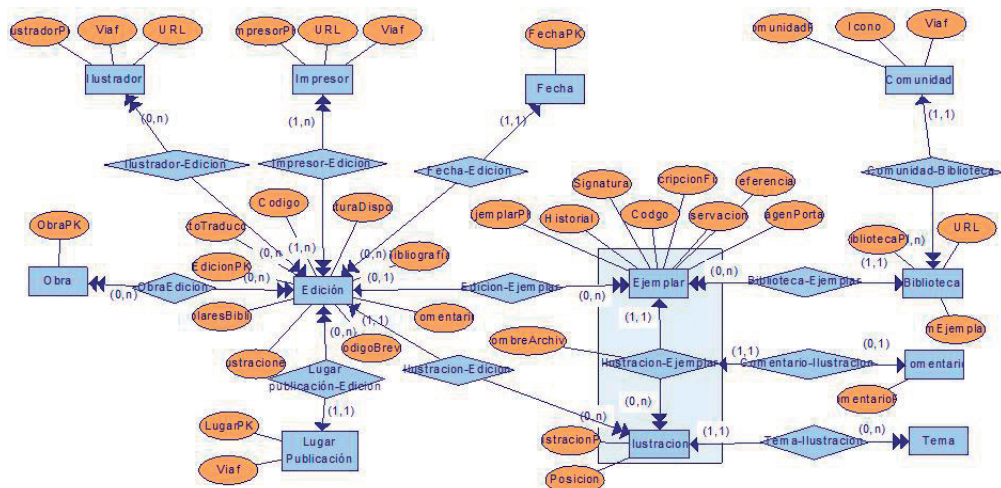


Fig. 1. Modelo conceptual de datos.

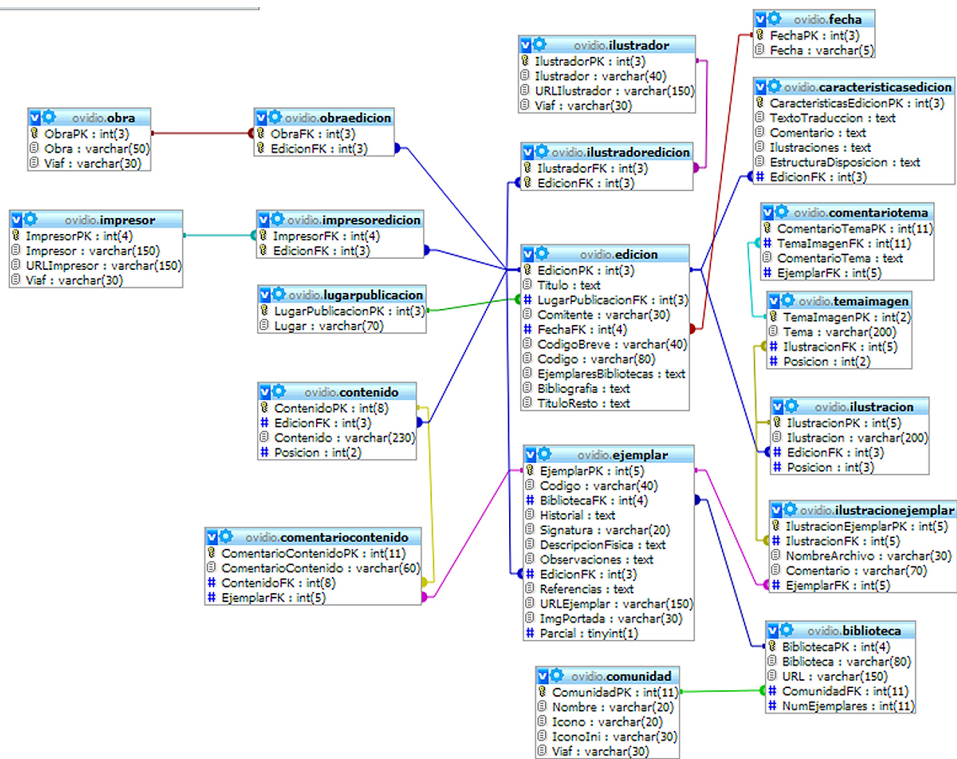


Fig. 2. Modelo lógico relacional.

estructura con grandes semejanzas a los modelos en copo de nieve empleados para modelar almacenes de datos corporativos⁴ con los criterios de búsqueda a modo de dimensiones y las ediciones y ejemplares, en el corazón del modelo, en paralelismo con las tablas de hechos. Esta semejanza calculada facilita las tareas de recuperación de la información que, al igual que en los modelos para la inteligencia de negocio, es el objetivo principal de nuestro sistema.

