

Resumen tema 3: Sindicación de contenidos

Lenguajes de marcas y Sistemas de Gestión de la Información

29 de noviembre de 2022

CIFP Carlos III - Cartagena

Santiago Francisco San Pablo Raposo

2º curso DAW

Contenido

# Índice de ilustraciones.

# Índice de tablas.

Resumen tema 3.

# 1.- Sindicación de contenidos.

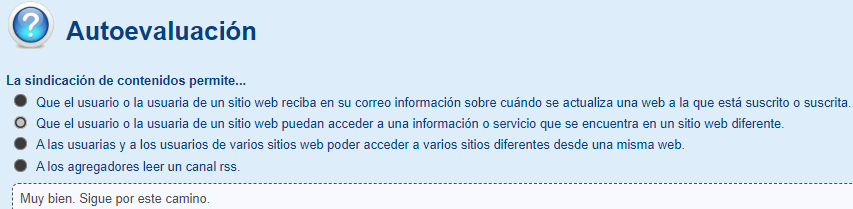
**La sindicación de contenidos permite a un sitio utilizar los servicios o contenidos ofertados por otra web diferente**. Esos servicios junto con los **metadatos** que tiene asociados en el sitio original, forman los **feed** o **canales de contenidos**.

**La redifusión de contenidos web suele realizarse bajo una licencia de normas de uso** en lugar de mediar un contrato para regular los derechos de los contenidos.

**En la actualidad** la redifusión web **consiste en ofrecer un contenido desde una fuente web**, cuyo origen está en una página web, **para proporcionar a los usuarios la actualización del mismo**.

Suele codificarse en XML, aunque es válido hacerlo en cualquier lenguaje que se pueda transportar mediante el protocolo HTTP.

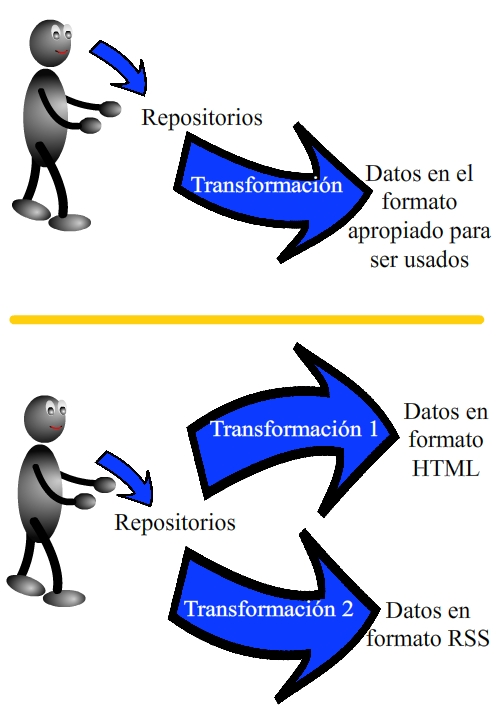
Para leer una fuente, o canal, hay que suscribirse a ella utilizando un agregador.



## 1.1.- Características.

**Para que una web sea suministradora de un canal**, en su cabecera hay que incluir, por debajo del elemento <title>, **un enlace al canal de contenidos**.

|  |  |
| --- | --- |
| RSS | <link rel="alternate" type="application/rss+xml" title="tituloEnlace" **href="http://www.misitio.com/fichero.rss"** /> |
| Atom | <link rel="alternate" type="application/atom+xml" title="tituloEnlace" **href="http://www.misitio.com/fichero.atom"** /> |

Actualmente, es **habitual el uso de algún** [**CMS**](https://rockcontent.com/es/blog/cms/) (Sistema de gestión de contenidos o Content Management System). Si es así, **los contenidos se encuentran en un repositorio** (o varios), y antes de ser servidor al cliente en el formato adecuado, **sufren algún tipo de transformación**. La figura de la izquierda muestra la estructura del flujo de la información.

Forma, Flecha

Descripción generada automáticamenteAl utilizar un CMS, la transformación puede replicarse. Además de tener más de una entrada de información, podríamos tener varias salidas. **Por ejemplo**: podemos generar tantos ficheros HTML como canales RSS.

