

Hacia la arquitectura de la información 3.0: pasado, presente y futuro

Por Roberto García, Federico Botella y Mari-Carmen Marcos



Roberto García es ingeniero en informática por la Univ. Politècnica de Catalunya (UPC), Máster en Comercio Electrónico por La Salle Barcelona y doctor por la Univ. Pompeu Fabra (UPF). Investiga la web semántica, su aplicación a la gestión de contenidos y su propiedad intelectual. Es profesor de la Univ. de Lleida (UdL) y miembro del grupo de investigación en IPO (Interacción Persona Ordenador) e integración de datos Griho (Grup de Recerca en Interacció Persona Ordenador i Integració de Dades), donde investiga cómo hacer llegar la web semántica a los usuarios. Coordina las asignaturas de arquitectura de la información en el Máster de Interacción Persona-Ordenador de la UdL.
<http://rhizomik.net/~roberto>



Federico Botella es ingeniero en informática y doctor ingeniero en informática por la Universitat d'Alacant. Es profesor titular en el Departamento de Estadística, Matemáticas e Informática de la Universidad Miguel Hernández de Elche (UMH) donde imparte docencia en temas de arquitectura de la información, usabilidad y accesibilidad. Es investigador asociado al Instituto Universitario Centro de Investigación Operativa y miembro del grupo de investigación registrado en la UMH "Sistemas web de soporte a las decisiones (Webdecision)" donde lleva a cabo su investigación sobre personalización y minería de uso de la Web.
<http://webdecision.umh.es>



Mari-Carmen Marcos es doctora y licenciada en documentación. Es profesora de documentación en el Departamento de Comunicación de la Universitat Pompeu Fabra de Barcelona. Coordina y da clases en el Máster Online en Documentación Digital. Es codirectora y profesora del Máster Online en Buscadores, además de participar en otros cursos de postgrado. Su línea de investigación es el diseño de interacción en la recuperación de información en la web, que lleva a cabo en los grupos DigiDoc y Grupo de Investigación de la Web de la UPF. Miembro del think tank Estrategia y Prospectiva de la Información (ThinkEPI).
<http://www.mcmarcos.com>

Resumen: La arquitectura de la información (IA), aunque presente en cierta manera desde hace más de 50 años, ha tomado especial relevancia con la llegada de la Web y el aumento de la cantidad de información a nuestro alcance. La disciplina cuenta con unos pilares claros, que se fundamentan en los contenidos gestionados pero sobre todo en los usuarios, sin olvidar su contexto. También se identifica una serie de componentes a combinar a la hora de elaborar una IA y una metodología básica. Pero la cantidad de información a nuestro alcance no para de crecer, así que se continúa innovando en nuevas formas de organización, buscadores más eficaces o tecnologías de web semántica que permiten a los ordenadores encargarse de la parte menos creativa de enfrentarse a esa avalancha de información.

Palabras clave: Arquitectura de la información, Navegación, Organización, Buscador interno, Comportamiento del usuario, Experiencia de uso, Web semántica, Web 3.0, Ontologías, Metadatos.

Title: Towards information architecture 3.0: past, present and future

Abstract: The discipline of information architecture (IA), although present in some way for over 50 years, has taken on special relevance with the advent of the Web and the explosion of the amount of available information. Nowadays, IA has clear pillars based on the managed content but especially on users, without forgetting the context. This article discusses components to combine when developing IA, and a basic methodology is identified. As the amount of information around us is still growing, IA continues to innovate with new forms of organization, more efficient search engines and Semantic Web technologies that let computers take care of the less creative part of coping with the avalanche of information.

Keywords: Information architecture, Navigation, Browsing, Organization, Internal search engine, User behavior, User experience, Semantic web, Web 3.0, Ontology, Metadata.

García, Roberto; Botella, Federico; Marcos, Mari-Carmen. "Hacia la arquitectura de la información 3.0: pasado, presente y futuro". *El profesional de la información*, 2010, julio-agosto, v. 19, n. 4, pp. 339-347.

DOI: 10.3145/epi.2010.jul.02

Artículo recibido el 18-05-10

Aceptación definitiva: 15-06-10

1. Introducción

1.1. ¿Qué es la arquitectura de la información?

El término arquitecto de la información suele atribuirse a **Richard S. Wurman** en su libro *Information architects* de 1997 (Wurman, 1997) donde lo define como la persona que organiza patrones inherentes a los datos haciendo claro lo complejo. Sin embargo, ya en 1959 dos miembros del *Departamento de Organización de Máquinas* de IBM, **Lyle R. Johnson** y **Frederick P. Brook**, comienzan a usar el término arquitectura en el campo de la informática (Ronda-León, 2008). Pero fueron **Rosenfeld** y **Morville** quienes popularizaron la profesión de arquitecto de la información con la publicación de su libro *Information architecture for the World Wide Web* en 1998 (el famoso libro del oso polar, por tener uno dibujado en su cubierta) (Rosenfeld; Morville, 1998). En este libro se extiende el concepto de arquitectura de la información (IA) al diseño de sitios web y se establecen los principios fundamentales de esta disciplina.

http://es.wikipedia.org/wiki/Arquitectura_de_la_información

Podemos definir de manera informal el concepto de arquitectura de la información como:

- Diseño estructural de un espacio de información compartido.
- Combinación de organización, etiquetado y esquemas de navegación para sitios web e intranets.