El diseño de la arquitectura del sitio sigue una estructura básicamente jerárquica con lo que se combinan una navegación muy intuitiva y una gran agilidad.

Se ha incluido, junto con el menú de opciones y los botones de navegación para poder avanzar y retroceder en las imágenes..., una barra de estado que informa en todo momento del significado de la información presentada. El ejemplar, núcleo de nuestro modelo de datos, es también el elemento central de la navegación y la barra de estado muestra el ejemplar activo para poder volver fácilmente a él.

EL NUEVO HORIZONTE: LA BIBLIOTECA DIGITAL OVIDIANA ABIERTA

El proceso de digitalización de contenidos bibliográficos cumple con un doble cometido. El primero la perdurabilidad, haciendo que estos contenidos no desaparezcan por el deterioro del soporte papel en el que están impresos. El segundo objetivo es hacerlos accesibles a un gran número de usuarios con el objeto de sacar el máximo rendimiento de los mismos. Esta disponibilidad se incrementa al publicarlos en la web. La Biblioteca Digital Ovidiana cumple con estos dos objetivos proporcionando una herramienta funcional y amigable mediante la que es posible acceder a la información sobre ediciones ilustradas de Ovidio en las bibliotecas españolas, a la información específica de cada uno de los ejemplares y al contenido iconográfico de los mismos.

Pero, al igual que en el ámbito profesional se llevan a cabo procesos de inteligencia de negocio en los que se emplea todo el potencial de los datos disponibles en la organización para sustentar la toma de decisiones de gestión, estos datos pueden ser empleados de forma eficiente para, mediante la investigación, extraer de ellos todo su potencial. Para ello se debe, por un lado abrir la biblioteca para incrementar la visibilidad de los datos y dotarle de contenido semántico para hacer más eficiente y automático su tratamiento al incluir junto con los objetos digitales información que posibilita a las máquinas entender su contenido.

De acuerdo con estos objetivos se está llevando a cabo el proceso de apertura, para hacer accesible su contenido a través de buscadores y otras interfaces, e integración de la biblioteca digital ovidiana en Hispana, con el fin de que sus datos puedan ser recolectados tanto por ella como por Europeana. De esta forma se incrementará la visibilidad de los datos. Además, al hacer públicos los metadatos y con el enriquecimiento semántico de los mismos se conseguirá no sólo perfeccionar los resultados obtenidos por los buscadores sino la posibilidad de implementar servicios de alto valor añadido capaces de sacar todo el partido a la riqueza de la información que guarda la biblioteca.

⁴ Más conocidos por su nombre en inglés, *Data Warehouse*. Puede verse Chaudhuri (1997) para obtener más información.

El recolector Hispana dentro del ámbito del proyecto Europeana

Europeana es un proyecto europeo que comenzó en 2005 y cuyo objeto es centralizar en una gran biblioteca digital el acceso a contenidos culturales procedentes de diferentes instituciones de países de la unión europea lo que constituirá “el punto común de acceso a los fondos de las bibliotecas, archivos y museos de toda Europa” (Comisión Europea, 2011).

En cuanto a la organización de Europeana, esta no almacena contenidos sino que se encarga de recopilar conjuntos de metadatos que describen los recursos. La responsabilidad de los documentos se organiza en tres niveles:

- Las bibliotecas digitales asociadas son las responsables del mantenimiento de los documentos y sus metadatos.
- Un agregador, a nivel estatal generalmente, agrega los datos de las bibliotecas.
- Europeana recopila los datos de los agregadores nacionales y los hace disponibles a los usuarios. Cuando un usuario quiere recuperar un documento le remite a la biblioteca de origen.