**Para saber más**: [¿Qué es un CMS?](https://rockcontent.com/es/blog/cms/)

Texto

Descripción generada automáticamente con confianza baja

## 1.2.- Ventajas de la redifusión de contenidos.

**¿Por qué deberíamos utilizar los canales de contenidos de otros propietarios?**

* **Aumentar** el **tráfico** de nuestro sitio web.
* Incita a los usuarios a visitar más frecuentemente el sitio web.
* Favorece el posicionamiento del sitio en buscadores.
* Ayuda a **establecer relaciones entre distintos sitios web** dentro de la comunidad.
* **Permite a otras personas añadir características a los servicios del sitio web** (por ejemplo, **notificaciones de actualizaciones** mediante mensajes instantáneos), aunque se requiera de tecnologías adicionales.
* **Enriquece Internet** impulsando la tecnología semántica y fomentando la reutilización.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

# 2.- Ámbitos de aplicación.

Se puede sindicar cualquier tipo de información, incluso vídeos de YouTube.

* Weblogs: ayudaron mucho a la popularización de la redifusión web.
* Sindicación de todo tipo de información en formato XML.
* **Podemos ofrecer contenidos propios para que sean mostrados en otras páginas web de forma integrada**, lo que aumenta el valor de la página que muestra el contenido y también nos genera más valor, ya que normalmente **la redifusión web siempre enlaza con los contenidos originales**.
* **Puede aplicarse a todo tipo de contenidos**: texto, audio, vídeos e imágenes.
* De esta forma, los suscriptores pueden de esta forma mantener una actualización profesional de las fuentes de contenidos a los que está suscrito. **Así, el usuario puede estar al día en**:
  + temas relacionados con su profesión.
  + recibiendo las noticias e informaciones en su blog
  + o en su programa agregador de noticias.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

# 3.- Tecnologías de creación de canales de contenidos.

Para sindicar contenidos hay que utilizar alguno de los estándares de sindicación, los cuales están basados en XML, y se agrupan en dos estándares:

|  |  |
| --- | --- |
| Estándar | Características y versiones |
| RSS | Really Simple Syndication. |
| Además **se utiliza en la conexión con sistemas de mensajería instantánea**, la **conversión** de **RSS en mensajes de correo electrónico**, o la capacidad de **transformar los favoritos del navegador en RSS**. |
| Ha sido desarrollado por tres organizaciones, lo que ha dado lugar a 7 formatos diferentes entre sí:   * **RSS 0.90**: creada por la empresa Netscape en 1999. Se basa en la especificación RDF de metadatos. * **RSS 0.91**: versión simplificada de RSS 0.90. Se detuvo su desarrollo por falta de éxito, aunque la empresa UserLand Software decidió usar esta versión para desarrollar blogs. * **RSS 1.0**: creado a partir del RSS 0.90. Es más estable. Permite definir una cantidad mayor de datos que en versiones anteriores de RSS. * **RSS 2.0**: UserLand Software rechazó el estándar RSS 1.0 por considerarlo complejo y continuó el desarrollo del formato RSS 0.91, publicando las versiones 0.92, 0.93 y 0.94. Hasta que se publicó RSS 2.0 para que cumpliera todas las normas de XML |
| Atom | Estándar del grupo de trabajo *Atom Publishing Format and Protocol* (Formato y protocolo de publicación Atom) de la IETF en el RFC4287. |
| Se desarrolló como una **alternativa a RSS, con el fin de evitar la confusión creada por la existencia de estándares** similares para la sindicación de contenidos, entre los que existía cierta incompatibilidad.  No sustituyó a dichos estándares, pero sí que convivió con ellos. |
| Se caracteriza por su flexibilidad. |

**Debes conocer**: las especificaciones de estos estándares:

* [RDF Site Summary (RSS) 1.0 (resource.org)](https://web.resource.org/rss/1.0/)
* [Site Outline (RSS 2.0 at Harvard Law)](https://cyber.harvard.edu/rss/)
* [RFC 4287: The Atom Syndication Format (rfc-editor.org)](https://www.rfc-editor.org/rfc/rfc4287)

# 4.- Estructura de los canales de contenidos.

**Caso práctico**: El primer paso es generar un fichero en RSS que ha de verificar las normas XML. Además de definir en cuál estándar se va a trabajar, hay que definir un canal donde se establecerá el sitio web asociado al mismo. En dicho canal, se definirán tantos ítems como sitios web se pretendan sindicar.