El arte y la ciencia de organizar información para ayudar a las personas a desempeñar eficientemente sus necesidades, mejorando la usabilidad y la capacidad de los usuarios de encontrar lo que buscan.

- Disciplina y comunidad de práctica que busca llevar los principios de diseño y arquitectura al mundo digital.

En la siguiente sección profundizaremos en la definición de IA a partir de sus profesionales: los arquitectos de información.

1.2. La profesión del arquitecto de información

Uno de los mayores retos a los que se enfrentan estos profesionales es el tener que vender su trabajo a sus clientes y/o a sus jefes ya que, a diferencia del de un arquitecto convencional, no es un trabajo visible a primera vista. Incluso en ocasiones es difícil de percibir por otro colega, por lo que el arquitecto debe esforzarse constantemente en explicar y demostrar los beneficios que aporta la IA en las áreas donde se aplica. Es difícil calcular de manera exacta el retorno de una inversión en IA ya que hay muchos factores implicados.

Determinar dónde empieza y dónde termina la IA es complicado. Comprende desde el diseño gráfico

hasta la ingeniería de la usabilidad, pasando por el periodismo o la biblioteconomía, pero cualquier campo basado en la información y su uso será un buen lugar donde aplicar arquitectura de la información.

1.3. Conocer al usuario

La arquitectura de la información pone en su punto de mira a los usuarios, y en particular a sus necesidades de información. Éstas pueden variar ampliamente, y hacen que los usuarios muestren conductas específicas de búsqueda. Los arquitectos de la información tienen que conocer esas necesidades y comportamientos para hacer sus diseños acordes con ellos.

A veces los usuarios abandonan la búsqueda sin haber conseguido su objetivo, en un estado de satisfacción parcial o de frustración. También puede suceder que durante el proceso encuentren una información que les haga cambiar parcial o totalmente sus planes. El comportamiento de los usuarios puede conocerse, por ejemplo, estudiando los ficheros de logs –registro de todas las páginas visitadas de un sitio web-. De esta forma se conocen las rutas de navegación más comunes, el tiempo de permanencia en cada página, en cuál se abandona, etc.

SCIPEDIA
“El trabajo de arquitecto de la información, a diferencia del de un arquitecto convencional, no es visible a primera vista”

Register for free at <https://www.scipedia.com> to download the version without the watermark

2. Principios y componentes

Una buena IA debe ser invisible para los usuarios, no interferir sino ayudar a su interacción. A primera vista, los elementos más prominentes son los relacionados con el diseño gráfico, como la gama de colores, la organización del texto con diferentes disposiciones en columnas, o las tipografías. Pero mirando más detenidamente se detectan cuatro sistemas:

- Organización: presentan la información del sitio de diferentes maneras, siguiendo esquemas que agrupan o diferencian componentes de información, por ejemplo, siguiendo un orden cronológico o teniendo en cuenta las audiencias a las que se dirige.

- Navegación: ayudan a los usuarios a moverse por el contenido, como por ejemplo los menús o el índice del sitio.

- Etiquetado: describen categorías, opciones y enlaces utilizando términos significativos, por ejemplo las etiquetas utilizadas en los menús de navegación.

– Búsqueda: permiten que los usuarios busquen contenidos concretos, habitualmente mediante palabras clave. Además, se pueden ofrecer mecanismos para acotar la búsqueda.

Podemos entonces empezar a estudiar estos elementos en las páginas de cualquier sitio web, especialmente en las *homes* o páginas principales. Cuando están bien diseñadas, éstas tratan de anticipar las necesidades generales de información de los usuarios, encaminándolos hacia los contenidos que les puedan interesar. La forma más común de elaborar una arquitectura de la información es ir de lo general al detalle (*top-down*).

La otra opción, denominada del detalle a lo general (*bottom-up*), enfoca el problema partiendo de contenidos individuales y no de una visión general. Esos contenidos están descritos mediante metadatos y palabras clave.

A partir de ellos se construyen mecanismos que ayudan a los usuarios a entender de qué trata el contenido, cómo usarlo o cómo encontrarlo. Esta aproximación *bottom-up* está adquiriendo importancia ya que cada vez es más común que los usuarios lleguen hasta los contenidos de un sitio sin pasar por la página principal, desde buscadores, enlaces desde otros sitios o agregadores de noticias.

Por lo tanto, una buena IA, aparte de los aspectos *top-down* más clásicos, debe ser robusta y tener en cuenta aspectos *bottom-up* que permitan una navegación simple y práctica desde cualquier punto del sitio. A continuación veremos con qué piezas cuenta un arquitecto de la información para tal fin, y cómo proceder para combinarlas.

“Nuestro conocimiento del mundo depende de nuestra habilidad para organizar la información”

2.1. Sistemas de organización

Nuestro conocimiento del mundo viene determinado por nuestra habilidad para organizar la información. La forma en que organizamos, etiquetamos e interrelacionamos la información influye en la manera en que las personas la comprenden.

Los arquitectos la organizan y etiquetan mediante esquemas con el propósito de que los usuarios sepan encontrar respuestas a sus preguntas.