Europeana centra su actuación en cuatro líneas estratégicas (Europeana, 2012):

1. Agregar contenidos para construir una fuente abierta de patrimonio europeo confiable.
2. Facilitar la transferencia de conocimiento, la innovación y la promoción de la herencia cultural.
3. Distribuir su patrimonio libremente a los usuarios estén donde estén, siempre que lo deseen.
4. Atraer a los usuarios en las nuevas formas de participación mejorando la experiencia y promoviendo la comunicación entre los conservadores/gestores de los objetos digitales y usuarios mediante el uso de herramientas de la web 2.0.

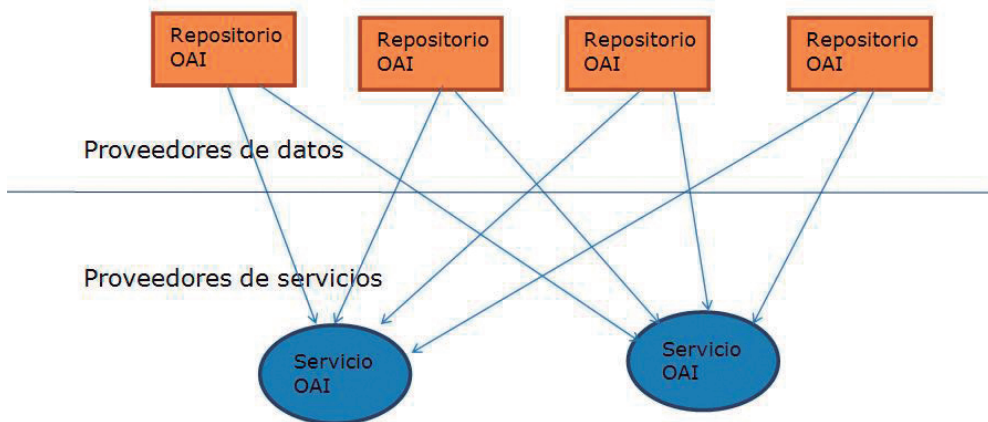


Fig. 3. Arquitectura básica del protocolo OAI-PMH.

Hispana reúne las colecciones digitales de archivos constituyendo un agregador de contenidos de las bases de datos de colecciones digitales empleando el protocolo OAI-PMH (Barrueco, 2003).

Este protocolo de interoperatividad permite la recopilación de metadatos descriptivos basados en Dublin Core⁵ en formato XML⁶. Según este protocolo existen dos tipos de agentes:

- Proveedores de datos (data providers). Gestionan los objetos y metadatos asociados a través de repositorios digitales y los ponen a disposición de otros sistemas vía OAI-PMH.
- Agregadores (havester): Recopilan los metadatos provenientes de diferentes proveedores de datos y los almacenan en una base de datos que posteriormente puede ser consultada por aplicaciones web o usuarios finales. Todo ello es gestionado por lo que se conoce por un proveedor de servicio.

El intercambio de metadatos posibilita el desarrollo de buscadores que indexen los contenidos de recursos digitales permitiendo a los usuarios acceder a información de diferentes repositorios en una sola consulta.

Hispana realiza la función de havester recopilando datos de distintas bibliotecas digitales que hacen la función de data providers. Por su parte Europeana, como ya se comentó, es un havester especial que recolecta datos de los distintos recolectores nacionales realizando así una labor de centralización.

Enriquecimiento semántico de los metadatos según las directrices de la web 2.0

La web, tal y como se concibió originalmente, consiste en un conjunto de páginas con texto conectadas entre sí mediante hipervínculos. La información contenida en ella resulta comprensible para las personas que la consultan pero no así para las máquinas que sólo ven en ella una secuencia de palabras sin estructura. Es por ello por lo que cuando queremos localizar una información mediante un buscador este nos devuelve simplemente páginas en las que aparecen los términos que buscamos y no aquellas en las que se encuentra el significado de lo que queremos consultar. Si tuviéramos, por ejemplo, la figura 4 (Consulta de un ejemplar como texto plano) la página aparecería entre los resultados de una búsqueda de "Metamorfosis Tarragona" y la ilustración de la portada en una búsqueda de "Imágenes Tarragona".

