APLICACIÓN DE LA AI PARA LA ESTRUCTURACIÓN DE CONTENIDOS MULTIMEDIA

SISTEMAS PARA LA ARQUITECTURA DE LA INFORMACIÓN

Son los responsables de la 'findability' (capacidad para ser encontrado) del producto de comunicación digital y de los contenidos que hay en él.

- Sistemas de organización
- Sistemas de etiquetado
- Sistemas de navegación
- Sistemas de búsqueda
- Vocabularios o lenguajes documentales

En el vídeo «Arquitectura de la información» descubrirás los aspectos clave de la Arquitectura de la información.

*Fuente: **Arquitectura de la información** by Rafael Pedraza Jiménez at https://www.youtube.com/watch?v=3MddqPRv-dw. License by owner of copyright.*

SISTEMAS DE ORGANIZACIÓN

Según la organización de la información en una web se logrará que los contenidos sean más fácilmente encontrados por los usuarios. Para lograr este objetivo es necesario un diseño adecuado del sistema de organización.

Esquemas de organización

Son aquellos que se encargan de organizar los grupos de ítems de información que contiene una página web en grupos, a partir de un críterio base. Estos pueden ser de dos tipos:

a. Exactos: Son aquellos bien definidos y que son mutuamente excluyentes y sirven para conocer cosas previamente conocidas:

- Alfabéticos
- Cronológicos
- Geográficos…

b. Ambiguos: Sirven para localizar contenidos que desconocemos. Por ejemplo:

- Tema

- Tarea
- Audiencia
- Metáfora (no hay que abusar de este sistema, pues si no está bien desarrollada su alta ambigüedad desorienta al usuario)
- Híbrido (utilización de varios de los esquemas anteriores)

Estructuras de organización

Su objetivo es organizar los grupos de ítems de información resultantes de los esquemas anteriores, mostrando las dependencias lógicas que existen entre ellos.

- **Estructura jerárquica:** Una buena arquitectura de información incluye una jerarquía o taxonomía entre sus componentes. Estos permiten al usuario ubicarse y obtener un modelo mental de la estructura del sitio que está visitando.

- **Basada en registros (o en el modelo de base de datos):** Este tipo de estructura es idónea para los contenidos bien estructurados y homogéneos.
- **Estructura hipertextual (o en red):** Debe ser complemento de otros tipos de estructuras y no la única opción de organización de la información. Permite reflejar relaciones menos estructuradas y más creativas existentes entre los contenidos.
- **Estructura secuencial:** Con un principio y un final claro y un único itinerario de consulta o exploración
- **Estructura en tabla (o matriz):** Únicamente para aquellos contenidos que admitan su representación en una tabla, y debe ser utilizado en combinación con otras estructuras.

SISTEMAS DE ETIQUETADO

Se debe utilizar un lenguaje claro y preciso que no lleve a confusión y que ayude a encontrar lo que nos interesa sin revelar ningún dato más.

Llegado el momento de definir el sistema de etiquetado, debemos tener

- La ambigüedad intrínseca al lenguaje, tanto sintáctica, léxicas (sinonimia y polisemia) como semántica (uso de metáforas, ironía, etc.)
- La arbitrariedad, evitando utilizar términos con un significado diferente con el que normalmente se asocian.
- La desorientación que producen las etiquetas que no anticipan ni dan pista sobre lo que esconden.
- Las etiquetas que utilizamos se asociación al branding de la empresa, y unas etiquetas mal elegidas pueden dar una mala impresión de la misma.

Las etiquetas pueden ser textuales o icónicas, estas deben utilizarse siempre junto a las primeras porque son intrínsecamente ambiguas.

Los sistemas de etiquetado con **formato textual** que deben planificarse son:

- **Enlaces contextuales.** Deben identificarse visualmente como tales, diferenciar los que sean externos y utilizarse de forma consistente. No deben ser ambiguos y se debe tener en cuenta el contexto en el que se insertan, siendo además independientes del mismo.
- **Títulos:** Deben estructurar de manera coherente el contenido y su diseño ha

de reflejar su jerarquía. Tienen que ser descriptivos y coherentes con los literales utilizados en los otros sistemas de etiquetado. Deben ser consistentes a lo largo del sitio tanto en su ubicación como en su aspecto.