Un esquema de organización define las características compartidas por los diferentes elementos de contenido y la forma de agruparlos:

a) Esquemas exactos: dividen la información en

grupos bien definidos y mutuamente excluyentes. Son de uso frecuente: orden alfabético, cronológico y geográfico. Son los más adecuados cuando se conoce el elemento a buscar.

b) Esquemas ambiguos: dividen la información en categorías que no tienen una definición exacta. Son difíciles de diseñar y mantener, y también pueden ser difíciles de usar. Sin embargo, son a menudo más útiles que los anteriores porque no siempre se sabe lo que se está buscando. Ejemplos: por tema, por tarea o por audiencia.

c) Esquemas híbridos: combinan diferentes esquemas de organización. Deben combinarse con cuidado porque pueden resultar confusos al mezclar diferentes formas de ver la información. Se recomienda utilizar un diseño claro que permita al usuario distinguir claramente los diferentes criterios utilizados, por ejemplo, enmarcando cada uno en bloques diferentes.

Por otra parte se definen los tipos de relaciones entre elementos de contenido, que pueden ser: de jerarquía, de metadatos y de hipertexto:

– Jerarquía: es la estructura clave para una organización *top-down*, de lo general a lo particular. Hay que cuidar un adecuado balance entre amplitud (opciones por cada nivel de la jerarquía) y profundidad (número de niveles en la jerarquía). Respecto a la amplitud, el factor limitante son las habilidades de exploración visual de las personas. Respecto a la profundidad, si los usuarios deben hacer clic por más de dos o tres niveles, seguramente se cansarán. Por ejemplo, el trabajo de Miller y Remington (2004) muestra que para 512 elementos y en las condiciones más comunes de cierto nivel de ambigüedad entre las opciones, la más usable es una de 2 niveles (con 16 opciones en el primero y 32 en el segundo, o viceversa). Los resultados son peores para la de 3 niveles (8 opciones en cada nivel).

– Metadatos: es la estructura típica para aproximación *bottom-up*, de lo detallado a lo general. Etiquetando documentos y otros objetos con metadatos de vocabularios controlados es posible la búsqueda, navegación y filtrado avanzados, así como los vínculos, por ejemplo a productos relacionados.

– Hipertexto: en esta estructura se encuentran elementos de información y enlaces que los interrelacionan. Aporta gran flexibilidad pero comporta un riesgo de complejidad y confusión. Sin un contexto claro los usuarios pueden perderse fácilmente en la red de enlaces. Raramente es la estructura de organización principal. Normalmente complementa estructuras jerárquicas o basadas en metadatos.

2.2. Sistemas de navegación

Utilizamos herramientas de navegación para trazar nuestro recorrido, para determinar nuestra posición y

Register for free at <https://www.scipedia.com> to download the version without the watermark

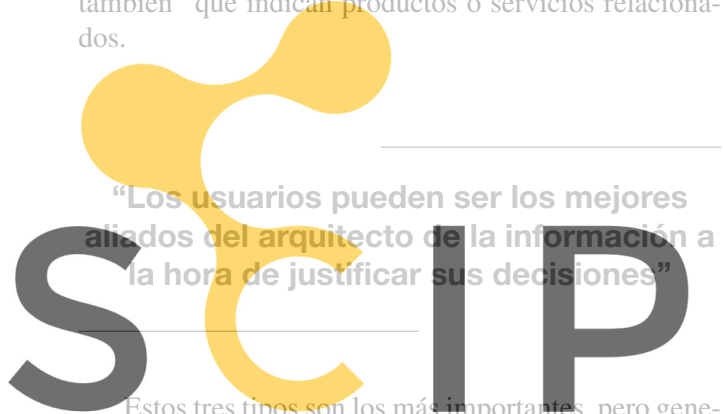
para encontrar el camino de vuelta en una arquitectura de la información.

Tipos:

a) Global: concebidos para estar presentes en todo momento, normalmente bajo la forma de una barra de navegación en la parte superior.

b) Local: el sistema de navegación global se complementa con uno o más sistemas locales que permiten explorar áreas concretas. Por ejemplo, en un periódico la barra de navegación global se amplía para proporcionar opciones locales de navegación para cada categoría de noticias.

c) Contextual: algunas relaciones entre contenidos no se ajustan a la organización jerárquica de la navegación global o local, dependen del contexto actual donde se encuentra el usuario. Por ejemplo los enlaces “Ver también” que indican productos o servicios relacionados.



Estos tres tipos son los más importantes, pero generalmente se complementan con otros como mapas del sitio o índices alfabéticos que permiten una vista global de un sitio web. Los índices alfabéticos permiten el acceso directo a los contenidos. Y las guías ofrecen normalmente navegación lineal ajustada a los requisitos particulares de una audiencia, tarea u objetivo específico.

2.3. Sistemas de etiquetado

Los sistemas anteriores se basan en el uso de etiquetas textuales e icónicas que permiten representar el contenido que hay detrás de una opción del sistema de navegación o un enlace. Por ejemplo, la etiqueta “Contactar con nosotros” representa el contenido que posiblemente incluye dirección, teléfono y fax.

El objetivo de una etiqueta es comunicar la información eficientemente, sin ocupar demasiado en la pantalla ni en espacio cognitivo del usuario. Los principales tipos de etiquetas de texto son:

- Enlaces contextuales: enlaces a trozos de información en otras páginas o en otros puntos de la misma página.
- Encabezamientos: etiquetas como los titulares que describen el contenido que les sigue.

– Opciones de los sistemas de navegación: etiquetas que representan las opciones disponibles.

– Términos de indexado: palabras clave, etiquetas (*tags*) y los encabezamientos que representan contenido para la búsqueda o la navegación.

