

SINDICACIÓN DE CONTENIDOS EN EL AYUNTAMIENTO DE MADRID

José Luis Cano Giner
Mercedes Lozano Quirce

***Dirección General de Desarrollo de TI.
Informática del Ayuntamiento de Madrid***

1 INTRODUCCIÓN

En la actualidad, la Web del Ayuntamiento de Madrid (www.munimadrid.es) ofrece a los usuarios más de 200.000 contenidos de una gran diversidad de tipologías, que incluye noticias, actividades y eventos, publicaciones, entidades, organismos, instalaciones, trámites y un largo etcétera.

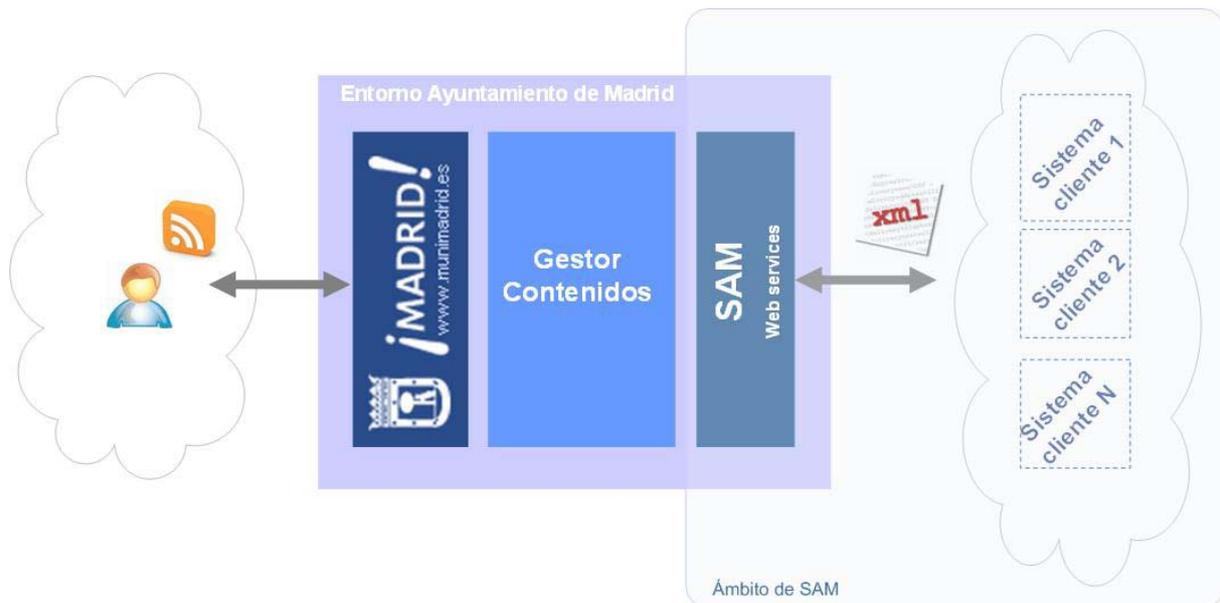
Toda esa información se mantiene actualizada y evoluciona permanentemente, centralizando las aportaciones de las diversas Unidades Gestoras del Ayuntamiento que diariamente incorporan, estructuran y clasifican nuevos contenidos y actualizan los existentes.

Este gran volumen de información, disponible en el Gestor de Contenidos Corporativo, constituye un enorme valor tanto para usuarios finales, incluyendo los propios informadores presenciales y telefónicos del Ayuntamiento, como para otros sistemas y aplicaciones que pueden hacer uso de él.

En esta línea, la Web del Ayuntamiento de Madrid ofrece desde el año 2008 un mecanismo de sindicación de contenidos, mediante canales RSS, orientado a usuarios finales, que les permite suscribirse a contenidos en función de diversas temáticas, secciones o ubicación geográfica.

Sin embargo, una vez habilitados dichos canales RSS, se detectó la necesidad adicional de implantar un sistema que permitiera compartir con otros portales y/o aplicaciones internas y externas, de forma automatizada y masiva, la ingente cantidad de información que ofrece la Web corporativa del Ayuntamiento.

