

# **Aplicación de la teoría de la relevancia a una red social de noticias**

**Yolanda Mauri Pérez**

Universidad de las Ciencias Informáticas, La Habana, Cuba, ymauri@uci.cu

**Ernesto de la Soledad García**

Universidad de Matanzas Camilo Cienfuegos, Matanzas, Cuba, ernesto.soledad@umcc.cu

**Ramiro Manuel Vázquez Villar**

Grupo Empresarial Comercial Caracol (Sucursal Casa Matriz), La Habana, Cuba, rmvazquez@caracol.cu

## **RESUMEN**

El sistema 4x4News es una red social cuyo eje central es un lector de canales RSS que muestra en una matriz de dimensión cuatro las dieciséis noticias más relevantes. El criterio de relevancia viene dado por parámetros estadísticos que toman en cuenta la interacción del usuario con la noticia ya que no siempre la última noticia publicada es la más relevante para el usuario. 4x4News permite a cada miembro de la comunidad crear un perfil que incluye las características de las noticias que desea ver, recibir una cobertura de un acontecimiento determinado, crear almacenes de noticias o definir la frecuencia de actualización de la matriz. Los usuarios podrán interactuar con personas de dentro y fuera del sistema recomendando, comentando noticias y compartiendo los titulares a través de las redes sociales. 4x4News es una aplicación web con soporte de datos desarrollada con los frameworks Symfony 2.2, Bootstrap 2.2.1 y jQuery 1.8.0.

**Palabras claves:** lector RSS, noticia, red social, relevancia.

## **ABSTRACT**

4x4News is a social network whose centerpiece is an RSS reader that shows in a four-dimensional matrix the sixteen most relevant news. The relevance criterion is given by statistical parameters that take into account the user's interaction with the news because not always the latest news published is the most relevant to the user. 4x4News allows each member of the community to manage a profile that includes news features to view, receive coverage of a particular event, create news stores or define the refresh rate of the matrix. Users can interact with people inside and outside the application recommending and commenting on news or share headlines through other social networks. 4x4News is a web application support data, it was developed with Symfony 2.2, Bootstrap 2.2.1 and jQuery 1.8.0 frameworks.

**Keywords:** news, relevance, RSS reader, social network.

## **1. INTRODUCCIÓN**

Como parte del flujo constante de datos en Internet se encuentran las redes sociales que según el ONTSI (Oficina Nacional de las Telecomunicaciones y de la Seguridad Informática) devienen como una herramienta de democratización de la información que transforman a las personas en receptores y productores de contenidos, permitiéndoles comunicarse fácilmente y crear comunidades (Urueña, y otros, 2011). Ofrecen servicios que permiten compartir a través de ellas los contenidos de otros sitios web. Lo anterior puede verse como una

limitante para los usuarios ya que en la red social solo acceden a la información que sus contactos hayan compartido. La arquitectura de la información de las redes sociales tiende a resaltar los elementos visuales en la presentación del contenido, permitiendo al lector retener a corto y largo plazo el mensaje que se le transmite.

Otro medio para difundir noticias lo constituyen los canales RSS, Really Simple Syndication. Este término hace referencia al sistema completo por el cual una página publica información a través de los canales o fuentes RSS y otras personas pueden acceder a los titulares, noticias o actualizaciones utilizando su lector de RSS.

Básicamente un archivo RSS es un XML que contiene la información reciente como la fecha de publicación, el titular o imágenes asociadas del sitio que provee el servicio. Otras terminologías asociadas a este tema son los canales o fuentes RSS (feeds) que son el medio al cual se suscriben los usuarios para recibir la información. Las aplicaciones dedicadas a recolectar la información de los feeds se denominan lectores RSS (RSS Reader).

Por lo general estas aplicaciones presentan buena arquitectura de la información pero a la hora de organizar los titulares no realizan un buen análisis de la relevancia de la noticia, teniendo en cuenta que no siempre la última noticia publicada es la más relevante. En los lectores de RSS el usuario está limitado a ver los titulares de las noticias impidiéndole personalizar el contenido de acuerdo a lo que él considere relevante.

