

Título:

Plataforma para la investigación de contenidos sindicados: desarrollo del sistema ReSync y aplicación a los medios de comunicación hispano-mexicanos

Autores:

- Manuel Blázquez Ochando
Departamento de Biblioteconomía y Documentación, Facultad de Ciencias de la Documentación, Universidad Complutense de Madrid.
C/Santísima Trinidad, 37. CP 28010 Madrid.
Tlfno. +34 913946619
1
- Esmeralda Serrano Mascaraque
Departamento de Ciencias Sanitarias y Medicosociales, Facultad de Documentación de la Universidad de Alcalá.
C/San Cirilo s/n. CP 28805 Madrid
Tlfno. +34 885 24 15
esmeralda.serrano@uah.es

Resumen:

La diversificación de las vías de transmisión de información en la red ha configurado un complejo mapa de contenidos en el que los canales de sindicación juegan un papel importante, requiriendo un tratamiento especial que asegure su procesamiento y estudio científico. El objeto de estudio es el desarrollo del programa ReSync, que lleva a cabo dichas funciones, de forma tal que proporciona una plataforma de experimentación con la información recopilada masivamente a través de múltiples canales o fuentes de información. Esta función propia de parsers y agregadores de contenidos, tratará de ser ampliada, para superar sus límites. En este sentido se llevan a cabo pruebas de importación, tratamiento y recuperación de contenidos con los principales canales de sindicación de medios de comunicación hispano-mexicanos, de tal forma que puedan obtenerse el número total de noticias, tratamiento de la información en prensa, radio, televisión e internet de una misma noticia, datos cuantitativos del ratio de publicación y estadísticas de la topografía y estructura de los canales de sindicación.

Palabras clave:

Recuperación de información, sindicación de contenidos, parsers, gestión de la información, datamining, medios de comunicación

Introducción

Uno de los medios de transmisión y redifusión de información más importantes de la red son los canales de sindicación de contenidos. Éstos son utilizados por múltiples herramientas y sitios web para transmitir el contenido publicado a modo de resumen, titulares o texto completo, a un usuario determinado, que en un momento concreto se suscribe a sus fuentes. Este sencillo mecanismo tiene uno de sus máximos exponentes en los medios de comunicación digitales, ya sea radio, televisión o prensa. Por lo tanto resulta habitual comprobar la publicación de muy diversos canales de sindicación en directorios específicos que representan cada una de las categorías de noticias o informaciones de que disponen los medios. Ante esta producción de contenidos existen programas especializados en su lectura, denominados lectores o agregadores de canales de sindicación, con los que el usuario avanzado puede solucionar sus necesidades informativas de manera efectiva. ¿Pero qué ocurre cuando se desea abordar el estudio de la publicación de contenidos, en canales de sindicación de todos los medios de comunicación de un país, aplicando un método científico? El primer problema que se plantea es la cantidad de canales de sindicación susceptibles de ser analizados. Por ejemplo, programas como Google Reader se bloquean fácilmente cuando se supera la importación de 1.700 canales de sindicación (BLÁZQUEZ OCHANDO, M., 2011). El segundo problema es la categorización o clasificación de los mismos, ya que no siempre consta esta información en su etiquetado. El tercer problema es la recuperación y organización de todos los contenidos de cada canal. Esto se debe a la necesidad de recopilar en todo momento y de forma continuada la cambiante información publicada en dichos canales. Debe tenerse en cuenta que muchos canales de sindicación sólo muestran las últimas 15 o 20 noticias, por lo que según se sucede la publicación continua de las mismas, algunas de ellas terminan por desaparecer del canal. Aunque existen soluciones de gestión para canales de sindicación como el *Feeding Manager System* (PREECHAVEERAKUL, L. and Wongsawat, N., 2007), aún no se ha desarrollado una solución específica. En el presente trabajo se propone el programa ReSync, que permite la investigación de tal objeto de estudio, introduciendo operaciones y funciones muy concretas para el tratamiento de los canales de sindicación y sus contenidos, como se explicará a continuación.