En Codina y Rovira (2006) se dan dos definiciones complementarias de web semántica⁷

"Definición 1. La visión de la Inteligencia Artificial: La web semántica es un conjunto de iniciativas destinadas a promover una futura web cuyas páginas estén organizadas, estructuradas y codificadas de tal manera que los ordenadores sean capaces de efectuar inferencias y razonar a partir de sus contenidos.

⁵ Dublin Core: Modelo de metadatos orientado a la descripción de recursos digitales.

⁶ XML: eXtensible Markup Language. Es un metalenguaje para la definición de vocabularios orientados al intercambio de datos estructurados.

⁷ En García (2005) puede verse una visión general de esta tecnología.



Fig. 4. Consulta de un ejemplar.

Definición 2. La visión del procesamiento robusto: La web semántica es un conjunto de iniciativas destinadas a convertir la World Wide Web en una gran base de datos capaz de soportar un procesamiento sistemático y consistente de la información”.

La Web Semántica (Fernández, 2008) es por tanto la nueva web comprensible no sólo por los humanos sino por sistemas inteligentes. En ella se incluyen anotaciones que describen formalmente la información, el conocimiento, que contiene. Gracias a esto será posible desarrollar sistemas que exploten este conocimiento para inferir nueva información útil para los usuarios. Las anotaciones antes citadas son los metadatos.

Según la norma ISO/UNE 23081 (ISO, 2012) “los metadatos se definen como datos que describen el contexto, contenido y estructura de los documentos, así como su gestión en el tiempo”.

Los metadatos de una biblioteca digital son la información acerca de la estructura de los datos almacenados en la misma. Según definen Hernando-De-Larramendi (2009) “aquella información que permite la descripción pormenorizada de copias digitales y de sus originales, los requisitos técnicos relacionados con la captura de las imágenes o generación de textos y el almacenamiento de los mismos y sus características, y la descripción de todos y cada uno de los procesos que dan lugar a los objetos digitales definitivos”.

Existen distintos tipos de metadatos para la gestión archivística (Rosado Pacheco y Enrique Muñido):

1. Metadatos administrativos. Para la gestión de recursos de información.
2. Metadatos descriptivos. Sirven para representar o identificar los objetos de información digital.
3. Metadatos de conservación. Destinados a gestionar la preservación de los elementos de información.
4. Metadatos técnicos. Para la gestión automática de los objetos.
5. Metadatos de uso. Relativos al nivel de utilización de los servicios de información.

Para poder ser recolectados, los metadatos deberán codificarse empleando un lenguaje, modelo, de metadatos, común. Para el proyecto Europeana se diseñó un modelo de metadatos unificado denominado ESE en el que se definían, mediante un conjunto de etiquetas, atributos como la descripción, el autor, el editor, el formato o el tipo de los objetos digitales. Este modelo, fácil de aplicar por su simplicidad, se ha comprobado que es insuficiente dada su falta de flexibilidad y su incapacidad para describir aspectos semánticos propios del contenido de los objetos de información. Esto, junto con su estructura plana que imposibilitaba la inclusión de enlaces externos y la extensibilidad a modelos más especializados, ha motivado su remplazo por el nuevo modelo de Europeana denominado European Data Model (EDM). EDM amplía las posibilidades de búsqueda y posibilita el enriquecimiento semántico de los datos soportando objetos más complejos. Según Doerr (2010) "Ofrece una mayor expresividad y flexibilidad y permite representación más rica y verdaderamente semántica de los millones de objetos de todo tipo de las comunidades de patrimonio cultural en Europeana".

Por otra parte, la Web Semántica no sólo necesitan tener acceso a los datos, también debe estar disponibles las relaciones entre ellos. Hay que crear una web de datos en lugar de una mera colección de conjuntos de datos (W3C). Esta colección de conjuntos de datos relacionados entre sí en la web también se conoce como Linked Data. Los Uniform Resource Identifier (URI), cuyo subconjunto más conocido son los URL, son un mecanismo para identificar de forma inequívoca cualquier recurso en esta red. En la Web Semántica juegan también el rol de identificadores de objetos del mundo real, físicos y no de información (Peis, 2003).