2.4. Sistemas de búsqueda

En un modelo simple de recuperación de información el usuario llega al sitio web con una necesidad de información, realiza una acción (por navegación o por búsqueda) y recibe una respuesta. Este modelo no suele ser adecuado, pues se centra estrictamente en lo que ocurre mientras el usuario está interactuando. Se omite el contexto de las necesidades de información (todo lo que ocurre antes y después de que el usuario pulse el botón), pasa por alto la comprensión de lo que pasa por la mente del usuario.

Para plantear un modelo de recuperación de información útil hay que contemplar al menos cuatro situaciones de búsqueda:

a) Item conocido: se sabe lo que se está buscando y cómo encontrarlo.

b) Exploratoria: no se está seguro de lo que está buscando y se va aprendiendo con el proceso de búsqueda.

c) Exhaustiva: se quiere todo lo que hay sobre un tema concreto.

d) “Rebúsquedas”: recuperar una información útil que previamente se había encontrado.

La siguiente sección abarca de forma más específica el tema del buscador, ahora ya pensando en el entorno de la Web, dada la importancia que tiene para la IA de sitios web.

3. El buscador interno del sitio web

Lou Rosenfeld, en un trabajo publicado en febrero de 2008 afirma que para llegar a encontrar una información en una web podemos optar por tres caminos:

- navegar (*browsing*)
- hacer una consulta en el buscador local (*searching*)
- preguntar a las personas responsables del sitio web (*asking*)

Estas tres opciones deben estar presentes de forma simultánea de manera que el usuario opte por la que mejor le parezca en cada momento. En la mayoría de los sitios web los usuarios usan primero la navegación, si no encuentran lo que necesitan usan el buscador, y si así tampoco, envían un mensaje a la dirección de contacto de la web.

Register for free at <https://www.scipedia.com> to download the version without the watermark

El buscador, además de servir como una opción más para satisfacer la necesidad de información del usuario, resulta ser un elemento clave para conocerlo, y, como se dijo, podemos hacerlo a partir de los ficheros log que genera. Gracias al buscador podemos saber:

a) En qué momento han dejado de navegar por el sitio para plantear una consulta en el buscador, señal de que muy posiblemente no han encontrado lo que buscan navegando, síntoma de un posible problema en el sistema de organización, de navegación o de etiquetado.

b) En qué momento el buscador da paso a la navegación, posiblemente porque no han encontrado lo que necesitaban, señal de un posible error en el buscador o bien debido a que las palabras usadas por el usuario no son las que utiliza el sitio web.

c) Qué busca exactamente, si realiza una segunda consulta o refina la que había introducido.

d) Qué vocabulario usa, que no es siempre el que hemos implementado, y nos permite tomar decisiones de cara a hacer cambios en el sitio web.

Con esto no queremos decir que los posibles problemas de navegación que tenga un sitio web se tengan que solventar con la instalación de un buscador interno. El buscador es una forma más de acceso a la información, no sustituye a una correcta navegación; es más, una correcta arquitectura en cuanto a organización, navegación y etiquetado ayudarán a que el buscador sea más eficiente.

– Están estructurados en campos, por ejemplo el catálogo de una biblioteca en el que los usuarios consultan las obras de un autor o buscan en el campo del título.

– Han ido creciendo de forma natural, sin que se haya planificado, y eso ha hecho que la navegación original no dé cabida a todo de una forma coherente.

b) El sitio tiene los contenidos fragmentados, es decir, ha habido distintos webmasters y probablemente distintas políticas de creación de las distintas partes del sitio. Es el caso que encontramos a menudo en los sitios web de universidades, donde cada departamento o cada facultad cuenta con un espacio en el servidor, pero bajo el mismo dominio, donde crear y mantener sus contenidos. En casos así, un buscador permite encontrar información en todo el dominio, sin tener que aprender la arquitectura de cada parte.

c) El sitio tiene contenidos dinámicos, esto es, se actualizan con mucha frecuencia de manera que lo publicado anteriormente cambia de lugar para dar paso a la información más actual. El caso más común es el de los sitios web de prensa. En estos casos, un buscador será de gran utilidad para localizar contenidos de días anteriores sobre un determinado tema.

Otras veces son los usuarios del sitio los que hacen que debamos tener un buscador, pues ellos esperan que lo haya porque lo han visto siempre en otros sitios de contenidos similares al nuestro: la web de una biblioteca, de una librería, de una tienda online, de un periódico o de una agencia de viajes.

En otras ocasiones son los webmasters quienes están interesados en contar con un buscador porque quieren hacer uso de los ficheros log para conocer el comportamiento de los usuarios y mejorar la arquitectura del sitio web.

Ahora bien, todas estas situaciones no tendrán el peso suficiente para que nos decidamos a instalar un buscador si no se da la condición más importante: tener tiempo y conocimientos para configurarlo correctamente y mantenerlo.

A pesar de ello, algunos webmasters insisten en implementar un buscador en el sitio web como la gran solución a un problema de navegación. Evidentemente esto es un error, no es nada recomendable instalar un buscador si la arquitectura no es correcta, pues no se podrá sacar de él todo el partido posible. ¿Por qué? Simplemente porque habrá problemas de etiquetado de contenidos, es decir, el buscador no podrá encontrar información cuando el usuario realice su consulta con una palabra y en el sitio web se haya usado otra para indicar lo mismo.