Con este objetivo final, en el año 2009, el Organismo Autónomo Informática del Ayuntamiento de Madrid, desarrolló el Sistema de Sindicación del Ayuntamiento de Madrid (al que nos referiremos en adelante como **SAM**) cuyo objetivo final es facilitar que otros sistemas puedan acceder a los contenidos disponibles en el Gestor de Contenidos Corporativo, y ofrecer esta información **completa, estructurada** (XML) y **clasificada**.



2 SAM- SINDICADOR

2.1 VISIÓN GENERAL

Mientras que el mecanismo de sindicación a través de canales RSS esta orientado a usuarios, y ofrece una información mínima del contenido junto con un enlace al detalle del mismo en el portal, el sistema de sindicación SAM está orientado a sistemas, internos o externos, ofreciendo información completa y estructurada del contenido.

Actualmente el SAM facilita información relativa a los siguientes tipos de contenidos:

- **Actividades:** Información de los eventos celebrados en el municipio, con independencia de quien los promueva, con datos sobre lugar de celebración, fecha y hora, precio, venta de entradas, etc. Excepcionalmente se incluyen programas no municipales.
- **Entidades, organismos e instalaciones:** Contenidos que poseen dirección postal, teléfonos, fax, correo e., servicios prestados, horarios, etc., dependiendo del organismo de que se trate. Si se ubican en territorio municipal incluyen también información de posicionamiento geográfico.
- **Trámites:** Describen cualquier tramitación o gestión que el ciudadano pueda hacer ante la Administración.
- **Publicaciones:** Contienen información relativa a boletines, anuarios, folletos. Pueden ser de carácter periódico y estar organizados jerárquicamente en capítulos.
- **Enlaces:** Contenido auxiliar que contiene un enlace a una URL externa o a un fichero.

Los clientes pueden descargar contenidos en función de la **tipología** de contenido (tipos y subtipos), **taxonomías** en las que esté clasificado y, adicionalmente incluso, en

función del valor de **atributos** específicos del contenido, lo que permite a los clientes de sindicación una amplia flexibilidad para concretar el tipo de información a descargar.

El sistema ofrece exclusivamente contenidos que se encuentren publicados en la Web del Ayuntamiento de Madrid, por tanto de carácter público, y que hayan sido creados o modificados dentro de un rango cronológico que el cliente debe indicar en cada solicitud de descarga.

De este modo un cliente puede realizar descargas periódicas obteniendo exclusivamente las novedades del periodo y si lo desea mantener su información actualizada.

2.2 DISEÑO

El diseño del sistema de sindicación se ha realizado con las siguientes premisas básicas:

- Facilitar la comunicación con cualquier aplicación, con **independencia del lenguaje de programación y la plataforma** de ejecución.
- Utilizar protocolos de intercambio basados en **estándares** abiertos.
- Independizar el formato de intercambio de datos del formato particular de contenidos utilizado internamente en la plataforma de gestión de contenidos de Munimadrid.
- Escalabilidad. Es decir, el sistema de sindicación permite la incorporación de nuevos tipos de contenidos que se definan en el futuro dentro del gestor de contenidos y que puedan resultar de interés para otros sistemas.
- La descarga de contenidos debe realizarse bajo demanda. De forma que sea el sistema cliente el que solicita al sistema de sindicación la descarga de contenidos.

Bajo estas premisas, se ha diseñado un sistema de sindicación que expone un conjunto de servicios Web (Web services), que permiten a una aplicación externa la descarga de contenidos publicados en la Web del Ayuntamiento de Madrid, estableciendo, mediante parámetros definidos en cada servicio, qué tipo de información se desean descargar y el período de actualización.

Las características principales de este sistema se describen a continuación.

2.3 CARACTERÍSTICAS

Información estructurada

El formato de descarga de los contenidos es XML. El XML de cada contenido está estructurado en función de los atributos que componen el contenido. La DTD¹ que define la estructura del XML es común para todos los tipos, con independencia de su tipología, aunque los nombres y formatos de los atributos son específicos de cada tipo de contenido.