- **Opciones del sistema de navegación:** Deben ser consistentes y coherente siguiendo un mismo patrón visual y de ubicación a lo largo de todo el sitio. Deben describir adecuadamente los contenidos que incluyen y ser entre ellos lo más excluyentes posibles.
- **Términos de indización:** Son el conjunto de etiquetas utilizadas para describir cualquier tipo de contenido en un entorno web y facilitar su localización, búsqueda y recuperación. Su enlace permite acceder al listado de todos los contenidos indizados con ese término y además alimentan el índice inverso del sistema de búsqueda. Cuando los propios usuarios o los autores de los contenidos proponen las etiquetas (tag) hablamos de marcadores sociales. Una folksonomía es el resultado de la agregación de las etiquetas o tags propuestos por uno más usuarios. Se pueden visualizar en forma de listado, nubes de etiquetas, los más populares, etc.

SISTEMAS DE NAVEGACIÓN

Son estructuras que ordenan y agrupan los contenidos de una página web bajo unas categorías que forman una clasificación. Nos permiten:

- Identificar las relaciones entre los contenidos de la web y entre esos contenidos y la página que se está visitando en ese momento.
- Habilitar y facilitar la navegación entre esos contenidos.
- Orientarnos, saber dónde estamos, que hay aquí, de dónde venimos y cómo podemos ir hacia donde deseamos ir.
- Formar una imagen mental del tamaño y estructura del sitio.

Google creó Google Sitemaps para que los desarrolladores web puedan publicar listas de enlaces de sus sitios, debido a que algunas webs tienen un número grande de páginas dinámicas que sólo están disponibles cuando los usuarios emplean los formularios correspondientes. En estos casos los archivos de Sitemaps pueden usarse para indicar a una araña web cómo deben ser encontradas las páginas. Actualmente Google, MSN, Yahoo y Ask.com soportan este protocolo conjuntamente, ahora llamado simplemente *Sitemaps*. Sitemaps suele usar XML, aunque también se admiten fuentes web RSS y archivos de texto.

Un ejemplo de ello serían los mapas de sitio. Un mapa de sitio web es una lista de las páginas de un sitio web accesibles por parte de los buscadores y los usuarios. Puede ser tanto un documento en cualquier formato usado como herramienta de planificación para el diseño de una web como una página que lista las páginas de una web (ya realizada), organizadas comúnmente de forma jerárquica. Esto ayuda a los visitantes y a los bots de los motores de búsqueda a hallar las páginas de un sitio web.

Los mapas de sitio pueden mejorar el posicionamiento en buscadores de un sitio, asegurándose que todas sus páginas puedan ser encontradas. Esto es especialmente importante si el sitio utiliza menús Adobe

Los usuarios pueden navegar de maneras muy distintas: con dirección, cuando

buscan una información determinada, sin dirección, cuando ojean la web sin un objetivo claramente preestablecido, además puede navegar en amplitud o en profundidad.

Tipos de sistemas de navegación

a. Compuestos por:

- Sistemas constantes o globales (permanecen de forma constante en el sitio)
- Sistemas locales (que deben estar correctamente articulados con el sistema constante),
- Sistemas contextuales (que permiten identificar contenidos relacionados y enlazar con ellos. Es importante que los contenidos o páginas críticas se enlaces desde otros contenidos).

b. Sistemas complementarios. Son recursos para localizar información y para orientarse que suelen ser páginas propias e independientes.

- Mapa del sitio.
- Índices, listado de términos que representan el contenido del sitio, normalmente ordenados alfabéticamente.
- Guías, para introducir a los nuevos usuarios en los contenidos y la funcionalidad de una parte concreta del sitio: wizards, configuradores, visitas guiadas, tutoriales, etc.

c. Sistemas no básicos.

