# Aplicación de lenguajes de marcas a la sindicación de contenidos

Sitio: <u>Aula Virtual do IES de Teis</u> Impreso por: Deivid Durán Durán

Curso: Linguaxes de Marcas e Sistemas de Xestión de Información 2021-22 (DAM-A)

Data: Venres, 24 de Xuño de 2022, 00:12

Libro: Aplicación de lenguajes de marcas a la sindicación de contenidos

# Táboa de contidos

- 1. ¿En qué consiste la sindicación de contenidos?
- 2. ¿Cómo funciona la sindicación de contenidos web?
- 3. Ejemplos de páginas que permiten sindicación de contenido
- 4. Formato RSS
- 5. Formato ATOM
- 6. Validación de feeds
- 7. ¿Cómo se integran los canales de RSS en los sitios web?
- 8. Ventajas de la sindicación de contenidos

# 1. ¿En qué consiste la sindicación de contenidos?

¿En qué consiste la sindicación de contenidos?

Sindicación es una traducción literal del inglés, *syndication*, que en significa **redifusión de contenidos** principalemente periodísticos. En el ámbito de la infomática se relaciona principalmente con redifusión de contenidos **web**, principalmente noticias y blogs.

La sindicación de contenidos web permite a los sitios web distribuir contenidos a usuarios interesados mediante una **subscripción**. Los contenidos pueden ser propios o de terceros. En este último caso, la redifusión de contenidos web suele realizarse bajo una licencia de normas de uso, o mediante un contrato que regule los derechos de los contenidos.

Popularmente, se ha extendido el término RSS (*Really Simple Syndication*) para hacer referencia a la sindicación de contenidos web en términos generales, pero como veremos, RSS es un posible formato de sindicación de contenidos basado en XML.

El icono que se asocia a este tipo de sindicación es habitualmente similar a este:



# 2. ¿Cómo funciona la sindicación de contenidos web?

Las fuentes de los contenidos se guardan en un fichero en formatos **basados en XML**. Estos ficheros indican dónde se encuentran los contenidos y pueden contener datos adicionales sobre categorías, autores, descripciones, etc.

Para hacer uso de esas fuentes en formato XML se necesitan programas específicos, lectores RSS o también llamados agregadores de *feeds*. Uno que gozó de gran popularidad fue <u>Google Reader</u>, hasta que dejó de ofrecer sus servicios en 2013.

Actualmente es posible encontrar otras opciones de agregadores o lectores de RSS:

- Aplicaciones web: Feedly e Inoreader,
- Navegadores con extensiones para leer RSS: FeedBro, extensión para el navegador Firefox, RSS Feed Reader para Chrome
- Clientes de correo: Outlook o Thunderbird
- Programas de escritorio: Winds

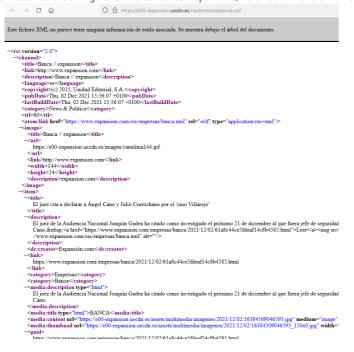
En cualquiera de las opciones, el usuario ha de suscribirse a una página o fuente sobre la que desea recibir las actualizaciones. De esta forma, tendrá acceso **desde una única aplicación** a los contenidos ahorrando así tiempo en visitar todas las webs de forma individualizada.

Los lectores RSS son capaces de interpretar los ficheros basados en XML recibidos mostrando su contenido de una forma más visual y comprensible por los humanos.

3. Ejemplos de páginas que permiten sindicación de contenido

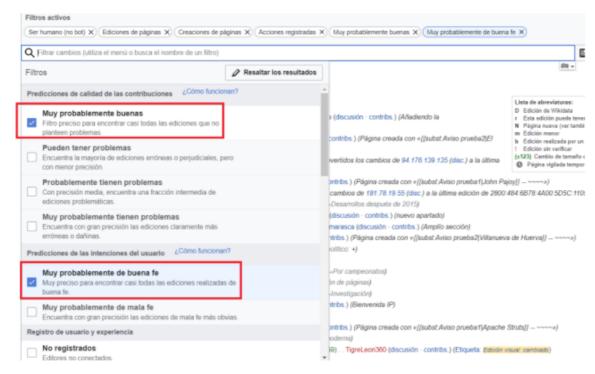
#### • Diarios digitales como Expansión, especializado en Economía: https://www.expansion.com/rss/

Si hacemos clic en alguna de las temáticas disponibles, veremos un fichero XML similar al siguiente:



#### • Wikipedia:

1. Visita <a href="https://es.wikipedia.org/wiki/Wikipedia:Portada">https://es.wikipedia.org/wiki/Wikipedia:Portada</a> . A continuación, haz clic sobre el enlace "Cambios recientes" para ver las páginas de la Wikipedia editadas recientemente. Seguidamente, en filtros activos añade los filtros: "Muy probablemente buenas" y "Muy probablemente de buena fe" para evitar ediciones probablemente irrelevantes



1. Haz clic sobre el enlace de Atom, en la sección de la izquierda



Sobre la página resultado, veremos un documento en formato XML.