Desarrollo del programa ReSync

El programa ReSync ha sido desarrollado para funcionar en servidores web basados en Apache, PHP y MySQL. A partir de este cuadro se han conformado las siguientes funciones y mecanismos del programa; 1) Importación de canales OPML, 2) Mantenimiento, 3) Actualización de canales, 4) Clasificación de canales, 5) Actualización de ítems, 6) Informes, 7) Búsquedas y consultas, 8) Gestión de administradores.



Figura 1. Portada del programa ReSync

1) Importación de canales OPML

La función de importación ha sido diseñada para recopilar todos los canales de sindicación que pueda contener un archivo OPML. De hecho éste resulta el método más efectivo para distinguir diversos tipos o categorías de fuentes según los archivos dispuestos. Por ejemplo los canales de sindicación de medios de comunicación españoles y mexicanos están divididos en dos archivos OPML que posteriormente son importados por el sistema y almacenados en una tabla de la base de datos para su posterior tratamiento.



Figura 2. Función de importación de archivos OPML

2) Mantenimiento

El mantenimiento de los canales de sindicación es esencial para el correcto funcionamiento del programa ReSync. Esto es debido a que permite comprobar y verificar los canales de sindicación importados, detectando errores de construcción sintáctica de los archivos XML, errores de caracteres e incluso extensiones no acordes con las estructuras de los principales formatos de sindicación RSS1.0, RSS2.0 o ATOM. Durante la ejecución del programa de mantenimiento, se visualiza en pantalla cada canal de sindicación y una indicación de validación o descarte del canal, así como la explicación del tipo de error acaecido. Al final la operación se eliminan los canales de sindicación defectuosos, quedando disponibles únicamente aquellos que han pasado todas las pruebas de carga, apertura, ejecución y lectura.

3) Actualización de canales

Depurados los canales de sindicación, se procede a su actualización. Esta función sólo actualiza la información descriptiva del canal, el título, su temática, autor y colaboradores, fecha de publicación y descripción general o resumen de los ítems o entradas que contiene el canal. En la consecución de este proceso se utilizan programas *PARSER del tipo New SimpleXMLElement* (BLÁZQUEZ OCHANDO, M., 2010, pp.217-227) que permiten recorrer todos los elementos del canal de sindicación, hasta obtener los propios de la meta-descripción del canal.

4) Clasificación de canales

Dado que con el proceso anterior no siempre se obtiene la categoría temática y el tipo de medio de comunicación del que procede cada canal de sindicación, las funciones de clasificación permiten al documentalista buscar los canales de sindicación por su título y

descripción y establecer las temáticas a las que corresponden en bloque o de forma selectiva, canal a canal.

5) Actualización de ítems

Cuando los canales están actualizados y clasificados, el siguiente paso es la actualización de ítems. Esta función recupera todos los contenidos publicados en los canales de sindicación. Para ello se emplean programas PARSER del tipo New DOMDocument (DOMElement class, 2011) que recorren la estructura de todos los ítems de los canales de sindicación, seleccionando mediante XPath (BLÁZQUEZ OCHANDO, M., 2010, pp.98-102) los contenidos correspondientes a su título, contenido completo, fecha de publicación, autor, categoría, fuente e identificador URI. Esta información es indexada y almacenada en una tabla *items* en la base de datos MySQL, directamente relacionada con la tabla *feeds* que contiene el canal de sindicación del que dependen dichas entradas.

Pero el proceso de actualización de ítems no acaba con el repaso de todos los canales disponibles. Esto se debe a la continua actualización de los contenidos de los canales de sindicación. De hecho, esta fluctuación de contenidos unida a la carencia de canales que guarden el historial completo de publicaciones, obligan a que el programa ReSync sea capaz de ejecutarse permanentemente para detectar cualquier cambio en los canales de sindicación, especialmente en sus entradas. Esta particularidad se lleva a cabo mediante un bucle temporizador que mantiene la ejecución del PARSER mientras no se supere la fecha y hora especificadas por el documentalista. Por ejemplo, de esta forma es posible programar el sistema para que se ejecute durante el tiempo que sea necesario de cara a un estudio que contemple 24 horas de publicaciones o un servicio permanente que dure 6 meses.