Para construir un canal de contenido, es necesario crear un fichero (RSS o Atom), basado en XML. **Estará formado por los siguientes elementos básicos**:

* **Declaración del documento xml y la definición de la codificación empleada** en el documento.
  + Codificación preferentemente en **UTF-8**.
* **Un canal** en el que **se determina el sitio web asociado a la fuente web** a la que hace referencia el fichero. Éste, además de su propia definición, estará formado por:
  + **Secciones**: cada una referencia a la web que contiene uno de los servicios que se van a ofrecer.
    - **En un canal pueden incluirse tantas secciones como se quiera**, lo que hace que un canal de contenido pueda tener un tamaño enorme si contiene un gran número de enlaces independientes.
  + **No existe restricción de cantidad de canales de contenidos** que se pueden ofrecer.

## 4.1.- RSS.

Ejemplo de fichero RSS:

|  |  |
| --- | --- |
| Elementos imprescindibles | Código |
| Declaración de XML | <?xml version="1.0" encoding="UTF-8"?>  <rss version="2.0"> |
| Creación de un canal e ítems | <channel>  <title></title> <!--Imprescindible -->  <link></link> <!--Imprescindible -->  <description> <!--Imprescindible -->    </description>  <language>es-es</language> <!--Imprescindible -->  <item> <!--Imprescindible -->  <title></title> <!--Imprescindible -->  <guid></guid> <!--Imprescindible -->  <description> <!--Imprescindible -->    </description>  <pubDate>Sun, 17 Nov 2022 18:45:00 GMT</pubDate>  </item>  </channel> |
| Cierre de RSS | </rss> |

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

## 4.1.- Atom.

|  |  |
| --- | --- |
| Elementos imprescindibles | Código |
| Declaración de XML | <?xml version="1.0" encoding="UTF-8"?>  <rss version="2.0"> |
| Creación de un canal e ítems | <channel>  <title></title> <!--Imprescindible -->  <link></link> <!--Imprescindible -->  <description> <!--Imprescindible -->    </description>  <language>es-es</language> <!--Imprescindible -->  <item> <!--Imprescindible -->  <title></title> <!--Imprescindible -->  <guid></guid> <!--Imprescindible -->  <description> <!--Imprescindible -->    </description>  <pubDate>Sun, 17 Nov 2022 18:45:00 GMT</pubDate>  </item>  </channel> |
| Cierre de RSS | </rss> |

<?xml version="1.0" encoding="utf-8"?>

<!--Se crea el canal, se determina el estándar a utilizar y el idioma del documento -->

<feed xmlns="http://www.w3.org/2005/Atom" xml:lang="es-es"> <!-- Canal -->

<title type="text">Ejemplo de Atom</title>

<updated>2011-01-14T19:17:46Z</updated> <!-- Fecha de actualizacion del canal -->

<link rel="self" type="application/atom+xml" href="http://www.infoalisal.com/feed/atom" /> <!-- Hace referencia a sí mismo -->

<link rel="alternate" type="text/html" href="http://www.infoalisal.com" /> <!-- Hace referencia a la web -->

<id>http://www.infoalisal.com/feed/atom</id> <!-- Hace referencia a la fuente web -->

<!-- Se definen las secciones (entradas que forman el canal -->

<entry>

<title></title>

<link rel="alternate" type="text/html" href="http://www.infoalisal.com/seccion01" />

<updated>2011-01-14T19:19:46Z</updated>

<id>http://www.infoalisal.com/seccion01</id>

<author>

<name></name>

<email></email>

</author>

<summary></summary>

</entry>

<entry>

<title></title>

<link rel="alternate" type="text/html" href="http://www.infoalisal.com/seccion01" />

<updated>2010-12-14T19:17:55Z</updated>

<id>http://www.infoalisal.com/seccion02</id>

<author>

<name></name>

</author>

<summary></summary>

</entry>

</feed>

**En la etiqueta <updated>** (fecha y hora de actualización), se utiliza un formato muy concreto:

**CCYY-MM-DDTHH:MM:SSZ**, donde:

* T es el separador entre la fecha y la hora
* Z indica que la hora hace referencia al sistema de tiempo universal, esto es la hora zulú, o la hora del meridiano de Greenwich.

**Etiquetas <link> del canal**: Definen el canal. Son necesarias dos:

* Uno al propio fichero .atom, cuyo valor del atributo rel del elemento link será "self".
* Otro al fichero web que oferta ese canal, en este caso rel="alternate".