Por otro lado, a veces los webmasters tienen que buscar soluciones alternativas. Si no les dejan cambiar

SCIPEDIA

Register for free at <https://www.scipedia.com> to download the version without the watermark

“Google y Yahoo! utilizan tecnologías de web semántica para mejorar las búsquedas, y se observa un mejor posicionamiento de los sitios que las incorporan”

3.1. ¿Necesita el sitio web un buscador interno?

No siempre resulta imprescindible un buscador. **Rosenfeld y Morville** (1998) nos guían sobre qué tener en cuenta a la hora de decidir si el sitio web debe tener un buscador interno. Veamos tres casos:

a) Hay muchos contenidos navegables. En tal caso será recomendable instalar un buscador, pero más aún si los contenidos cumplen alguna de estas características:

– Son técnicos y, por lo tanto, sus usuarios conocen las palabras específicas para hacer las consultas; por ejemplo los modelos de locomotoras de una determinada marca.

la arquitectura, al menos un buscador podría ayudar a mejorar los problemas de recuperación de información del sitio. Aun así, si estuviéramos en este caso, deberíamos pensar bien si lo que necesitamos es un buscador. Si los contenidos no son excesivos, no están fragmentados, no son dinámicos y no hay otras motivaciones que lo hagan imprescindible, en lugar de un buscador sería mejor plantearse la construcción de un índice de tipo A-Z con las palabras más importantes.

3.2. Consideraciones previas a la instalación del buscador

Rosenfeld y Morville (1998) insisten en la utilidad de los datos de nuestra web para mejorarla y plantean una cuestión muy interesante: ¿quién debe hacerse cargo de todo lo que está relacionado con el buscador local? Por supuesto los técnicos, que saben cómo implementarlo. En cambio, para que un buscador tenga éxito, hay otros componentes que no son técnicos y que son fundamentales:

- calidad de los contenidos del sitio web;
- redacción de los contenidos;
- interfaz de búsqueda, de resultados y de reformulación de la búsqueda.

La complejidad para la implementación de un buen buscador ha hecho que este trabajo quede en manos de los técnicos, mientras que los diseñadores web se han dedicado a otras tareas. En cambio es necesario que los diseñadores formen parte del equipo de elaboración del buscador para asesorar en cuanto a los contenidos y la redacción, así como para definir las distintas interfaces del buscador.

Una tarea más que pueden (y deben) realizar los diseñadores de interacción (y entre ellos los arquitectos) es analizar el comportamiento de los usuarios mediante las consultas realizadas y los clics que siguen a las consultas. La interpretación de estos datos no tiene por qué recaer en los técnicos, pues afecta sobre todo a la experiencia de uso y de ellos van a salir propuestas para cambiar la arquitectura del sitio.

4. La IA en la elaboración de sistemas interactivos

La labor del arquitecto de la información se enmarca dentro del proceso de elaboración típico de un sistema interactivo, preferiblemente uno centrado en el usuario en el que además de las fases comunes (análisis, diseño, implementación y mantenimiento), se contemplen otras más enfocadas a la participación (como prototipado y evaluación con usuarios) (Granollers et al., 2005). La IA se decide fundamentalmente durante las fases de análisis y diseño, aunque teniendo muy en cuenta la visión del usuario mediante prototipos y pruebas.

4.1. Análisis

Desde una aproximación *top-down* se definen los dos o tres niveles superiores de las estructuras de organización y navegación del sitio. Desde el punto de vista *bottom-up* se detectan tipos de documentos y un primer esquema de metadatos. Para ello se estudian los contenidos, los usuarios y el contexto.

Se analiza una muestra representativa del contenido para revelar los patrones y las relaciones existentes, y de los metadatos que se pueden utilizar para mejorar la estructura, organizar y proporcionar acceso al mismo. Para cada elemento de contenido se determinan metadatos estructurales (título, secciones), metadatos descriptivos (tema, idioma) y metadatos administrativos (autor, fecha de creación).

Normalmente se considera apropiado integrar unos cuantos métodos de estudio de usuarios y tener en cuenta que los usuarios pueden ser los mejores aliados del arquitecto de la información a la hora de justificar las decisiones de diseño o rediseño de una web. Por una parte tenemos técnicas de análisis de datos generados por los usuarios: estadísticas de uso, como los citados logs, registros de búsquedas, que indican en qué términos está interesado el usuario, y los datos de atención al cliente o soporte técnico.

Por otro lado tenemos técnicas de análisis con grupos de usuarios, como las encuestas, los estudios de campo y las reuniones con grupos de usuarios (*focus groups*). Finalmente, tenemos las técnicas que estudian a los usuarios de manera individual, como las entrevistas con implicados, las pruebas con usuarios y las ordenaciones de tarjetas (*cardsorting*).

Este método se basa en un conjunto de tarjetas etiquetadas con categorías, subcategorías y contenidos individuales dentro del sitio. Se pide al usuario que organice en pilas unas decenas de ellas y que dé un nombre a cada pila. Para llevar a cabo las ordenaciones se puede utilizar simplemente lápiz y papel, pero también existen herramientas informáticas que incluso permiten realizarlas remotamente y analizar los resultados mediante métodos estadísticos. Normalmente, con unos 15 usuarios será suficiente (Nielsen, 2004) para obtener un modelo mental común de los usuarios para organizar la información.