- Sistemas de personalización, estructuras de navegación proactivas que se autodiseñan en función de lo que espera el usuario, ofreciéndole enlaces a partir de su perfil. El principal problema es que los comportamientos pasados no son garantía de inferir sus comportamientos futuros.
- Sistemas de customización: estructuras reactivas que permiten que el usuario pueda diseñar su propio sistema de navegación. El principal problema es que a los usuarios les pasen desapercibidas estas opciones, que si sus necesidades cambian deben redefinirlas o que simplemente no tengan tiempo de hacerlo.
- Sistemas de navegación visual, permiten explorar usando recursos icónicos o visuales.
- Sistemas de navegación social, inferidos a partir del comportamiento de la mayoría de usuarios que visitan la página (lo más popular, lo más comprado, etc.)

SISTEMAS DE BÚSQUEDA

Utilizados normalmente para localizar información a partir de una necesidad concreta. Ofrecen los resultados que coinciden con los definidos por el usuario en la ecuación de búsqueda. Los problemas a superar son:

- Ruido: contenidos recuperados no pertinentes, se mide con el índice de precisión.
- Silencio: documentos pertinentes no recuperados, se mide con el índice de exhaustividad.

Los sistemas de búsqueda pueden ser:

a. **Reactivos**: reaccionan frente a la conducta informativa del usuario.

b. **Proactivos**: ofrecen proactivamente la información al usuario sin que tenga que reclamarla continuamente:

- Sistemas de difusión selectiva de información (DSI), ofrecen una actualización informativa automatizada sobre un tema concreto a partir de la sindicación de contenidos.
- Sistemas de workshop, suministran automáticamente la información dentro de un proceso a partir del perfil del usuario.
- Agentes inteligentes: infieren el perfil de interés informacional de un usuario a partir de criterios como su histórico de comportamiento o su similitud con el de otros usuarios.

c. **Mixtos.** RSS, por ejemplo, es reactivo porque has de suscribirte y proactivo porque desde ese momento los recibes automáticamente.

Recomendaciones

- Evaluar la necesidad de un vocabulario controlado y definir los metadatos implicados en su descripción.
- Seleccionar como términos de indización ciertos componentes de los contenidos (título, url, etiquetas de los enlaces o títulos de las imágenes)
- Si se indiza el texto completo de todos los contenidos, introducir recursos como la búsqueda por campos para refinar el resultado y evitar una tasa de ruido elevada.
- Evaluar si se indiza solo algunos contenidos, algunas zonas de esos contenidos o algunos componentes determinados de ellos.
- Indizar los contenidos según el tipo de usuario o audiencia al que va dirigido.
- Inferir patrones en el comportamiento de los usuarios.
- Tratar de indizar las páginas según el tipo de contenido.
- Incluir siempre que se pueda:
    - Corrector ortográfico: corrigen automáticamente la ortografía dando alternativas.
    - Herramientas fonéticas: para identificar términos con grafías diferentes pero fonéticamente idénticas
    - Herramientas de procesamiento del lenguaje natural: obvian las palabras vacías e introducen AND entre los términos propuestos.
    - En función del proyecto, evaluar la inclusión de sugerencias de términos en la caja de búsqueda, que se va adaptando a lo que se escribe y que permiten identificar las búsquedas con más resultados.
- Poder refinar y mejorar manualmente los resultados de las búsquedas más habituales.
- Formar a los autores del contenido sobre cómo deben redactar los títulos o etiquetar el contenido.
- La caja de búsqueda debe estar identificada como tal y cerca del sistema de navegación.
- Diseñar la interfaz de búsqueda teniendo en cuenta a los usuarios, el tipo de necesidad informativa, la cantidad de información recuperada y el tipo de contenidos a recuperar.
- En la página de resultado:
    - indicar los términos por los que se ha buscado
    - indicar el número de resultados mostrados y encontrados
    - poner en negrita las coincidencias
    - permitir refinar la búsqueda (no solo semánticamente sino por fecha, lengua, tipo de documento, etc.)
    - permitir realizar búsquedas únicamente en los resultados

    - permitir personalizar el número de resultados a mostrar.
    - poder imprimir o enviar por email los resultados
    - poder guardar la búsqueda, para ello es necesario que sea mediante GET para que tenga una URL propia.
    - ofrecer opciones de ordenación (por precio, por relevancia, por cronología, etc.)