EDM, para su integración en la iniciativa Linked Data, se sustenta en la combinación de varias tecnologías o estándares para compartir datos estructurados en la web (Doerr, 2010).

- El estándar RDF (Resource Description Framework), permite representar información estructurada sobre cualquier recurso mediante triples con la forma (sujeto, predicado, objeto). Los objetos se identifican mediante URIs⁸ sobre las que se declaran propiedades y se establecen relaciones. Estas propiedades junto a los tipos de sujetos y objetos se definen en las ontologías⁹. RDF es el lenguaje de representación de la Web Semántica y el metamodelo adoptado por EDM.
- Las ontologías se definen por medio de RDF Schema (RDFS) y del estándar Web Ontology Language (OWL).
- EDM reutiliza algunas ontologías de referencia como Simple Knowledge Organization System (SKOS) que proporciona elementos para representar un sistema estructurado de conocimiento como pueden ser tesauros, clasificadores o listas de encabezados de materias.
- Dublin Core (DC) proporciona términos para describir las características esenciales de los objetos culturales (creadores, relaciones con otros recursos, índices por materias, etc.). Es un modelo sencillo que facilita además la reutilización de los datos en EDM ya que DC está muy extendido y lo usan muchas aplicaciones.

⁸ Un *Uniform Resource Identifier* o *URI* (en español "identificador uniforme de recursos") es una cadena de caracteres corta que identifica inequívocamente un recurso en la web.

⁹ Empleamos el término ontología como sinónimo de conceptualización.

- Friend-Of-A-Friend (FOAF): es una ontología pensada para describir personas y sus perfiles y se emplea fundamentalmente en redes sociales.
- OAI Object & Exchange (OAI-ORE): es el marco estructural para definir conjuntos de recursos en la red así como recursos complejos formados mediante agregación de otros más sencillos. Mediante OAI-ORE es posible asociar un objeto con diferentes instancias del mismo representadas mediante un conjunto de recursos presentes en distintas ubicaciones.

La información de la consulta antes comentada se podría representar con la estructura de metadatos mostrada en la figura 5 (Grafo RDF ejemplo).

El archivo XML del ejemplo anterior sería:

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:j.0="#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:edm="http://www.europeana.eu/schemas/edm/"
  xmlns:gss="http://www.w3.org/2001/11/IsaViz/graphstylesheets#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:ore="http://www.openarchives.org/ore/terms/"
  xmlns:dct="http://purl.org/dc/terms/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:op="http://www.ovidiuspictus.es/terminos#"
  xmlns:xhtml="http://www.w3.org/1999/xhtml/vocab#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xml:base="file:/C:/Users/Paloma/Documents/documentos/investigación/ovidio/hispana/ejemxml">
  <op:comunidad rdf:about="op:comunidad_1">
    <rdf:object>http://viaf.org/viaf/170986710</rdf:object>
    <dc:title>Cataluña</dc:title>
  </op:comunidad>
  <rdf:Description rdf:about="op:listaobras_99">
    <rdf:li>
      <op:obra rdf:about="op:obra_1">
        <rdf:object>http://viaf.org/viaf/180872106</rdf:object>
        <rdf:value>Metamorfosis</rdf:value>
      </op:obra>
    </rdf:li>
    <foaf:isPrimaryTopicOf rdf:resource="file:/C:/Users/Paloma/Documents/documentos/investigación/ovidio/hispana/ejemxml#obra"/>
  </rdf:Description>
  <rdf:Description rdf:about="op:portada_103">
    <edm:object>M.BPT.Ven.1509.p.jpg</edm:object>
    <edm:isShownAt>http://www.ovidiuspictus.es/images/images/M.BPP.Par.1802.t4/M.BPP.Par.1802.t4.p.jpg</edm:isShownAt>
    <edm:type>IMAGE</edm:type>
  </rdf:Description>
  <op:biblioteca rdf:about="op:biblioteca_20">
    <op:autonomia rdf:resource="op:comunidad_1"/>
    <dc:title>Tarragona. Biblioteca Pública de Tarragona</dc:title>
  </op:biblioteca>
  <op:edicion rdf:about="op:edicion_104">
    <dc:title>Accipe Studioso Lector<br />P. Ouidii Metamorphosin cum luculentissimis<br />Raphaelis Regii enarrationibus: quibus<br />plurima as-
```

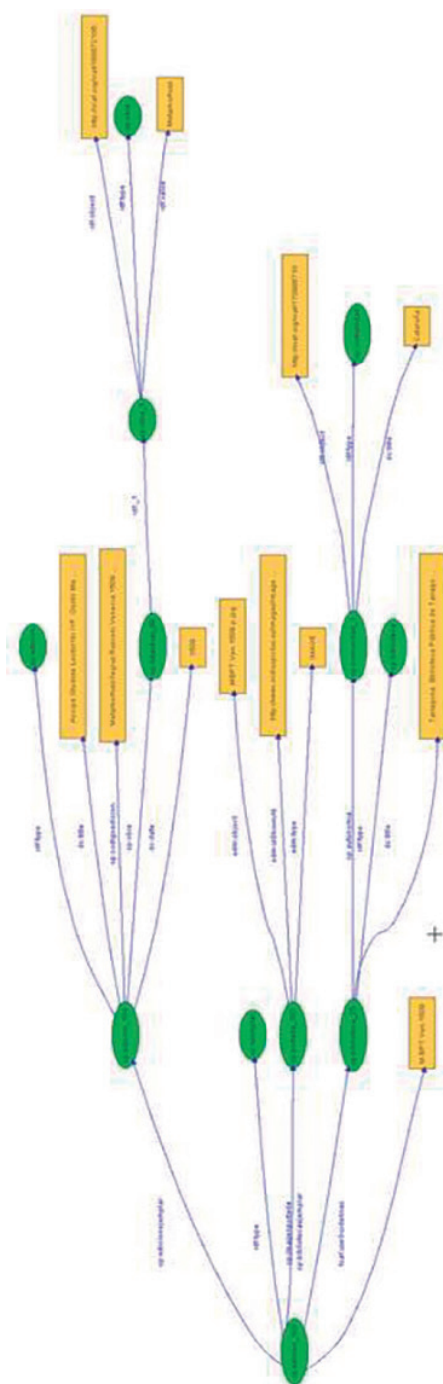


Fig. 5. Grafo RDF ejemplo.

cripta sunt: que in exempla
ribus antea impressis non inveni
untur. Que sint rogas: Inter
legendum faciletibi
ocurrent.
Cum Gratia et Privilegio.

Colofón:
Ad lectorem
Siquid forte litterar[um] immutatione. transpositione. inuersione
appositione. omissione aliaue depravatio[n]e offenderis stu
diose lector id correctionis difficultati ascribas ro
gat Georgius de Rusconibus Mediolane[n]sis
cuius industria Raphael Regius i[n] hoc
opere describendo usus est. Ve
netiis Principe felicis.
Leonardo Lau
reda[n]o die. ii.
maii. M.D.
IX.</dc:title>

<op:codigoedicion>Metamorfosis.Regius.Rusconi.Venecia.1509</op:codigoedicion>

<op:obra rdf:resource="op:listaobras_99"/>

<dc:date>1509</dc:date>

</op:edicion>

<op:ejemplar rdf:about="op:edicion_103">

<op:edicionejemplar rdf:resource="op:edicion_104"/>

<op:imagenportada rdf:resource="op:portada_103"/>

<op:bibliotecaejemplar rdf:resource="op:biblioteca_20"/>

<foaf:useGuidelines>M.BPT.Ven.1509</foaf:useGuidelines>

</op:ejemplar>

</rdf:RDF>

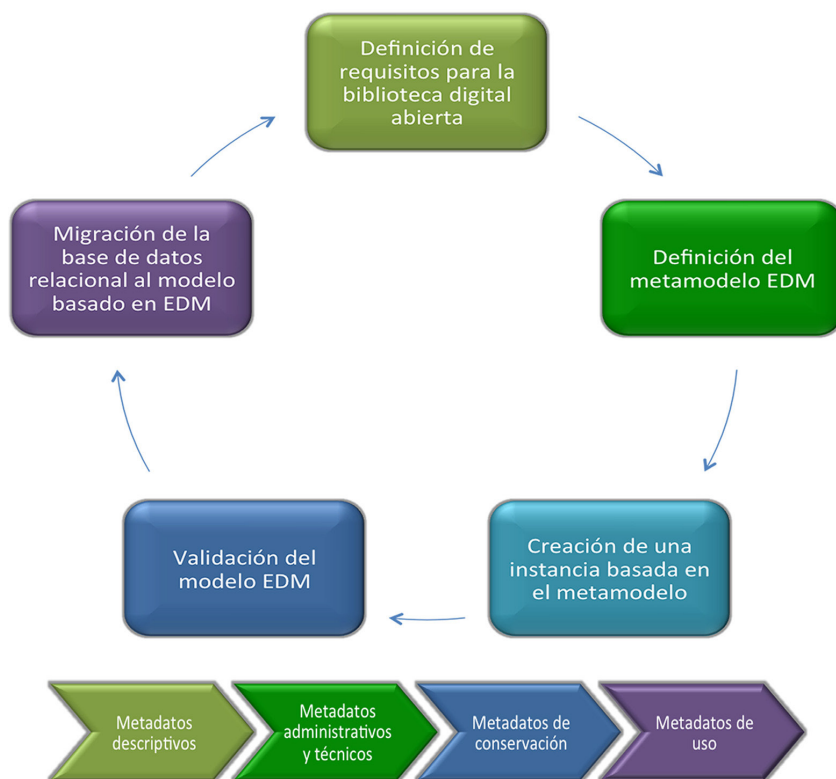


Fig. 6. Estructura evolutiva del proceso.

Fases del proyecto