Es importante no basarse únicamente en los valores numéricos que se obtienen por el análisis estadístico de los resultados del *cardsorting*. Mucha de la información que se obtiene viene de escuchar los comentarios de los usuarios mientras organizan las tarjetas: conociendo por qué las personas colocan ciertas tarjetas juntas se obtiene una visión más profunda de sus modelos mentales que del simple hecho de clasificar tarjetas en la misma pila.

4.2. Diseño

Otra parte importante de la labor del arquitecto de la información se lleva a cabo durante la fase de diseño, cuando se establecen las bases para la implementación creando modelos, maquetas y esquemas detallados de los metadatos, que los diseñadores gráficos, programadores, autores de contenidos y el resto de equipo utilizarán.

Los modelos muestran las relaciones entre las páginas y otros componentes de contenido, y se pueden utilizar para ilustrar la organización, la navegación y los sistemas de etiquetado. Son una herramienta excelente para explicar diferentes aproximaciones arquitectónicas y discutir las con los clientes o colaboradores. En la figura 1 se muestra un modelo a tres niveles de granularidad que incluye grupos de páginas, páginas y componentes de página.

Se trata de un modelo sencillo, de los primeros pasos en la fase de diseño. A medida que va evolucionando la organización que permite con usuarios y otros implicados, el modelo se va detallando con la jerarquía completa de la información desde la página principal a las páginas destino o flechas que capturan el flujo de navegación.

Las maquetas muestran una perspectiva arquitectónica de una página o una plantilla individual, por ejemplo dónde se sitúan los sistemas de navegación tal y como se ve en la figura 2. Las maquetas constituyen la intersección entre IA y diseño gráfico. Se emplean en las páginas más importantes del sitio tales como páginas principales, las interfaces de búsqueda o las plantillas que se aplican a múltiples páginas.

Finalmente, antes de pasar a la implementación, el arquitecto de la información define los metadatos que se utilizan para describir los contenidos partiendo de los ya identificados durante el análisis. Fundamentalmente determina los valores de aquellos metadatos que están delimitados por algún tipo de vocabulario controlado (por ejemplo idioma o moneda) y los que por el

contrario corresponderán a texto libre o valores no acotados (por ejemplo título o descripción). Los vocabularios controlados pueden ser simples listas de términos equivalentes o los más complejos tesauros, que además de relaciones de equivalencia incluyen relaciones jerárquicas y asociaciones, como por ejemplo la asociación “término relacionado” entre “coche” y “taller”.

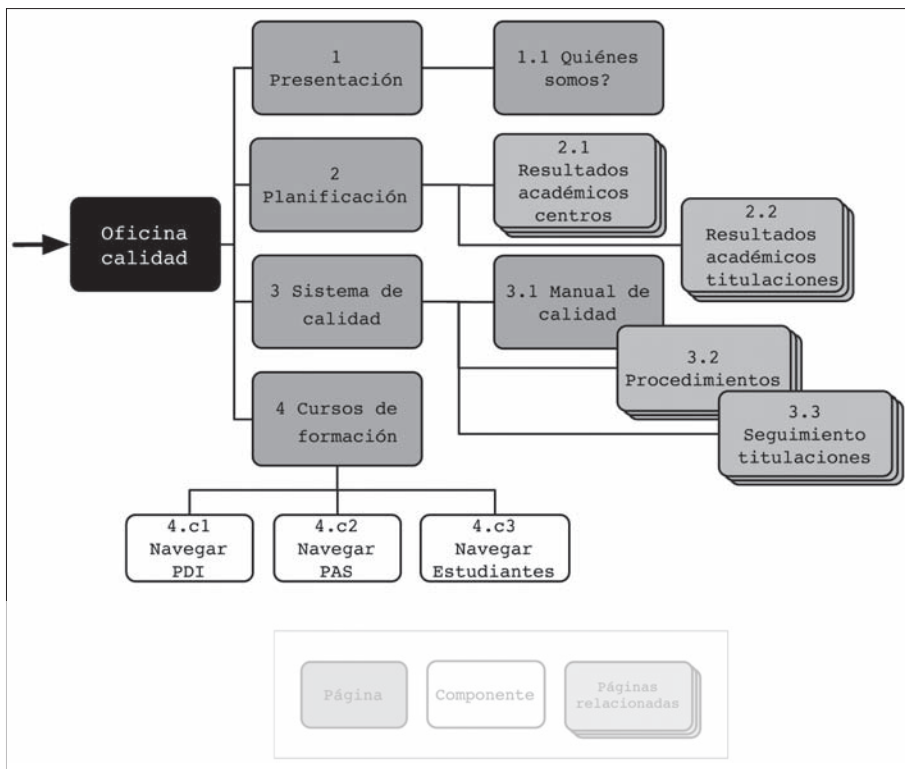


Figura 1. Modelo de una IA elaborada en la fase de diseño



Figura 2. Maqueta de una IA elaborada en la fase de diseño

“Con las ontologías de web semántica se pretende tener un modelo que integre todos los elementos de descripción de los contenidos”

5. IA y web semántica

La web semántica proporciona un conjunto de herramientas de representación de significados más complejas y potentes que las habitualmente utilizadas en IA. Por ejemplo, considerando los tesauros, los más sofisticados de los vocabularios controlados mencionados hasta el momento, permiten evitar las ambigüedades ya que operan a nivel de términos y no de significados.

Con las ontologías de web semántica, una herramienta de representación de conocimiento más potente que los tesauros, se pretende tener un modelo que integre todos los elementos de descripción de los contenidos, tanto los atributos a utilizar para describirlos como los posibles valores que éstos pueden tomar.