¹ DTD son las siglas del término inglés **Document Type Definition**. La DTD define la estructura y sintaxis de un documento XML. Su función principal es garantizar la consistencia entre todos los documentos que utilicen la misma DTD.

```
<!ELEMENT contenido (tipo, atributos+)>
<!ELEMENT tipo (#PCDATA)>
<!ELEMENT atributos (atributo+)>
<!ELEMENT atributo (#PCDATA | atributo)*>

<!ATTLIST atributos
  idioma ( es | en | fr ) #REQUIRED>
<!ATTLIST atributo
  nombre CDATA #REQUIRED>
```

DTD que define la estructura del XML de los contenidos

Información personalizada para cada cliente.

El sistema de sindicación permite *personalizar*, para cada cliente de sindicación, el conjunto de atributos del XML de descarga de cada tipo de contenido. Esto es posible puesto que el cliente debe identificarse en cada solicitud de descarga de contenidos. Sirva de ilustración el siguiente ejemplo: la información de los contenidos relativos a instalaciones u organismos puede incluir imágenes. En este caso, sería posible definir diferentes configuraciones para diferentes clientes, de forma que para unos clientes se incluyan las imágenes de las instalaciones y para otros no. De este modo se optimiza el conjunto de atributos en función de las necesidades particulares de cada cliente y se minimiza el tráfico entre ambos sistemas.

Adicionalmente, la configuración del sistema permite establecer restricciones relativas a la información que se permite descargar a cada cliente.

Información accesible

El sistema de sindicación está orientado a ofrecer a los sistemas clientes información accesible. De esta forma se garantiza que:

- El código HTML contenido en los atributos de texto enriquecido está validado según el nivel AAA de accesibilidad.
- Todos los ficheros que forman parte de un contenido incluyen la información del tamaño del fichero.
- Las URLS que se incluyen como valor de algunos atributos de los contenidos, se presentan codificadas conforme a los estándares HTML de W3C.

Información multidioma

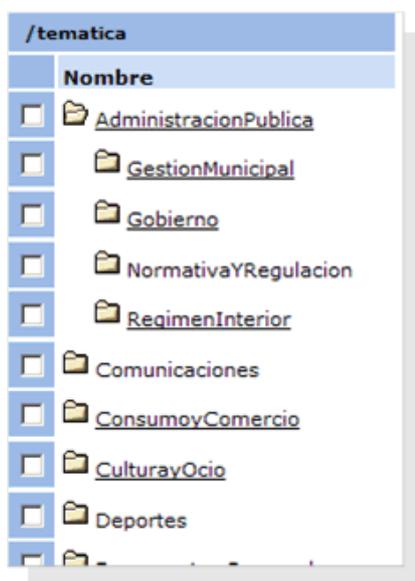
La Web del Ayuntamiento de Madrid ofrece información en varios idiomas (castellano, inglés y francés). Los sistemas clientes tienen la posibilidad de descargarse la información de los contenidos en cualquiera de los idiomas en los que se presente el contenido en la Web.

Filtros de la información a descargar

El cliente define qué tipo de información desea descargar en función de:

- **Tipo de contenido:** El cliente debe indicar en cada descarga un tipo de contenido.

- **Subtipos de contenido:** Si el tipo de contenido esta dividido en subtipos, pueden indicarse varios.
- **Taxonomías:** para concretar más la descarga el cliente puede establecer un filtro en función de las taxonomías. Existe en el gestor de contenidos un árbol de más de 2000 taxonomías agrupadas jerárquicamente en los grupos principales: *temática*, *ubicación geográfica*, *contenido*, *titularidad* y *usuario* (al que esta orientado el contenido).



Fragmento del árbol de taxonomías para el grupo *temática*

- **Atributos específicos:** opcionalmente, existe la posibilidad de realizar filtros con atributos concretos de los contenidos.