Con el fin de implantar un modelo de datos acorde con las tecnologías de Web Semántica y Linked Data para la integración de la biblioteca digital ovidiana a Europeana se ha diseñado una planificación que se puede ver en figura 6 (Estructura cíclica del proceso).

Con el fin de garantizar el éxito del proyecto se ha optado por un modelo de proceso en espiral en el que el modelo irá refinándose en cada una de las iteraciones. Las iteraciones que se planean realizar son las siguientes:

En este momento nos encontramos en la primera iteración, describiendo los metadatos descriptivos. Para ello se ha definido el metamodelo a partir del modelo conceptual de datos. Se ha representado una instancia de ejemplar según el anterior modelo de metadatos. Es la correspondiente al M.BPP.Par.1802.t4, una edición de las Metamorfosis de París del siglo XIX que tiene cuatro tomos y se encuentra en la Biblioteca Pública de Pontevedra. Esta instancia deberá ser validada. En las próximas iteraciones del proceso de desarrollo se irán incluyendo los metadatos técnicos, administrativos, de gestión y de uso.

BIBLIOGRAFÍA

- Barrueco, J.M., "OAI-PMH: protocolo para la transmisión de contenidos en Internet", *El profesional de la información*, 12, 2 (2003), pp. 99-106.
- Chaudhuri, S., "An overview of data warehousing and OLAP technology", *ACM Sigmod record*, 26, 1 (1997), pp. 65-74.
- Chen, P., "The Entity-Relationship Model: Toward a Unified View of Data", *ACM Transactions on Database Systems (TODS)*, 1, 1 (1976), pp. 9-36.
- Codina, L. y Rovira, C., *La Web Semántica*, 2006.
- Comisión Europea, *European's Information Society. Thematic portal*, 27 de octubre de 2011, Iniciativa sobre Bibliotecas Digitales:: Europeana: <http://ec.europa.eu/information_society/activities/digital_libraries/europeana/index_es.htm> [recuperado el 25 de junio de 2013].
- Doerr, M.G., "El Modelo de Datos de Europeana (EDM)", *Meeting: 149. Information Technology, Cataloguing, Classification and Indexing with Knowledge Management*, Gotemburgo, Suecia, 2010: WORLD LIBRARY AND INFORMATION CONGRESS: 76TH IFLA GENERAL CONFERENCE AND ASSEMBLY.
- Europeana, *Strategic Plan 2011-2015*, 2012, <http://www.pro.europeana.eu/c/document_library/get_file?uuid=c4f19464-7504-44db-ac1e-3ddb78c922d7&groupId=10602> [Recuperado el 28 de junio de 2013].
- Fernández, R.C., *Representación del Conocimiento. Web Semántica*, Madrid, Universidad Carlos III de Madrid, 2008.
- García, N.F., "La Web Semántica: fundamentos y breve estado del arte", *Novática: Revista de la Asociación de Técnicos de Informática*, 178 (2005), pp. 6-11.
- Hernando-De-Larramendi, L.D.-M.-P.-C., "Datos y metadatos: la normalización dinámica de los elementos y de los procesos constituyentes de una Biblioteca Virtual", *FESA-BID'09*, 2009, pp. 109-116.
- ISO, *UNE-ISO/TR 23081-3 Metadatos para la gestión de documentos*, 2012.
- Offutt, J., "Quality Attributes of Web Applications", *IEEE Software*, 19, 2 (2002), pp. 25-32.
- Olsina, L.E., "Specifying Quality Characteristics and attributes for Web Sites", *Proc. 1st. ICSE Workshop on Web Engineering*, Los Angeles, ACM, 1999.