Las ontologías se asemejan bastante a herramientas de modelado de bases de datos o de programación orientada a objetos. La principal ventaja, aparte de tener un mayor grado de expresividad, es que las ontologías web están pensadas para poder ser publicadas en la Web.

Esto facilita compartirlas y reutilizarlas, ya que las clases (categorías), las propiedades (atributos y relaciones) o los valores que éstas pueden tomar quedan publicadas en la Web mediante una url que se puede utilizar para recuperar su definición y para reutilizarla en nuestra IA.

Pero no sólo las ontologías quedan visibles en la Web para facilitar su reutilización, es posible publicar también las descripciones que hagamos con ellas, los metadatos concretos que utilizamos para describir un elemento de contenido. De esta forma se fomenta la reutilización también de los metadatos permitiendo construir aplicaciones web que mejoran la experiencia del usuario al combinar información de múltiples fuentes, lo que se denomina *mashup* o remezcla, por ejemplo de la iniciativa de *Datos enlazados abiertos* (Bizer et al., 2009).

5.1. Tecnologías de web semántica

La web semántica (Pollock, 2009) incluye una serie de lenguajes de representación que permiten tener ontologías y metadatos publicados en la web. La base es rdf (*resource description framework* o marco de descripción de recursos). Se trata de un lenguaje para representar metadatos. Sobre esta base se han elaborado varios lenguajes para representar ontologías web. El más sencillo es rdfs (*rdf schema*). Para representar ontologías más complejas está owl (*web ontology language* o lenguaje de ontologías web). Cabe destacar también sparql, un lenguaje para hacer consultas sobre metadatos semánticos y ontologías.

<http://www.w3c.es/divulgacion/guiasbreves/websemantica>

Rdf se sitúa en la línea de los llamados microformatos. Como éstos, su objetivo es poder ofrecer información tanto para las personas como para los ordenadores, lo que se consigue al marcar semánticamente los elementos de manera que no haya ambigüedades. Por ejemplo, si utilizamos la palabra “Paris” en el contenido de la página html, podemos informar al ordenador (mediante el marcado semántico) si nos referimos a “Paris” la capital de Francia o si “Paris” se refiere a Paris Hilton, la persona.

<http://microformats.org>

5.2. Aplicaciones de la web semántica

La mayor limitación de los microformatos es su poca flexibilidad. Cada tipo de objeto sobre el que se quiera dar información requiere su propia plantilla que define cómo utilizar elementos html para ese caso particular. No se pueden reutilizar otros formatos previos y para cada nuevo caso se debe pasar por un proceso de consenso sobre el nuevo microformato. Además, no se puede mezclar más de un microformato para la misma descripción de un objeto.

Rdfa es la propuesta de la web semántica para metadatos en html. Se basa en una serie de tecnologías y propuestas que permiten solventar las limitaciones de los microformatos. Rdfa ha sido adoptado por los principales buscadores (*Yahoo!* y *Google*) como una forma de proporcionar resultados de búsqueda de mayor calidad (figura 3). Se espera que esto fomente la adopción de esta tecnología ya que se ha empezado a observar un mejor posicionamiento de sitios que incorporan rdfa (Brinker, 2009). Y *Facebook* también acaba de empezar a utilizarlo para integrar páginas externas dentro de su grafo social.

<http://developer.yahoo.com/searchmonkey/>

<http://knol.google.com/k/google-rich-snippets-tips-and-tricks>

<http://opengraphprotocol.org>

6. Conclusiones

Como arquitectos de la información, nuestro propósito principal es aprender sobre las principales necesidades de información de nuestros usuarios así como sobre sus conductas de búsqueda de información. Una mejor comprensión de lo que ellos realmente quieren de nuestro sitio nos ayudará a determinar y priorizar qué componentes de la arquitectura debemos construir, lo que hará nuestro trabajo mucho más sencillo, especialmente considerando de cuántas formas puede ser diseñada una IA.

También tendremos información sobre usuarios que nos ayudarán a contrarrestar los otros componentes que a menudo influyen el diseño como el presupuesto, el