Para ver su código fuente, podemos pulsar la combinación de teclado CTRL+U y veremos algo similar a lo siguiente:

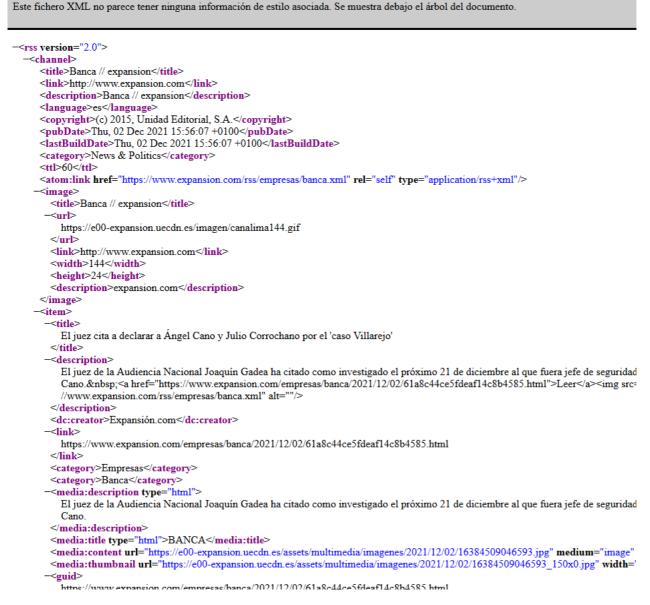
```
\rightarrow C \widehat{\omega}
\leftarrow
                                                                                                  🛆 view-source:https://es.wikipedia.org/w/api.php?hidebots=1&hidecategorization=1&hideWikibase=1&urlversion=1&days=7&limit=50&action
           <feed xmlns="http://www.w3.org/2005/Atom" xml:lang="es">
                        <id>https://es.wikipedia.org/w/api.php?hidebots=16amp;hidecategorization=16amp;hideWikibase=16amp;urlversion=16amp;days=76amp;limit=5066
                        <title>Wikipedia - Cambios recientes [es]</title>
<title>Wikipedia - Cambios recientes [es]</title>
<tink rel="self" type="application/atom+xml" href="https://es.wikipedia.org/w/api.php?hidebots=1&amp;hidecategorization=1&amp;hideWikibecategorization=1&amp;hideWikibecategorization=1&amp;hideWikibecategorization=1&amp;hideWikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikibecategorization=1&amp;hidewikib
                        <link rel="alternate" type="text/html" href="https://es.wikipedia.org/wiki/Especial:CambiosRecientes"/>
                        <updated>2021-12-03T12:01:31Z</updated>
                        <subtitle>Realiza un seguimiento de los cambios más recientes en la wiki en este canal.</subtitle><generator>MediaWiki 1.38.0-wmf.9</generator>
<link rel="alternate" type="text/html" href="https://es.wikipedia.org/wiki/Usuario:Martamarcosalvarez"/>
                                    <summary type="html">&lt;p&gt;Se ha creado la cuenta de usuario &lt;a href=&quot;/w/index.php?title=Usuario:Martamarcosalvarez&amp;
                                   <id>https://es.wikipedia.org/w/index.php?title=Aire&amp;diff=140116969&amp;oldid=140102731</id>
                                    <link rel="alternate" type="text/html" href="https://es.wikipedia.org/w/index.php?title=Aire&amp;diff=140116969&amp;oldid=140102731</pre>
                                                           <td colspan=&quot;2&quot; style=&quot;background-color: #fff; color: #202122; text-align: center;&quot;&gt;- Revisión and
&lt;td colspan=&quot;2&quot; style=&quot;background-color: #fff; color: #202122; text-align: center;&quot;&gt;Revisión del :
                                                            </tr&gt;&lt;tr&gt;
                  <td colspan=&quot;2&quot; class=&quot;diff-lineno&quot;&gt;Linea 1:&lt;/td&gt;
                 <td colspan=&quot;2&quot; class=&quot;diff-lineno&quot;&gt;Linea 1:&lt;/td&gt;
   38 &lt:/tr&at:
   39 &lt:tr&at
                 <td class=&quot;diff-marker&quot; data-marker=&quot;-&quot;&gt;&lt;/td&gt;
                  61t;td style=6quot;color: #202122; font-size: 88%; border-style: solid; border-width: 1px 1px 1px 4px; border-radius: 0.33em; border-color
                 <td class=&quot;diff-marker&quot; data-marker=&quot;+&quot;&gt;&lt;/td&gt;
                 61t;td style=6quot;color: #202122; font-size: 88%; border-style: solid; border-width: 1px 1px 1px 4px; border-radius: 0.33em; border-color
   44 </tr&gt
   45 <tr&gt
                <td class=&quot;diff-marker&quot;&gt;&lt;/td&gt;
                  61t;td style=6quot;background-color: #f8f9fa; color: #202122; font-size: 88%; border-style: solid; border-width: 1px 1px 1px 4px; border-
                  <td class=&quot;diff-marker&quot;&gt;&lt;/td&gt;
                 <td style=&quot;background-color: #f8f9fa; color: #202122; font-size: 88%; border-style: solid; border-width: 1px 1px 1px 4px; border-
   50 </tr&gt
   51 <tr&gt
                <td class=&quot;diff-marker&quot;&gt;&lt;/td&gt;
                  61t;td style=6quot;background-color: #f8f9fa; color: #202122; font-size: 88%; border-style: solid; border-width: 1px 1px 1px 4px; border-style: solid; border-width: 1px 4px; border-style: solid; border-width: 1px 4px; border-style: solid; border-width: 1px 4px; border-style: solid; border-style: solid; border-width: 1px 4px; border-style: solid; border-width: 1px 4px; border-style: solid; bord
                  <td class=&quot;diff-marker&quot;&gt;&lt;/td&gt;
                 61t;td style=6quot;background-color: #f8f9fa; color: #202122; font-size: 88%; border-style: solid; border-width: 1px 1px 1px 4px; border-style: solid; border-width: 1px 4px; border-style: solid; 
  57 </table&gt;</summary>
58 <author><name>158.99.4.163</name></author>
                        </entry>
```