Tomando como punto de partida la popularidad de las redes sociales y la cantidad de noticias que en todo momento hay disponibles en internet, el objetivo de la presente ponencia es caracterizar en cuanto a arquitectura, funcionalidades y tecnologías empleadas en el desarrollo de 4x4News, una red social donde a partir de un lector RSS se le muestran al usuario las noticias más relevantes para él y la comunidad.

En el presente artículo primero se abordan algunas características de los **lectores RSS en Internet** que se tomaron como referente para la investigación. Luego se exponen los aspectos que intervienen en el criterio de **relevancia de la noticia**, se explican la **arquitectura y las tecnologías** empleadas en la codificación del sistema y se detalla el funcionamiento de la **Red social de noticias 4x4News**. Finalmente se realizan breves **conclusiones** acerca de los resultados alcanzados con la implementación de este software.

## 2. DESARROLLO

### 2.1 LECTORES RSS EN INTERNET

En internet hay disponibles gran variedad de aplicaciones dedicadas a consultar canales RSS. TenByTen es un lector de código abierto y gratuito que presenta en una matriz de orden diez las cien imágenes asociadas a las palabras más recurrentes en los titulares que circulan en la red. Cada hora, el sistema analiza los flujos RSS de varias fuentes principales de noticias internacionales, y lleva a cabo un complejo proceso de análisis lingüístico ponderado en el texto que figura en sus principales noticias. Después de este proceso, se seleccionan las primeras cien palabras junto a las imágenes correspondientes extraídas de sus respectivas fuentes. Al final de cada semana, cada mes y cada año muestra una matriz con las cien imágenes y palabras más recurrentes en esos respectivos períodos de tiempo (Harris, 2013).

Por otro lado se tiene a Newsmap cuya filosofía es muy creativa ya que la página muestra las noticias extraídas desde las fuentes configuradas por el usuario seccionadas en cuadrados con colores y tamaños de fuente diferentes; los colores representan las categorías de las noticias y el tamaño de la fuente pone de manifiesto la actualidad de la misma (mientras mayor más reciente). Este sistema brinda una organización muy completa de las fuentes RSS a consultar, que pueden ser seleccionadas por el usuario a partir de un mapa interactivo que organiza los feeds por regiones del mundo (Newsmap, 2013).

Google Noticias forma parte de los servicios de la compañía Google, por lo que está directamente integrado con la red social GooglePlus. Este lector RSS consulta un número considerable de fuentes en dependencia del idioma seleccionado por el usuario. Muestra los artículos organizados en grupos, clasificándolos mediante un sistema que evalúa la frecuencia de los mismos y las características de las diferentes fuentes. Asociada a cada noticia se muestran una serie de titulares de temas relacionados (Google, 2013).

Otros espacios ideales para difundir los titulares son las redes sociales como por ejemplo Facebook, Twitter o GooglePlus, pero son redes de propósito general, donde los usuarios comparten además sus historias personales que no forman parte del acontecer noticioso. Esto provoca que los datos de las noticias que le son notificadas al usuario de la red estén mezclados con otra información menos notoria como actualizaciones de estados, fotografías o videos del resto de la comunidad.

## **2.2 RELEVANCIA DE LA NOTICIA**

En las soluciones anteriores se percibe claramente la intención marcada de brindar a los usuarios la noticia de último momento y más interesante para él, pero no toman en cuenta un aspecto significativo: la interacción del resto de los usuarios con el mismo titular. Puede decirse entonces que estos lectores pasan por alto que “no siempre la última noticia publicada es la más relevante para la comunidad de usuarios”.

Partiendo de esta afirmación el sistema 4x4News se basa en la teoría de la relevancia de (Wilson, y otros, 2004). La misma se ocupa de analizar la tendencia a la búsqueda de información más interesante y guía a las personas hacia la interpretación más adecuada de cada mensaje transmitido. La noticia que se publique en internet debe mostrar un equilibrio informativo entre el texto, hipervínculos e imágenes de forma tal que induzca a una interpretación adecuada del tema que aborde.