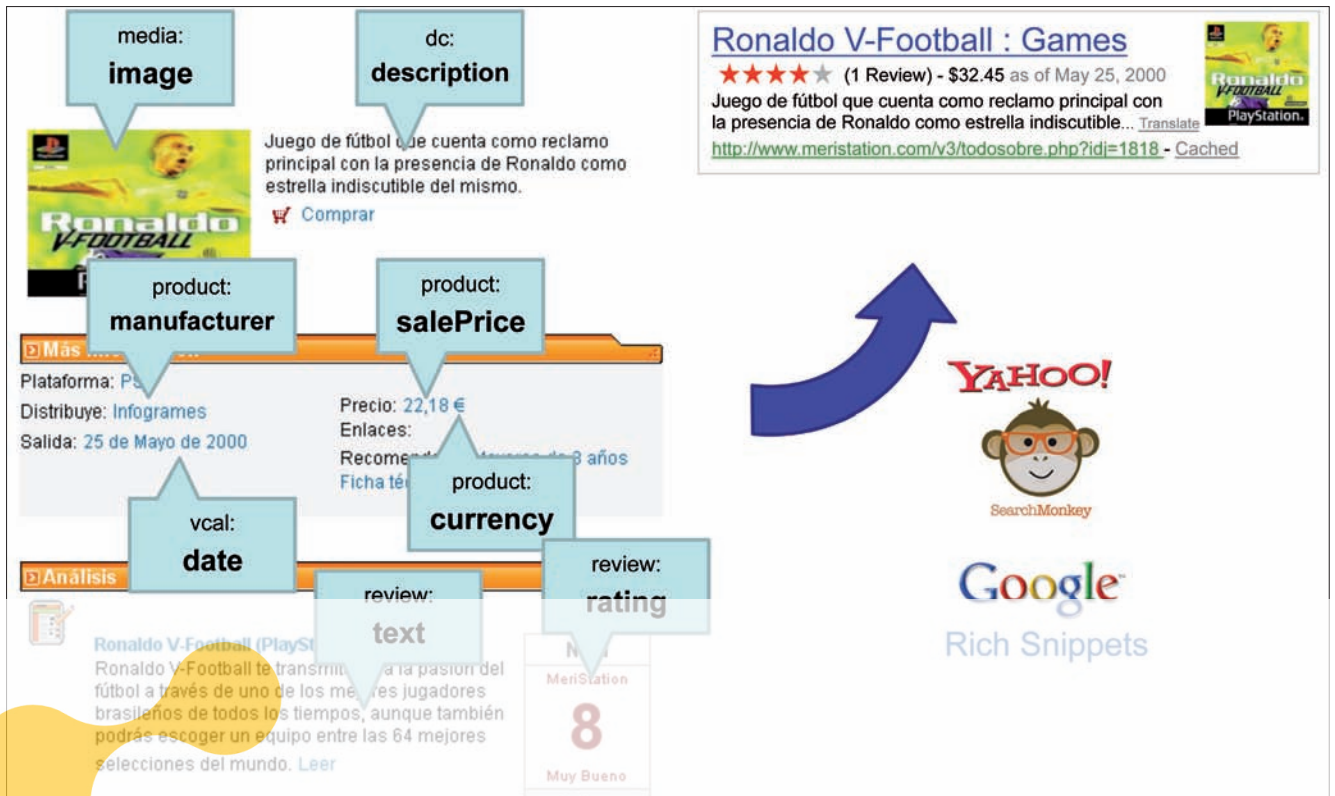


Figura 3. Resultado de búsqueda mejorado gracias a anotaciones semánticas no visibles de la página original basadas en ontologías de Yahoo! para productos, contenidos, u opiniones

SCIPEDIA

tiempo, las políticas, las tecnologías y las preferencias personales de los diseñadores.

Para construir una buena IA hoy en día ya no basta con realizar un buen análisis y un buen diseño, integrando un buscador interno y definiendo los metadatos de los contenidos, sino que debemos comenzar a ir un poco más allá contemplando tecnologías de web semántica, como la definición de una ontología o la inclusión de rdfs en nuestro sitio web. Esto puede facilitar la evolución y mantenimiento de nuestra IA, pero también mejorar el posicionamiento en buscadores como *Google* o *Yahoo!* o integrarla en el grafo social de *Facebook*.

7. Bibliografía citada

Bizer, Christian; Heath, Tom; Berners-Lee, Tim. "Linked data - The story so far". *International journal on semantic web and information systems*, 2009, v. 5, n. 3, pp. 1-22.

Brinker, Scott. "Best buy jump starts data web marketing". *Chief marketing technologist*, 11 Dec. 2009. <http://www.chiefmartec.com/2009/12/best-buy-jump-starts-data-web-marketing.html>

Granollers-Saltiveri, Toni; Lorés-Vidal, Jesús; Cañas-Delgado, José-Juan. *Diseño de sistemas interactivos centrados en el usuario*, Barcelona: UOC, 2005.

Miller, Craig S.; Remington, Roger W. "Modeling information navigation: implications for information architecture". *Human-computer interaction*, 2004, v. 19, n. 3, pp. 225-271.

Nielsen, Jacob. "Card sorting: how many users to test". *Jakob Nielsen's Alertbox*, 19 Jul. 2004. <http://www.useit.com/alertbox/20040719.html>

Pollock, Jeffrey T. *Semantic web for dummies*. Indianapolis, Indiana: Wiley, 2009. eISBN: 978-0-470-49818-7.

Ronda-León, Rodrigo. "Arquitectura de información: análisis histórico-cualitativo". *Metadatos*, 2008, n. 1. http://www.nosolousabilidad.com/articulos/historia_arquitectura_informacion.htm

Rosenfeld, Louis; Morville, Peter. *Information architecture for the World Wide Web*. Sebastopol: O'Reilly, 1998. ISBN: 978-1-56592-282-2.

Rosenfeld, Louis. *The road to finding is paved with data: web analytics and user experience*, 2008. <http://www.adobe.com/designcenter/thinktank/rosenfeld.html>

Wodtke, Christina; Govella, Austin. *Information architecture: blueprints for the Web, 2nd ed.* Berkeley: New Riders, 2009. ISBN 978-0-321-60080-6.

Wurman, Richard S. *Information architects*. Zurich: Graphis, 1997, ISBN 1888001380.

Roberto García. *Universitat de Lleida.*
rgarcia@diei.udl.cat

Federico Botella. *Universidad Miguel Hernández, Elche.*
federico@umh.es

Mari-Carmen Marcos. *Universitat Pompeu Fabra, Barcelona.*
mcarmen.marcos@upf.edu