6) Informes

Las funciones correspondientes a informes permiten obtener datos del número de canales de sindicación disponibles según sus categorías temáticas, tipos de medios de comunicación, medios de comunicación con más canales de sindicación, número de ítems o entradas recopiladas en el tiempo establecido, número de ítems según términos o frases de consulta y recuentos totales. Tales datos serán probados en los subsiguientes puntos relativos al análisis de los medios de comunicación hispano-mexicanos.

7) Búsquedas y consultas

Las operaciones de búsqueda también son incluidas en modo *MySQL Fulltext* (Full-Text Search Functions, 2010) siendo posible una rápida recuperación de información por medio de campos especiales denominados *indexer*, presentes en las tablas *feeds* e *items*, que almacenan los textos indexados y depurados en los procesos de actualización de contenidos. De esta forma el usuario puede recuperar todas las noticias de todos los medios de comunicación sobre un tema, términos o frase que sea de su interés.

8) Gestión de administradores.

En cuanto al factor administrador del programa ReSync, se permite la gestión y administración de los usuarios con derechos y privilegios de uso en el sistema, pudiendo editar o clasificar a la vez múltiples canales de sindicación, sin que el programa sufra problemas sincronía con respecto a las operaciones en la base de datos.

Pruebas de ejecución en ReSync: estudio de los medios de comunicación hispano-mexicanos

Para verificar el funcionamiento del programa ReSync, se procede al estudio de los medios de comunicación hispano-mexicanos, concretamente a través de sus canales de sindicación. Para su consecución se hace necesario elaborar un plan de trabajo con las siguientes actuaciones; 1) Elaboración de una muestra exhaustiva de los canales de sindicación de los medios de comunicación españoles y mexicanos, 2) Construcción de archivos OPML que contengan los canales de sindicación obtenidos, 3) Importar los canales de sindicación al programa ReSync, 4) Clasificar los canales según medios y categorías, 5) Actualizar ítems de los canales de sindicación durante 24 horas.

1) Elaboración de una muestra exhaustiva de los canales de sindicación de los medios de comunicación españoles y mexicanos

Para elaborar una muestra exhaustiva de los canales de sindicación en los medios de comunicación de ambos países, se hace imprescindible en primer lugar indagar en relación a los principales grupos de comunicación o conglomerados empresariales relacionados con la prensa, radio y televisión (MANCINAS CHÁVEZ, R., 2009, pp.159-325). En este sentido se han investigado la mayoría de los grupos de comunicación, concretamente 7 conglomerados mexicanos y 14 españoles, *véase la tabla 1*.

Grupos de comunicación analizados	
México	España
1. G. Imagen	1. G. Antena3
2. G. Televisa	2. G. Cope
3. G. OEM	3. G. Forta
4. G. Azteca	4. G. Godó
5. G. Multimedios	5. G. Intereconomía
6. G. Reforma	6. G. Prisa
7. G. Radio Fórmula	7. G. Planeta
	8. G. RTVE
	9. G. Unidad Editorial
	10. G. Vocento
	11. G. Voz
	12. G. Recoletos
	13. G. Moll
	14. G. Mediapro

Tabla 1. Principales grupos de comunicación mexicanos y españoles

Terminado este reconocimiento, el proceso de recopilación ha resultado complejo ya que se ha elaborado de forma manual, visionando todos los sitios web de prensa, radio y televisión disponibles en tales grupos. Ello se ha debido en parte a la falta de archivos OPML en la red que contengan recopilaciones ya dadas, de este tipo de recursos, especialmente de medios de comunicación hispano-mexicanos. También este hecho demuestra la práctica inexistencia de estudios similares partiendo de este enfoque.

2) Construcción de archivos OPML que contengan los canales de sindicación obtenidos