**Etiqueta <link> de un entry**: de tipo rel=”alternate”, debe enlazar a la fuente de la sección.

**Para saber más**: [15.2. Canales Atom (El tutorial Jobeet) (uniwebsidad.com)](https://uniwebsidad.com/libros/jobeet-1-4/capitulo-15/canales-atom-2)

# 5.- Validación.

Basta con entrar a una de los siguientes validadores y especificar la URL, copiar y pegar el código, o en su defecto, subir el fichero XML para que nos lo validen automáticamente estos servicios online.

**Una vez validado, suelen ofrecer una imagen del tipo "XML" o "RSS", de color naranja por lo general** (que indican que el canal es válido)**, que se puede incluir en la página principal, para enlazar a la dirección del fichero alojado en su dominio**.

* [W3C Feed Validation Service mediante URI](https://validator.w3.org/feed/" \l "validate_by_uri).
* [W3C Feed Validation Service mediante código](https://validator.w3.org/feed/#validate_by_input).
* [RSS Advisory Board](https://www.rssboard.org/rss-validator/).
* [googletransitdatafeed](https://github.com/mobilitydata/gtfs-validator).

# 6.- Utilización de herramientas.

Igual que con HTML, XHTML y XML, también existen editores para facilitar la creación de estos ficheros.

Poseen una interface que simplifica al máximo el trabajo con canales de contenidos. Un ejemplo de este tipo de herramientas es el [PSPad editor](http://www.pspad.com/es/).

**Características de estas herramientas**:

* Trabajar con **distintos estándares**.
* Importar documentos de texto CSV y HTML.
* **Editar HTML** con el editor WYSIWYG.
* **Editar documentos XML e imágenes**.
* **Actualizar las fuentes por vía FTP**.
* Tienen **capacidad para exportar documentos RSS a HTML**, CSV y JavaScript

Texto

Descripción generada automáticamente

# 7.- Directorios de canales de contenidos.

**Caso práctico**: después de tener el fichero fuentevalidado, es necesario sindicarle (registrar el fichero en un directorio de canales web).

**Permiten que el fichero RSS esté disponible para cualquiera**, además de **facilitar a los usuarios finales la búsqueda de información**, ya que los directorios de canales de contenidos **clasifican los ficheros RSS**.

**Para registrar un sitio en un directorio RSS**: es similar a registrar un sitio en un motor de búsqueda.

* [Syndic8 - The world's first BPM for E-commerce](https://syndic8.io/)

**Recomendaciones a la hora de crear contenido para los canales**:

Usar la codificación UTF-8 y las entidades XML que les sustituyen, es decir,

* ***&aacute;***en lugar de á,
* ***&eacute;***en lugar de é, ...,
* ***&lt;*** en lugar de <,
* ***&gt;*** en lugar de >,

**Para saber más**: [cómo promocionar un sitio web](https://developers.google.com/search/docs/crawling-indexing/ask-google-to-recrawl?utm_source=wmx&utm_medium=deprecation-pane&utm_content=submit-url).

# 8.- Agregación.

**Caso práctico**: Para poder leer los contenidos, es necesario un agregador de contenidos o lector de fuentes.

**¿Qué es un agregador o lector de fuentes?**

Es una aplicación de software para suscribirse a fuentes en formatos RSS y Atom. El agregador avisa al usuario o usuaria de qué webs han incorporado contenido nuevo desde nuestra última lectura y cuál es ese contenido.

**Tipos de agregadores**:

* **Los agregadores web (o agregadores en línea)**, son aplicaciones que residen en determinados sitios web y que se ejecutan a través de la propia web.
  + Recomendables cuando se accede desde distintos ordenadores.
  + **Ejemplos**: Es el caso de Google Reader, Bloglines o Netvibes.
* **Los agregadores de escritorio**, son aplicaciones/extensiones que se instalan en el ordenador del usuario o usuaria.
  + Su uso es aconsejable para quienes acceden a Internet siempre desde el mismo ordenador.
  + Su interfaz gráfica suele ser parecida a la de los programas de cliente de correo electrónico, con un panel donde se agrupan las suscripciones, y otro donde se accede a las entradas individuales para su lectura.
  + **Ejemplos de ellos son**: Newsgator y FeedDemon.

Interfaz de usuario gráfica, Texto, Escala de tiempo

Descripción generada automáticamente

# Bibliografía.

A continuación, presento la relación bibliográfica que he consultado para la realización de este trabajo.

**No hay ninguna fuente en el documento actual.**