Lenguajes documentales

La **indización** es la operación en la que se asigna a cada contenido una serie de términos (palabras clave) que representan el tema o temas sobre los que versa. Es una operación de análisis. La clasificación es la operación por la cual se asigna a cada contenido un único término que representa el tema principal sobre el que versa ese documento. Es una operación de síntesis.

La información resultante de la indización y la clasificación suele incorporarse al contenido mediante **metadatos** asociados para que funcione con los sistemas de organización, etiquetado, navegación y búsqueda.

Un se constituye a partir de un subconjunto de términos del lenguaje natural (acompañados a veces de números como en el caso de las clasificaciones) para facilitar la búsqueda y recuperación de la información contenida en los documentos.

Un lenguaje documental está formado por:

El vocabulario del lenguaje documental: subconjunto del lenguaje natural. Se distingue:

- El término de indización principal (descriptor): es unívoco. Identifica la representación estándar de ese concepto. Por ejemplo, “sacerdote”.
- Término de indización secundario: es sinónimo del principal y una buena representación del concepto que el primero también representa, pero se decide no identificarlo como la representación estándar. Por ejemplo: “cura”, “clérigo”

Relaciones semánticas entre los términos de indización, que pueden ser de:

- Equivalencia: sinónimos, acrónimos, abreviaciones, variantes léxicas, posibles errores ortográficos, cuasisinónimos.
- Jerarquía, que puede ser:
- Genérica: soltero-hombre
- Relación parte-todo: rueda-coche
- Relaciones de instanciación: Mediterráneo-mar
- Asociativa: por afinidad semántica o evocación. Por ejemplo, veneno-toxicidad.

Tipos de lenguajes documentales

a. Libres, por ejemplo extraer automáticamente de un texto las palabras que no pertenecen al fichero de palabras vacías.

b. Controlados:

- Anillos de sinónimos, entre los términos se establecen relaciones de equivalencia pero no de jerarquía ni de asociación. Al hacer una búsqueda se recuperan también los indexados por sus sinónimos. Reduce el silencio en las búsquedas pero aumenta el ruido.
- Fichero de autoridades, listado de términos principales (con sus

respectivos sinónimos) para describir y normalizar un conjunto de entidades (personas, organizaciones, lugares geográficos) ante la variedad de homónimos, sinónimos o nombres con los que puede ser denominada una persona, entidad, obra, tema o concepto. Se utiliza especialmente en la catalogación de los fondos de las bibliotecas.
- Lista de encabezamientos de materia (LEM), en el cual el encabezado condensa el tema del documento con uno o varios términos. Presentan relaciones asociativas, jerárquicas y de equivalencia. Se usan preferentemente en las bibliotecas y en los centros de documentación cuyos fondos son esencialmente enciclopédicos.
- Taxonomías, sus términos presentan relaciones de equivalencia y jerarquía sin embargo no asociativas.
- Clasificación, representa entre sus términos relaciones asociativas, jerárquicas y de equivalencia y algunas veces asocia un código identificativo a cada uno de sus términos. La más empleada en web es la clasificación facetada para poder clasificar simultáneamente desde distintos puntos de vista, o a partir de diferentes criterios, un mismo conjunto de objetos. Cada una de las clasificaciones o facetas es paralela a las demás en un mismo nivel semántico. Cada faceta sería una etiqueta del sistema de navegación global, o un campo de la base de datos; y las categorías de esa faceta serían las de uno de los sistemas de navegación local, o las de los posibles valores de un campo de la base de datos. Por ejemplo, cada faceta seria (corriente literaria, género literario, etc.) y las categorías (realismo, modernismo, …; novela, poesía, …).
- Se encuentran integrados en los sistemas de navegación y búsqueda. Sus términos presentan (términos relacionados, TR), (términos genéricos, TG y términos específicos, TE) y de

SISTEMAS GUIADOS POR IA

La inteligencia artificial está revolucionando la estructuración de contenidos y la recuperación de información en sistemas web. Mediante la automatización de la clasificación, la generación de metadatos, la optimización de la navegación y la personalización del contenido.