Los canales de sindicación obtenidos son editados inicialmente en un archivo OPML, que contiene la información mínima y básica de cada uno de ellos. Esto es el título y subtítulo si procede, la URL del canal de sindicación, la URL de la fuente o sitio web del que procede y un indicador de lugar para distinguir su procedencia geográfica.

```

1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <opml version="1.1">
3 <head>
4 <title>Medios de comunicación españoles</title>
5 <dateCreated>Sat, Nov 03 2010 10:35:14 GMT</dateCreated>
6 <dateModified>Sat, Nov 03 2010 10:35:14 GMT</dateModified>
7 <ownerName>Esmeralda Serrano; Manuel Blázquez Ochando</ownerName>
8 <ownerEmail>esmeralda.serrano@gmail.com; mblazquez.ccdoc@gmail.com</ownerEmail>
9 </head>
10 <body>
11 <outline text="El Mundo - Podcast resumen de la PORTADA" type="rss" xmlUrl="http://rss.elmundo.es"
12 <outline text="El Mundo - Podcast del resumen de noticias 60 SEGUNDOS" type="rss" xmlUrl="http://
13 <outline text="El Mundo - Portada" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo/rss/
14 <outline text="El Mundo - España" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo/rss/e
15 <outline text="El Mundo - Internacional" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo
16 <outline text="El Mundo - Unión Europea" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo
17 <outline text="El Mundo - Cultura" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo/rss/
18 <outline text="El Mundo - Ciencia y ecología" type="rss" xmlUrl="http://elmundo.feedsportal.com/e
19 <outline text="El Mundo - Madrid 24h" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo/r
20 <outline text="El Mundo - Barcelona" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo/rs
21 <outline text="El Mundo - Baleares" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo/rss
22 <outline text="El Mundo - Castilla y León" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmun
23 <outline text="El Mundo - Valencia" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo/rss
24 <outline text="El Mundo - Alicante" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo/rss
25 <outline text="El Mundo - Castellón" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo/rs
26 <outline text="El Mundo - País Vasco" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo/r
27 <outline text="El Mundo - Andalucía" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo/rs
28 <outline text="El Mundo - Sevilla" type="rss" xmlUrl="http://elmundo.feedsportal.com/elmundo/rss/

```

Figura 3. Archivo OPML con canales de sindicación de medios de comunicación españoles

El resultado de este proceso es la constitución de dos archivos OPML que contienen todos los canales de sindicación españoles *resync_spain.xml* y mexicanos *resync_mexico.xml*, relativos a medios de comunicación.

3) Importar los canales de sindicación al programa ReSync

El proceso de importación se lleva a cabo de forma automática con las funciones específicamente destinadas a este fin, obteniendo los primeros datos cuantitativos del estudio, véase *tabla2*. En total se ha recopilado una muestra de 3.324 canales de sindicación, de los que 1.002 corresponden a medios mexicanos y 2.322 a medios españoles.

	Canales recopilados	Canales defectuosos	Total Canales útiles
México	1.002	82	920
España	2.322	81	2.241
Total	3.324	163	3.161

Para más información puede consultar el listado de canales en: http://www.mblazquez.es/documents/resync-list_2011-01-14.html

Tabla 2. Canales de sindicación de medios de comunicación: cifras generales

Concluida esta operación, se necesita validar y depurar los canales de sindicación importados al sistema. En este punto se obtiene información de utilidad, ya que no todos los canales están disponibles, otros son defectuosos o no están contruidos de acuerdo a las especificaciones de cada formato de sindicación. En la muestra obtenida se advierten pocos canales defectuosos, concretamente 163 en los que su principal fallo estriba en la codificación de caracteres especiales. Ello da al traste con la labor de lectura y recuperación en muchos sistemas, por lo que en la *tabla3*, puede comprobarse qué medios han fallado la prueba y el número de canales afectados.

Medios de comunicación mexicanos			
Medio	URL general	Nº de canales defectuosos	Tipo de error
El Siglo de Durango	http://www.elsiglodedurango.com.mx/	8	Caracteres incorrectos
La Crónica	http://www.lacronica.com	7	
Frontera.info	http://www.frontera.info/	7	
El Imparcial	http://www.elimparcial.com/	7	
El Debate	http://www.debate.com.mx/	18	
El Mercurio	http://www.elmercurio.com.mx/	10	
Diario Provincia	http://www.provincia.com.mx/	12	
Periódico el Vigía	http://www.elvigia.com.mx/	1	
Boletín turístico	http://www.boletinturistico.com.mx/	4	
OEM en línea	http://www.oem.com.mx/	8	
Medios de comunicación españoles			
Medio	URL general	Nº de canales defectuosos	Tipo de error
ABC	http://www.abc.es/	9	Caracteres incorrectos
Hoy motor	http://www.hoymotor.com/	1	
Hoy cinema	http://www.hoycinema.com/	1	
Diario médico	http://www.diariomedico.com/	2	
Plus.es	http://www.plus.es/	1	
Cinemanía	http://www.cinemanía.es/	6	
Onda Cero	http://www.ondacero.es/	2	
RAC1	http://www.rac1.org/	1	
Icatfm	http://www.icatfm.cat/	1	
RTVC	http://www.rtv.es/	48	
El Economista	http://www.eleconomista.es/	1	
El Correo Web	http://www.elcorreoweb.es/	8	
Para más información puede consultar el listado completo en: http://www.mblazquez.es/documents/resync-errors_2011-01-14.html			