- Peis, E.H.-V., *Ontologías, metadatos y agentes: recuperación "semántica" de la información*, 2003.

Pressman, R.S., *Ingeniería del Software. Un enfoque práctico*, 7ª ed., México D.F., McGraw Hill, 2010.

Rosado Pacheco, I. y Muñido, Enrique J., "El viaje de los documentos: Del anaquel a la nube. Metadatos y principales repositorios de objetos digitales", *Códice. Gestión de la información*, s.f.

W3C, *Semantic Web*, Linked Data, s.f., <<http://www.w3.org/standards/semanticweb/data>> [Recuperado el 2 de julio de 2013].



RESUMEN

La *Biblioteca Digital Ovidiana* es un caso atípico de biblioteca digital que tiene un sistema con una compleja estructura de información y que cumple con el doble cometido de la perdurabilidad y disponibilidad de la información sobre las ediciones ilustradas de Ovidio en las bibliotecas españolas. Pero con el objeto de extraer de estos datos todo su potencial, se ha dispuesto abrir la biblioteca para incrementar la visibilidad de los datos y dotarla de contenido semántico, para hacer más eficiente y automático su tratamiento al incluir junto con los objetos digitales la información que posibilita a las máquinas entender su contenido. De acuerdo con estos objetivos, se está llevando a cabo el proceso de apertura, para hacer accesible su contenido a través de buscadores y otras interfaces, y de manera concreta para integración de la biblioteca en *Hispana*, con el fin de que sus datos puedan ser recolectados tanto por ella como por *Europeana*. De esta forma se incrementará la visibilidad de los datos y, al hacer públicos los metadatos y con el enriquecimiento semántico de los mismos, se conseguirá no sólo perfeccionar los resultados obtenidos por los buscadores, sino la posibilidad de implementar servicios de alto valor añadido capaces de sacar todo el partido a la riqueza de la información que guarda la biblioteca.

Palabras clave: Biblioteca Digital, Europeana, Web Semántica, Calidad Web.

ABSTRACT

The *Biblioteca Digital Ovidiana*, which meets the double mission of providing durability and availability to the information about the Ovid illustrated editions kept in Spanish libraries.

However, it has been arranged to become an open library, in order to extract the full data potential and be enable to also increase the data visibility. Additionally, the semantic content, which includes digital objects information that enables machines to understand its content, allows automatic data treatment. According to these goals, the opening process is being carried out to make the content accessible through search engines and other interfaces, including the integration in *Hispana*, and so that the data could be also harvested by *Europeana*. This will increase the visibility of the data and, by making public the metadata and by their semantic enrichment, will be improved not only the results obtained by the search engines, but also the possibility of implementing value-added services, capable of take full advantage of the rich information stored in the library.

Keywords: Digital Library, Europeana, Semantic Web, Web Quality.