La red social 4x4News muestra en matrices de orden cuatro grupos de dieciséis noticias ordenadas por un criterio de relevancia determinado por los siguientes aspectos:

- El número de fuentes donde se encuentra publicado el titular. La fuente no es más que el sitio proveedor del canal RSS.
- Similitud de los titulares. En ocasiones los títulos de las noticias no son completamente iguales, pero tienen un alto porcentaje de palabras coincidentes y por ende los contenidos de las noticias probablemente tengan puntos en común.
- Número de vistas al titular en 4x4News.
- Número de visitas desde 4x4News hacia las fuentes del titular.
- Valoración del titular que puede ser positiva o negativa.
- Cantidad de comentarios que haya recibido el titular dentro del sistema.
- Actualidad del titular tomando como base su fecha de publicación.

Abordar todos estos valores estadísticos dentro del valor de relevancia que se le confiere a un titular permite poner los intereses de cada usuario al servicio del resto de la comunidad.

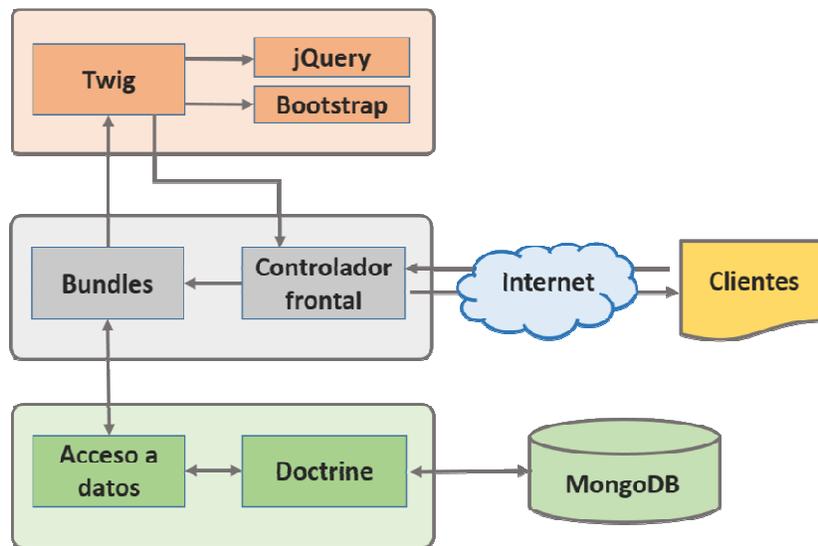
## **2.3 ARQUITECTURA Y TECNOLOGÍAS**

Para la implementación del lector RSS 4x4News se utilizó como framework de desarrollo principal Symfony 2.2 (Symfony 2.0, 2012) por lo que el sistema se sustenta en el estilo arquitectónico Modelo-Vista-Controlador (MVC). El uso de este marco de trabajo permitió estructurar el código fuente de forma tal que se faciliten las tareas de mantenimiento y la extensibilidad en la aplicación a partir del mecanismo de bundles. Para la gestión de usuarios se utilizó el FOSUserBundle (Gómez, 2012), como una mejora al componente de seguridad del framework se utilizó JMSSecurityExtraBundle (Symfony, 2013) y para la comunicación con las redes sociales GooglePlus, Facebook y Twitter fueron empleados FOSGoogleBundle (knpBundles, 2012), FOSFacebookBundle y FOSTwitterBundle (Symfony.es, 2013) respectivamente.