Tabla 3. Canales de sindicación defectuosos.

En cuanto a los medios de comunicación con mayor número de canales de sindicación destaca principalmente, Radio Televisión Española con más de 400 canales, muy por encima del resto, tal y como se puede apreciar en la *tabla4*. Esta cifra tan alta se debe al empleo sistemático de canales de podcasting para todos los programas de radio actualmente en vigor y para aquellos que dejaron de emitirse. Este caso hace que los canales de sindicación de RTVE sean un ejemplo de fonoteca en redifusión permanente.

México		España	
Medio	total	Medio	total
http://www.notisistema.com/	88	http://www.rtve.es/	427
http://www.tvazteca.com/	70	http://www.20minutos.es/	125
http://www.excelsior.com.mx/	68	http://www.eitb.com/	110
http://www2.esmas.com/	53	http://www.tv3.cat/	106
http://www.imer.gob.mx/	45	http://www.elmundo.es/	88
http://www.noroeste.com.mx/	33	http://www.libertaddigital.com/	83
http://www.cnnexpansion.com/	25	http://www.cincodias.com/	83
http://www.eluniversal.com.mx/	23	http://www.lavanguardia.es/	74
http://www.elsiglodetorreon.com.mx/	22	http://www.cope.es/	63
http://correo-gto.com.mx/	21	http://www.ondacero.es/	62

Tabla 4. Medios de comunicación mexicanos y españoles con más canales de sindicación

4) Clasificar los canales según medios y categorías.

Seguidamente se procedió a la clasificación de los canales atendiendo al tipo de medio de comunicación con las variables: radio, televisión, radiotelevisión y prensa. De los datos obtenidos destacan la prensa y la radio en España. En México destaca igualmente la prensa y en segundo lugar la televisión, debido esencialmente al vago empleo de los canales de sindicación para los servicios de podcasting, véase tabla5.

	Radio- televisión	Radio	televisión	Prensa	Otros
México	11	64	155	682	8
España	189	788	214	1.025	25
Total	200	852	369	1.707	33

Tabla 5. Distribución de canales de sindicación según tipología de medios

La clasificación temática se elaboró conforme a las siguientes categorías: noticias locales, regionales, nacionales, internacionales, última hora, política, sociedad, economía, justicia, finanzas, mercados, ciencia, tecnología, cultura, educación, salud, deportes, opinión, medios y entretenimiento. El resultado total obtenido es significativo, dado que permite observar cinco grandes temas de contenidos; noticias (677), opinión (536), entretenimiento (399), economía (294) y deportes (195).

Categoría temática	México	España	Total
Noticias de última hora	69	66	135
Noticias locales	72	155	227
Noticias regionales	60	93	153
Noticias nacionales	54	23	77
Noticias internacionales	45	40	85
Política, sociedad y justicia	20	32	52
Economía, finanzas, bolsas y mercados	67	227	294
Ciencia y tecnología	26	39	65
Cultura y educación	42	68	110
Salud	25	79	104
Deportes	67	128	195
Opinión	106	430	536
Medios	31	50	81

Entretenimiento	165	234	399
Otros	60	54	114

Tabla 6. Canales de sindicación según su temática

Alternativamente, se ha considerado el estudio de tres parámetros que si bien no son clasificatorios desde un punto de vista temático, sí lo son desde la topología o estructura de los medios de comunicación. Estos son la presencia de blogs en medios de comunicación, la utilización de canales Youtube y el empleo de canales de sindicación dedicados al podcast, véase tabla7.