Información actualizada

Además de los filtros relativos al tipo de información a descargar, los clientes de sindicación pueden aplicar filtros cronológicos y de este modo limitar las descargas exclusivamente a contenidos que hayan sido creados o modificados en un período indicado. Esta utilidad es fundamental para los sistemas clientes que deseen mantener actualizada la información obtenida de *Munimadrid*. Así un cliente que desee mantener la información actualizada diariamente, necesitará realizar una conexión diaria al sistema de sindicación para descargar los contenidos creados o modificados cada día.

Se facilita también información de los contenidos que son eliminados del portal en un periodo de tiempo.

2.4 SERVICIOS DISPONIBLES

A continuación se indica la relación de los servicios que incluye SAM:

- **Consulta de contenidos:** El servicio permite realizar una consulta de contenidos a partir de los parámetros de búsqueda mencionados anteriormente (tipo de contenido, taxonomías, atributos, margen cronológico). El resultado obtenido es un listado de identificadores de contenidos que cumplen los criterios de búsqueda y que son usados posteriormente en el servicio de *descarga de contenidos para obtener los XML de los contenidos*.

Para realizar una descarga de contenidos el cliente debe invocar primero el servicio de consulta y posteriormente utilizar los identificadores resultantes en el servicio de descarga. Esto permite al cliente realizar descargas paginadas en función del número de resultados obtenidos.

- **Descarga de contenidos:** Este servicio permite descargar contenidos a partir de los identificadores obtenidos en el servicio de consulta. El servicio devuelve como resultado la información de cada contenido en formato XML.
- **Consulta de tipos disponibles para el cliente:** Este servicio permite consultar a un cliente de sindicación los tipos de contenidos disponibles.
- **Consulta de subtipos disponibles:** Este servicio permite a un cliente de sindicación consultar cuales son los subtipos de contenidos disponibles para un tipo de contenido dado.
- **Consulta de taxonomías:** Este servicio permite consultar cual es el conjunto de taxonomías disponibles para el cliente.
- **Consulta de contenidos eliminados:** Este servicio facilita la lista de los contenidos de cierto tipo dado, que han sido eliminados del portal Munimadrid durante el intervalo de tiempo indicado.

3 PUESTA EN MARCHA. CASOS PRÁCTICOS

El sistema de sindicación SAM se encuentra desplegado de modo que tanto aplicaciones internas como externas puedan efectuar descargas de contenidos publicados en *Munimadrid*.

Actualmente existen dos clientes de sindicación operando en real, uno que accede al sistema de sindicación SAM desde la red interna del Ayuntamiento (Base de datos ciudad) y otro que accede desde el exterior a través de una conexión HTTPS (www.lavozdelaexperiencia.es).