### 4. Formato RSS

- RSS (Really Simple Syndication) ha tenido varias versiones desde su nacimiento en 1999 en el seno de la empresa Netscape.
- La versión más reciente es RSS 2.0
- Los archivos RSS 2.0 deben ser conformes con la especificación de XML 1.0
- Tienen extensión .rss o .xml

→ C m

- El elemento raíz es rss y debe tener un atributo obligatorio version
- El elemento raíz tiene un único subelemento channel y, este, a su vez debe tener los siguientes subelementos:
  - o title: Define el título del canal
  - o link: URL de la página que oferta el canal.
  - o description: Breve descripción del sitio que ofrece el canal RSS



https://e00-expansion.uecdn.es/rss/empresasbanca.xml

Un canal puede contener cualquier número de elementos <item>.

ELEMENTO ITEM.- También puede subelementos. Los más habituales son:

- title: Si está presente, define el título o el enlace hacia la noticia completa
- link: Si está presente, enlace para visualizar la noticia completa referida en title
- description: Si está presente, podría tener todo el contenido del artículo/reportaje/story/noticia o una breve sinopsis del contenido

Un *item* puede representar un artículo o reportaje (o *story*), al igual que en un periódico o revista; si es así, su descripción es una sinopsis de la historia y el enlace apunta a la historia completa.

Un *item* también puede estar completo en sí mismo, si es así, la descripción contiene el texto (se permite HTML codificado con entidades, sustituyendo < por &lt;, & por &lt;, & por &lt;, etc.) y el enlace y el título pueden omitirse. Todos los elementos de un *item* son opcionales, sin embargo, debe estar presente al menos o *title* o *description*.

Por lo tanto, un item podría tener:
•Solo description: Todo el contenido estaría aquí incluido

•Solo title

Se pueden consultar los diferentes elementos, su opcionalidad u obligatoriedad en la <u>especificación de RSS 2.0 del W3C</u>

En ese recurso, podéis consultar ficheros de ejemplo para diferentes versiones en Sample files.

### 5. Formato ATOM

```
| demi version="10"72 | demi-sourcehtps://es.wikipedia.org/w/paj.hpp?hidebots=1&hidecategorization=1&hidevikibase=1&uriser=1&uriser=1** | demi version="10"72 | demi version="10
```

La página resultado de visitar los cambios recientes de la wikipedia es un ejemplo de un fichero con formato ATOM para la sindicación de contenidos:

- ATOM es un estándar propuesto por el grupo de trabajo Atom Publishing Format and Protocol (Formato y protocolo de publicación Atom) de la IETE
- Un documento que use el formato ATOM deben ser documento XML bien formado
- Pueden tener extensiones .atom o .xml
- Debe tener un elemento raíz feed
- Todos los elementos del documento deben pertenecer al espacio de nombres http://www.w3.org/2005/Atom
- Cada elemento feed debe tener 3 elementos obligatorios:
  - title: Contiene un texto legible por humanos que suele coincidir con el título de la página web origen. No puede estar en blanco.
     id: URI que identifica el recurso y en el que probablemente se encontrará el recurso (URL)

**updated**: Indica la fecha y hora en la que el feed se ha modificado de forma significativa. Debe seguir un formato específico recogido en un documento técnico RFC 3339, basado en ISO8601.