Además, gracias a la implementación de búsquedas semánticas, las recomendaciones personalizadas, el análisis de sentimiento y la optimización de SEO, la IA hace que la recuperación de información sea más precisa, relevante y eficiente, transformando la manera en que los usuarios interactúan con los contenidos en línea.

Algunos ejemplos de su utilización en la **estructuración de contenidos** son:

- **Clasificación automática y generación de taxonomías.** La inteligencia artificial (IA) permite la clasificación automática de grandes volúmenes de contenido en categorías mediante algoritmos de aprendizaje automático, lo que mejora la organización y accesibilidad del contenido al identificar patrones y asignar etiquetas.
- **Generación de metadatos.** Los sistemas de IA generan metadatos de forma automática, proporcionando descripciones, palabras clave y etiquetas que facilitan la indexación y recuperación del contenido.
- **Optimización de la navegación.** La IA puede optimizar la estructura de navegación de los sitios web al analizar el comportamiento del usuario. Esta optimización ajusta menús y enlaces para proporcionar una experiencia

de usuario más intuitiva y eficiente.
- **Personalización del contenido.** Los algoritmos de IA personalizan la presentación del contenido según las preferencias y comportamientos previos de los usuarios, lo que asegura que cada usuario tenga una experiencia adaptada a sus necesidades.

En cuanto a la **recuperación del contenido:**

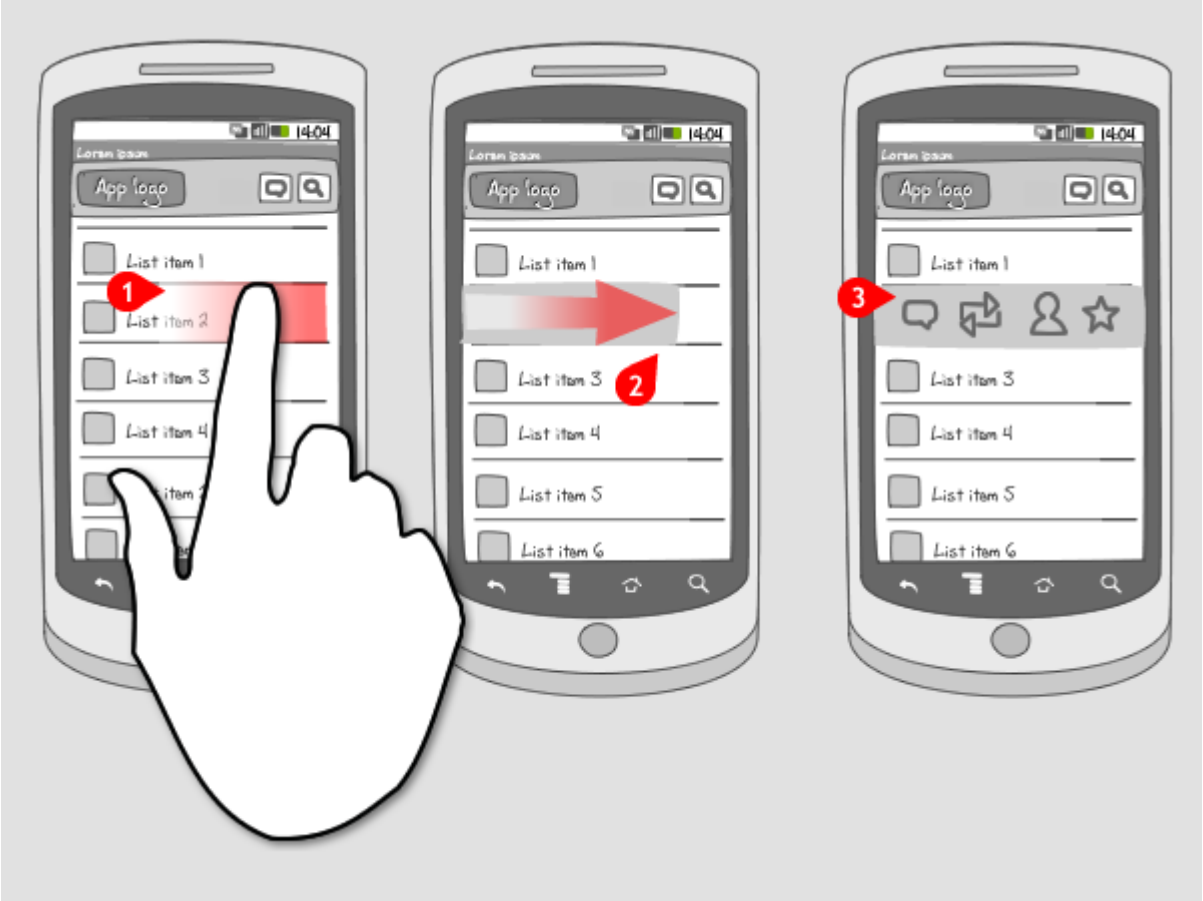
- **Búsqueda semántica.** La IA mejora las capacidades de búsqueda mediante la implementación de búsquedas semánticas que entienden el contexto y la intención detrás de las consultas de los usuarios, lo que facilita la localización de la información.
- **Recomendaciones personalizadas.** Mediante técnicas de aprendizaje automático, la IA analiza el comportamiento y las preferencias de los usuarios para ofrecer recomendaciones personalizadas de contenido. Esto no solo mejora la experiencia del usuario, sino que también incrementa la retención y el tiempo de permanencia en el sitio web.
- **Análisis de sentimientos y opiniones.** La IA analiza los comentarios y opiniones de los usuarios en tiempo real, proporcionando información útil sobre sus percepciones y sentimientos hacia ciertos contenidos. Ello permite ajustar y mejorar continuamente los contenidos ofrecidos en función del feedback recibido.
- **Optimización de SEO.** La IA ayuda a optimizar el contenido para motores de búsqueda mediante la identificación de tendencias y palabras clave relevantes, así como la mejora de la estructura del contenido para cumplir con las mejores prácticas de SEO.