3.1 LA WEB DE LOS MAYORES DE MADRID: WWW.LAVOZDELAEXPERIENCIA.ES

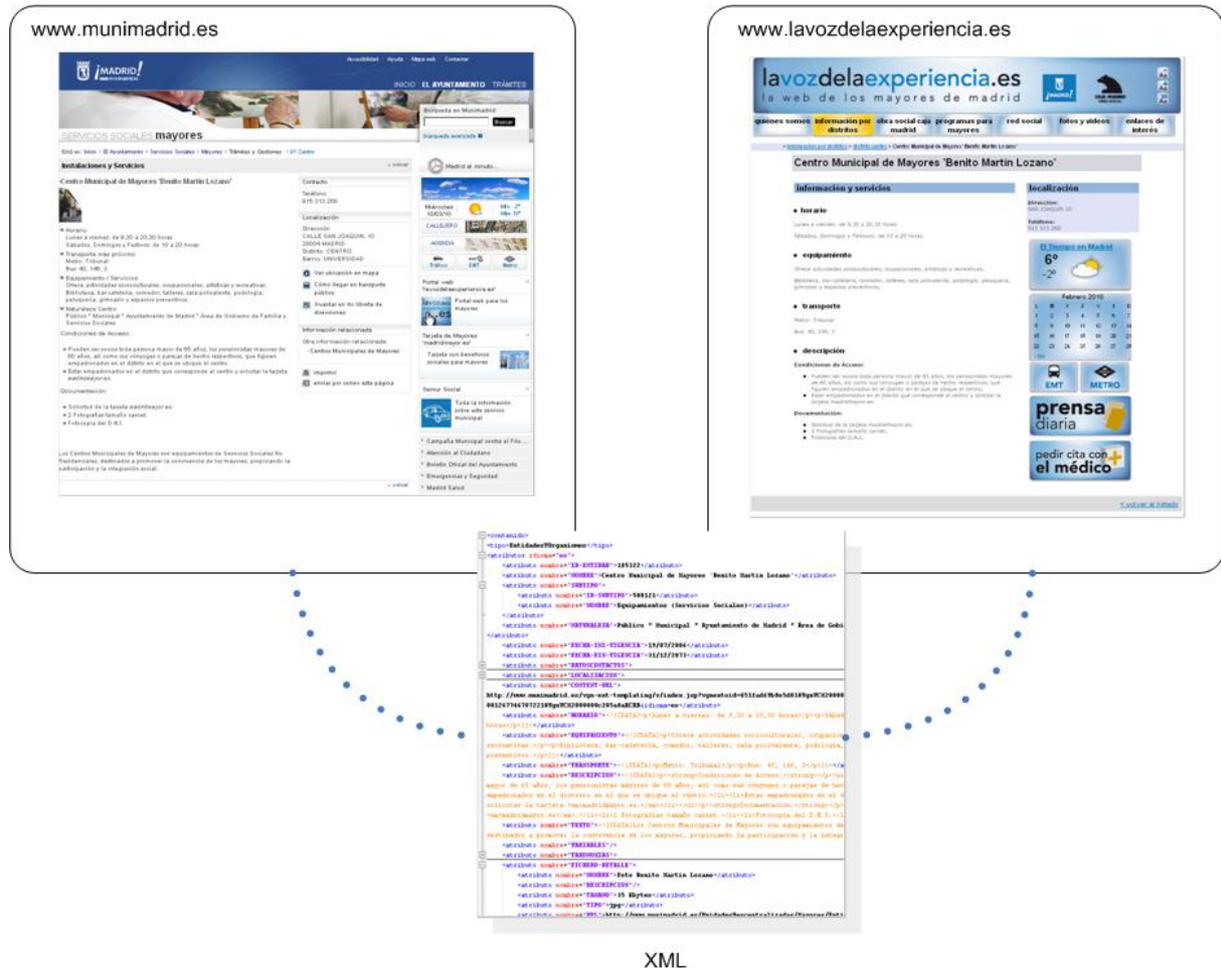
La voz de la experiencia es una Web orientada al colectivo de mayores de Madrid, promovida por el Área de Gobierno de Familia y Servicios Sociales del Ayuntamiento de Madrid, con el patrocinio de la Obra Social Caja Madrid.

La Web ofrece multitud de información orientada a los mayores: programas de mayores, voluntariado, instalaciones (centros de día, residencias, etc.), información multimedia y también ofrece una red social como canal de comunicación entre mayores.

Actualmente la Web de mayores, utiliza el sistema de sindicación SAM para obtener de *Munimadrid* información de actividades, entidades, organismos e instalaciones. Y presenta esta información integrada en su propio portal.

Para obtener la información relativa a los mayores utilizan filtros por tipología de contenido y taxonomías, principalmente de los grupos **temática**, **usuario** (orientado a mayores) y **ubicación geográfica** (para distinguir, por ejemplo, instalaciones Municipales de no municipales)

En un futuro inmediato está previsto que la Web de los Mayores obtenga de Munimadrid también información relativa a publicaciones, trámites y otros contenidos.



The image shows two web browser screenshots side-by-side. The left one is from www.munimadrid.es and the right one is from www.lavozdelaexperiencia.es. Both display information for the 'Centro Municipal de Mayores Benito Martín Lozano'. Below the screenshots, an XML snippet is shown, containing details such as the center's name, address, phone number, and opening hours. The XML is structured with tags like <nombre>, <ubicación>, <horario>, <equipamiento>, <transporte>, <descripción>, <condiciones de acceso>, and <documentación>.