• Un feed puede tener cero o varios elementos entry

Consulta la Especificación de ATOM del W3C para averiguar cuáles son los elementos obligatorios de entry

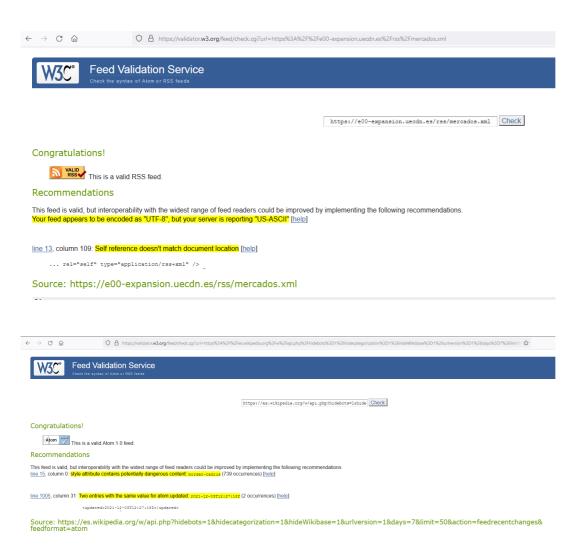
### 6. Validación de feeds

El W3C tiene un servicio de validación online de feeds de sindicación de contenidos, tanto para ATOM, como para RSS.

La validación se puede realizar a través de las URL o a través del propio documento XML.

Valida con esa herramienta las URLs generadas en los 2 ejemplos que hemos visto de la Wikipedia y de Expansión.

- https://e00-expansion.uecdn.es/rss/empresasauto-industria.xml
- https://e00-expansion.uecdn.es/rss/mercados.xml
- https://es.wikipedia.org/w/api.php?hidebots=1&hidecategorization=1&hideWikibase=1&urlversion=1&days=7&limit=50&action=feedrecentchanges&feedformat=atom



### 7. ¿Cómo se integran los canales de RSS en los sitios web?

Los canales RSS tienen que ser accesibles a través del propio sitio web.

</html>

El acceso a los ficheros de sindicación de contenidos (RSS o ATOM) se realiza a través del código HTML de la página web, que veremos en la próxima unidad. Para ello pueden darse varias posibilidades:

• En el documento HTML, se puede añadir un enlace al fichero XML, normalmente acompañado del icono de sindicación de contenidos.

• Se vincula en el documento HTML en el apartado <head>. Por ejemplo, un documento HTML básico es el siguiente:



La mayoría de CMS (Content Management System) o Sistema de Gestión de Contenidos, que permiten crear y publicar contenidos web sin tener conocimientosde HTML o CSS, como pueden ser Wordpress o Joomla tienen extensiones o plugins para generar ficheros RSS a medida que se publican nuevos contenidos.

Una vez que el fichero de sindicación de contenidos (RSS o ATOM) es accesible a través de una URL podía registrarse, además, en un **directorio de feeds:** un sitio web que ofrecía, ordenados por categorías, enlaces a diferentes ficheros de sindicación de contenidos. En su momento, favorecía el descubrimiento y consumo de usuarios interesados en una misma temática y aportaba mayor visibilidad y tráfico a la web generadora de contenidos. <u>Feedburner</u> o <u>RSS Network</u> son ejemplos de directorios de feeds. Feedburner ha anunciado que deja de prestar algunos de sus servicios y entró en un período de mantenimiento desde julio de 2021.

Hoy en día, el consumo de RSS se ha visto desplazado en su mayoría por el uso de las redes sociales (Twitter, Facebook, Instagram) o sitios de contenido audiovisual como YouTube que permiten subscripción directa.

# 8. Ventajas de la sindicación de contenidos

### Ventajas de la sindicación de contenidos

#### Como productores de contenido:

- Aumentar el tráfico del sitio web
- Favorece el posicionamiento del sitio en buscadores aumentando el número de referencias desde otras páginas.

#### Como consumidores de contenido:

- Las páginas web y blogs distribuyen a través de los canales RSS las últimas actualizaciones
- La decisión está del lado del usuario ya que él es quien elige a qué páginas web suscribirse y cuando darse de baja de estas páginas web
- Evita la generación de SPAM y, en caso de querer cancelar la suscripción es sencilla, el proceso es muy sencillo.

Lectura recomendada: Volver a RSS? URL