PATRONES DE DISEÑO

Los patrones de diseño son “**soluciones a un problema que se usa repetidamente** en contextos similares con algunas variantes de implementación” (Christopher Alenxander).

Los patrones de diseño responden a las siguientes características:

- Deben ser concretos, es decir, que deben servir de forma específica a resolver una cuestión o problema.
- Multiplataforma. Deben ser capaces de ser aplicables y utilizables a través de distintas plataformas.
- Se concretan en productos.
- Se aplican como sugerencias, no como requerimientos.
- Deben adaptarse al contexto de diseño.



*Fuente: **Android Patterns** by UNITiD at https://unitid.nl/androidpatterns/wp-content/themes/AndroidPatterns2/ads/banner1.html. License by owner of copyright.*

Los patrones de diseño especifican qué hacer para conseguir un objetivo concreto, en un contexto determinado. Habitualmente detallan cuándo, cómo y por qué puede aplicarse el patrón mediante explicaciones y ejemplos.

Utilidad de los patrones

Los patrones de diseño son especialmente útiles en el desarrollo de aplicaciones y productos digitales de comunicación, ya que permiten ahorrar tiempo de producción en el diseño y planteamiento, aprovechando el conocimiento previo aportado por soluciones exitosas. A través de estos, podemos aprender de los ejemplos, comparar alternativas de diseño, servir de inspiración, etc.

Ampliación de información sobre los Patrones de diseño

Patterns Revisited from **Dave Malouf**

*Fuente: **Patterns Revisited** by Dave Malouf at https://www.slideshare.net/dmalouf/patterns-revisited. License by owner of copyright.*

Grupo Ciberimaginario | Manuel Gertrudix - Alejandro Carbonell | 2025/2026 | Esta obra está bajo una Licencia Creative Commons Atribución 4.0 Internacional. Los contenidos citados se ajustan a lo regulado en el art. 32 del TRLPI de España

(cc) BY