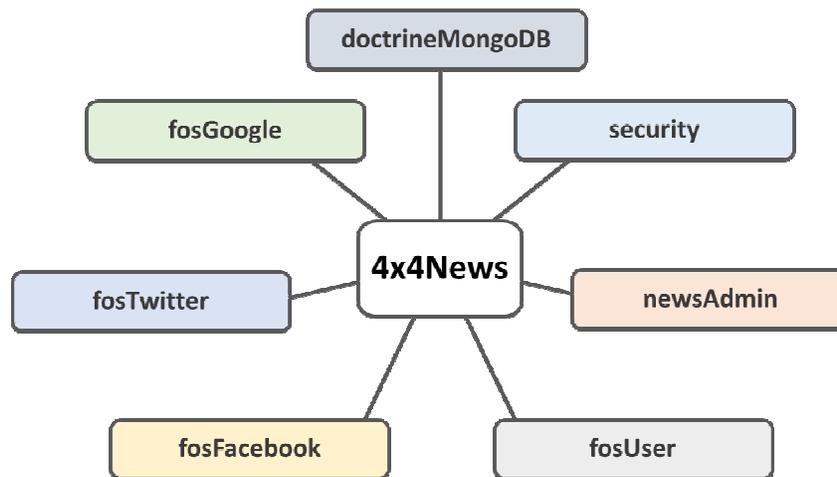
Para la presentación Symfony provee a los desarrolladores del motor de plantillas Twig que subdivide la estructura de los contenidos de cierta lógica que pueda existir a la hora de mostrar los mismos; así se evitan las recurrentes mezclas entre código HTML y PHP en la vista (Symfony 2.0, 2012). La estructura y el diseño de la aplicación se realizaron con HTML 5 (Freddy Vega, y otros, 2011), CSS 3 (Cederholm, 2010) y JavaScript, apoyados estos dos últimos por los frameworks Bootstrap 2.2.1 (Cochran, 2012) y jQuery 1.8.0 (Benedetti, y otros, 2011) respectivamente.

Debido a que la información del sistema crece aceleradamente en el tiempo y el rendimiento es vital, los datos se

encuentran alojados en el servidor de base de datos no relacional MongoDB (Chodorow, y otros, 2010). Para facilitar el acceso a los datos se utilizó el DoctrineMongoDBBundle (Symfony, 2013). La arquitectura del sistema queda explícita a través de las siguientes imágenes:



**Figura 1: Arquitectura de 4x4News**



**Figura 2: Bundles utilizados en 4x4News**

#### 2.4 RED SOCIAL DE NOTICIAS 4X4NEWS

El lector RSS 4x4News es una aplicación web con soporte de datos dedicada a recolectar titulares que se obtienen a partir de los feeds que definan sus usuarios. Los titulares se muestran en una matriz de orden cuatro con las imágenes que provee la fuente. En caso de no tener imagen se le asocia un color relacionado a la categoría a la que pertenezca la noticia y muestra el titular.

De acuerdo con el valor de relevancia calculado para cada noticia en la página inicial se muestran las primeras dieciséis más relevantes y a través de un paginado se puede acceder a los titulares más antiguos. Esta forma de mostrar el contenido ayuda al lector a captar el mensaje de que lo que está viendo es también lo más relevante para el resto de la comunidad.

Al entrar a la aplicación el usuario solamente tendrá acceso a las dieciséis noticias más relevantes de todo el sistema, véase la Figura 3: Interfaz de usuario de 4x4News. Una vez autenticado podrá utilizar el paginado para ver los titulares menos relevantes. 4x4News permite a sus usuarios crear un perfil que incluye las características de las noticias que desea ver, tal es el caso de: introducir los canales RSS que sean de su interés o seleccionar los que le propone el sistema, categorizar los canales, seleccionar las palabras claves que no deben faltar en el título o la descripción de las noticias, modificar el tiempo de actualización de la matriz, además de modificar los aspectos que puedan tomar partida en el valor de la relevancia.



**Figura 3: Interfaz de usuario de 4x4News**

Como parte de los perfiles de los miembros de la red se encuentra el “almacén” de noticias. En un “almacén” el usuario puede guardar los titulares que considere trascendental y que quiera tener al alcance de la mano en cualquier momento sin tener que hacer una búsqueda avanzada engorrosa para consultarlo nuevamente. Una vía muy eficiente para dar seguimiento un evento de cualquier índole son las “coberturas”. La “cobertura” es un mecanismo automático a través del cual el usuario recibe todas las noticias relacionadas con un mismo tema en un período de tiempo finito, tanto el tema y el tiempo los especifica el usuario. Para dar cobertura a un acontecimiento 4x4News por defecto consulta las fuentes definidas en el perfil, pero si el usuario lo desea puede indicarle al sistema que incluya en la búsqueda el resto de los canales que no están en su perfil.