	México	España	Total
Blogs	95	375	470
Youtube	24	21	45
Podcast	5	522	527

Tabla 7. Canales de sindicación en blogs, Youtube y podcast

De la tabla se puede observar una importante presencia de los blogs en los medios de comunicación españoles, así como de podcast, constituido en su mayoría por los canales de RTVE. En cambio el número de canales de videos en Youtube es muy similar al que disponen los medios mexicanos.

5) Actualizar ítems de los canales de sindicación durante 24 horas.

Finalmente el último punto del plan de trabajo, se centra en la recuperación de los contenidos publicados en los canales de sindicación durante un periodo de tiempo de 24 horas a partir del viernes 14 de enero de 2011 a las 11:55 hora española. El programa ReSync fue configurado para trabajar ininterrumpidamente durante ese tiempo, obteniendo los resultados que figuran en la tabla8.

	México	España	Suma total
Total contenidos publicados	24.369	36.161	60.530
Noticias nacionales	2.230	3.388	5.618
Noticias internacionales	1.673	4.090	5.763
Noticias locales y regionales	1.683	2.797	4.480
Política	34	113	147
Sociedad	212	581	793
Cultura y educación	523	965	1.488
Ciencia y tecnología	257	435	692
Economía	82	117	199
Bolsa	134	215	349
Mercados	79	186	265
Finanzas	90	61	151
Salud	562	578	1.140
Comunicación	92	201	293
Deportes	482	914	1.396
Fútbol	298	281	579
Música	225	635	860
Radio	236	836	1.072
Cine	252	756	1.008
Literatura y libros	83	470	553

Guerra, conflictos armados	152	378	530
Sucesos	177	333	510
Gobierno e institucional	1.221	2.162	3.383
Motor	100	273	373
Biblioteconomía y Documentación	37	71	108
Suma por categorías	10.914	20.836	31750
Resto	13.455	15.325	28780

Tabla 8. Contenidos recuperados en 24 horas

Tal como se muestra en la *tabla8*, la mayoría de los contenidos recuperados corresponden a noticias nacionales, internacionales, locales y regionales en los medios mexicanos y españoles, agrupando un total de 15.861 entradas. A continuación destacan las informaciones cuyo objeto y referente es el gobierno, la administración pública y las instituciones con más de 3.300 contenidos. Sucesivamente destacan cultura y educación con 1.488 entradas, deportes con 1.396 registros y finalmente las categorías de salud, radio y cine superando el millar de noticias.

Futuras líneas de investigación y mejoras

El programa ReSync es una plataforma aún en desarrollo y existen mejoras pendientes que lograrán refinar su eficacia e información. En este sentido se prevé la introducción de programas de análisis de frecuencias de palabras que constituyen el corpus de contenidos recopilados por el sistema, de tal forma que puedan conocerse los términos más profusos y aquellos más discriminatorios, en línea con las técnicas de recuperación de información.

Por otro lado, futuras investigaciones relativas a medios de comunicación deberán comprender el uso de canales de sindicación propios de la web social, como es el caso de Twitter y Facebook. De hecho es bien conocido que la mayoría de los medios de comunicación tienen presencia en las redes sociales y éstas disponen de canales de sindicación públicos que pueden ser aprovechados para su estudio conjunto, separado o paralelo a las noticias que los medios publican.

Por último y en relación a la mejora de las prestaciones de la herramienta, se tiene prevista la compatibilidad de ReSync con sistemas de webcrawling, de tal forma que el rastreo de la web, suponga una fuente de información directa a través de la que proveer canales de sindicación. Dicho de otra forma, lograr una alimentación automática del sistema. Este proceso aún siendo complejo liberaría al documentalista de una enorme cantidad de trabajo en el proceso de elaboración de muestras.

Conclusiones:

1. El programa ReSync constituye una plataforma válida para la investigación de los contenidos publicados en canales de sindicación, ya que tiene en cuenta su validación, clasificación y recuperación conformando el ciclo completo de tratamiento documental. En este caso concreto, está demostrada la capacidad del sistema para obtener datos de producción y publicación de contenidos de los distintos medios de comunicación mexicanos y españoles.