Ejemplo de un contenido de tipo instalación integrado en la Web de mayores

3.2 BASE DE DATOS CIUDAD

Base de datos ciudad, en adelante BDC, es el callejero informatizado para el conjunto de aplicaciones informáticas del Ayuntamiento de Madrid.

BDC dispone de coordenadas UTM tanto en proyección ED50 como en ETRS89, lo que permite ofrecer servicios directos y responder a preguntas tales como:

- El portal(es) más cercano(s) a una posición
- Las calles cercanas a una posición.
- El plano – guía de una posición.
- Transportes más cercanos a una posición

Un objetivo básico de BDC es suministrar valor añadido a los servicios, especialmente en el ámbito geomático. En esta línea, actualmente se está trabajando en proporcionar este

tipo de información a los navegadores GPS para su uso por los ciudadanos. Dada la memoria, todavía limitada, de estos dispositivos, la iniciativa se ha dirigido a almacenes de POI (*point of interest*) independientes (por ejemplo de autobuses interurbanos, urbanos, metropolitano, cercanías o paradas de taxi).

Dando un paso más allá y puesto que el Ayuntamiento de Madrid almacena en su Gestor de Contenidos Corporativo una gran cantidad de información actualizada y veraz, se ha desarrollado un cliente de sindicación que operando contra SAM, descarga un conjunto de información (principalmente relativa a entidades, organismos e instalaciones) que, utilizando filtros de taxonomías principalmente de carácter geográfico pasan a formar parte de una tabla que permite su posterior uso en análisis espacial.

Actualmente se obtienen del SAM los siguientes contenidos:

- Oficina Turismo Municipal
- Centros culturales Municipales
- Juntas Municipales
- Centros Salud Municipal
- Oficinas de atención al ciudadano
- Polideportivos
- Cines
- Farmacias
- Gasolineras
- Hostales
- Hoteles
- Monumentos de la ciudad
- Museos
- Salas de conciertos
- Teatros
- Puntos limpios
- Bibliotecas
- Centros de salud
- Colegios
- Hospitales
- Oficinas de correos
- Parking público
- Comisarías de policía
- Tanatorios
- Intercambiadores de transporte

Tanto los transportes como el resto de la información se incorporan a otro proceso automatizado en el que se extrae la información para los navegadores GPS.

El formato de intercambio elegido ha sido GPX, un formato libre para GPS, basado en una estructura de datos propia sobre un esquema XML, y que, por sus excelentes prestaciones se ha convertido en un estándar "de facto".

En la actualidad y a través del software distribuido por los propios fabricantes, o bien a través de software libre disponible en la red, prácticamente la totalidad de dispositivos GPS admiten este formato.

Los datos proporcionados por el sistema de sindicación SAM ofrecen un enorme valor añadido, ya que no sólo es posible identificar y navegar hacia un POI (un teatro, por ejemplo), sino que se puede acceder directamente a la página de detalle del contenido en la versión PDA de la Web Corporativa Municipal, y obtener toda su

información adicional: horarios de taquilla, compra de entradas e incluso obra representada y funciones disponibles para la misma.

Este hecho abre grandes posibilidades no sólo para el ciudadano (hoy en día, casi todos los nuevos teléfonos del mercado incluyen GPS y navegación Internet), sino también para los propios servicios municipales que trabajan diariamente en la ciudad, que podrían disponer de una herramienta de ubicación geográfica, de consulta, e incluso de actualización, de aquellos puntos de interés sobre los que aplica su tarea cotidiana.

4 MARCO NORMATIVO

El Sistema de Sindicación de Contenidos descrito se enmarca en el Acuerdo de la Junta de Gobierno del Ayuntamiento de Madrid de 14 de enero de 2010, por el que se disponen los "Criterios de la reutilización de documentos del sistema de gestión de contenidos publicados en el sitio Web municipal (www.munimadrid.es)".

El contenido completo de este Acuerdo se encuentra publicado en el Boletín Oficial del Ayuntamiento de Madrid, de 18 de enero de 2010, páginas 9 y 10, y en el Boletín Oficial de la Comunidad de Madrid de 28 de enero de 2010, páginas 80 y 81.