2. El estudio de los contenidos publicados en un canal de sindicación, depende del factor tiempo, ya que son regenerados y no siempre albergan el historial completo de entradas. La alta obsolescencia que comporta este fenómeno implica necesariamente la recuperación continua de la información, obligando a cualquier programa especializado la ejecución permanente, controlada mediante temporizador o bucles temporales.
3. Los principales medios de comunicación mexicanos constan de 920 canales de sindicación que publican más de 24.000 noticias al día. En España los medios constan de 2.241 canales que difunden más de 36.000 noticias diarias. Ello implica que los medios mexicanos publican 26,5 noticias por canal frente a 14 noticias por canal de los medios españoles, significando una diferencia considerable de 12,5 contenidos más.
4. La composición de los medios en México marca una preponderancia de la prensa con un 75%, seguida de la televisión y la radio con un 17% y un 7% respectivamente. Esta composición difiere de la estructura de medios en España, conformada por la prensa en un 46%, la radio en un 35% y la televisión en un 8%. Este cuadro dispar en torno a la radio se debe a la gran cantidad de canales de sindicación especializados en podcast cuyos contenidos superan el millar.
5. Existe una correlación entre los canales de sindicación dedicados a la opinión y el número de blogs detectados en las muestras. En España se clasifican 430 canales de sindicación especializados en opinión y se obtiene a su vez que al menos 375 canales corresponden a blogs. En México sucede el mismo fenómeno, de los 106 canales de opinión, se obtiene que al menos 95 corresponden a blogs. Esta relación queda demostrada teniendo en cuenta que la cifra de canales de blogs resulta ligeramente inferior a la de canales de opinión a los que con mucha frecuencia corresponden.
6. Resulta significativo el hecho de que no son suficientes 24 categorías temáticas para clasificar los contenidos recuperados en 24 horas en 3.161 canales de sindicación hispano-mexicanos. De hecho, en conjunto quedan 28.780 contenidos que deberán ser analizados para completar el cuadro temático de la producción informativa de ambos países.
7. La Biblioteconomía y Documentación, aún con una cifra reducida, tiene una presencia razonable, dentro de la producción informativa de los medios de comunicación constando en suma de 108 noticias o contenidos especializados. Estas cifras resultan considerables si se tiene en cuenta que la fuente o los canales de sindicación proceden de medios de comunicación generales.
8. Existe una clara estratificación de canales de sindicación que diferencian noticias locales, regionales, nacionales e internacionales, así como de aquellas consideradas de portada o de última hora, conformando una parte importante de todo lo publicado en canales de sindicación con 15.861 entradas.

Bibliografía:

- BLÁZQUEZ OCHANDO, M. 2010. [Tesis Doctoral] *Aplicaciones de la sindicación para la gestión de catálogos bibliográficos*. Universidad Complutense de Madrid.
- BLÁZQUEZ OCHANDO, M. 2011. *Probando los límites de Google Reader*. [online]. [visitado en 16 Enero 2011]. Disponible en: < HYPERLINK "http://www.mblazquez.es/documents/articulo_limites-google-reader.html" http://www.mblazquez.es/documents/articulo_limites-google-reader.html >
- *DOMElement class*. 2011. [online]. [visitado en 14 Enero 2011]. Disponible en: < HYPERLINK "http://www.php.net/manual/en/class.domelement.php" <http://www.php.net/manual/en/class.domelement.php> >
- *Full-Text Search Functions*. 2010. [online]. [Visitado en 23 Diciembre 2010]. Disponible en: < HYPERLINK "http://dev.mysql.com/doc/refman/5.0/en/fulltext-search.html" <http://dev.mysql.com/doc/refman/5.0/en/fulltext-search.html> >
- MANCINAS CHÁVEZ, R. 2009. [Tesis Doctoral] *La estructura mediática de México y el caso del Estado de Chihuahua: Prensa, Radio e Internet*. Universidad de Sevilla.
- PREECHAVEERAKUL, L. and N. WONGSAWAT. 2007. Feeding Manager System (FMS). *en: Proceedings of the 2007 International Conference on Internet Computing*. Las Vegas.