En torno al lector RSS también se pueden conocer nuevas personas y 4x4News los denomina “colegas”. Los miembros de la red pueden interactuar con personas de dentro y fuera del sistema recomendándole noticias o compartiendo los titulares a través de las otras redes sociales. En el momento en que se consultan los datos de un titular puede ser comentado, votado y pueden verse tanto las noticias relacionadas como las diferentes fuentes de las que proviene. Este lector RSS tiene además un peculiar sistema de estadísticas. En el mismo se pueden acceder a datos como las fuentes más consultadas, los titulares mejores votados y más visitados.

Para integrar a 4x4News dentro de otros sitios web, se implementó un servicio que brinda la información de los cuatro titulares más relevantes publicados en la totalidad de las fuentes RSS que consulta el lector. Permite además que al usuario que se autentique en él solo se le muestren las cuatro noticias más relevantes de acuerdo con los datos de su perfil.

### 3. CONCLUSIONES

Para concebir el 4x4News fueron tomadas en cuenta las características más significativas de los principales sistemas estudiados. Para realzar los elementos visuales del titular se utilizó una arquitectura de la información bastante similar a la del TenByTen; de Google Noticias se tomó el mecanismo para configurar los feeds y compartir las noticias a través de las redes sociales; mientras que el uso de colores que hace Newsmap sirvió para darle un tratamiento diferenciado a las categorías de las noticias.

Como aspecto novedoso se tiene la aplicación de la teoría de la relevancia de (Wilson, y otros, 2004), partiendo de la premisa de que no siempre la última noticia publicada es la más relevante. Los parámetros estadísticos definidos para calcular la relevancia guardan estrecha relación con la interacción entre el 4x4News y el usuario, convirtiendo a este último en el protagonista de la información que visualiza. Por interacción se entiende a las configuraciones que realiza el usuario junto a las lecturas, votos y comentarios que reciben los titulares.

El sistema implementado constituye una vía novedosa para mantener informada a la comunidad de lectores, donde la acción coordinada de los canales RSS y los usuarios contribuyen a la difusión masiva de las noticias más interesantes.

### REFERENCIAS

- Benedetti, Ryan y Cranley, Ronan . 2011. *Head First jQuery*. s.l. : O'Reilly Media, 2011. ISBN 978-1-4493-9321-2.
- Cederholm, Dan. 2010. *CSS3 for Web Designer*. New York : s.n., 2010. ISBN 978-0-9844425-2-2.
- Chodorow, Kristina y Dirolf, E Michael. 2010. *MongoDB The Definitive Guide*. s.l. : O'Really, 2010. ISBN: 978-1-449-38156-1.
- Cochran, David. 2012. *Twitter Bootstrap Web Development (1st edition)*. 2012. ISBN 978-1849518826.
- Freddy Vega, John y Van Der Henst, Christian. 2011. *Guía de HTML. El presente de la web*. 2011.
- Gómez, Raúl. 2012. Integración de SecurityBundle de Symfony2 y FOSUserBundle. *SolucionexWeb*. [En línea] 06 de 2012. <http://blog.solucionex.com/symfony2/integracion-de-securitybundle-de-symfony2-y-fosuserbundle>.
- Google. 2013. Google Noticias. [En línea] 2013. <http://www.google.com/reader/view/?hl=es>.
- Harris, Jonathan. 2013. TenByTen. *100 Words and Pictures that Define the Time*. [En línea] 2013. <http://tenbyten.org>.
- knpbundles. 2012. FOSGoogleBundle by bitgandtter. [En línea] 2012. <http://knpbundles.com/bitgandtter/FOSGoogleBundle>.
- Newsmap. 2013. Newsmap. [En línea] 2013. <http://newsmap.jp/>.
- Symfony 2.0. 2012. *The all in one book*. s.l. : Sensio Labs, 2012.
- Symfony. 2013. The Symfony Standard Edition Bundles (2.0). *Symfony*. [En línea] 2013. <http://symfony.com/doc/2.0/bundles/index.html>.
- Symfony.es. 2013. Bundles publicados por FriendsOfSymfony. [En línea] 2013. <http://www.symfony.es/bundles/FriendsOfSymfony/>.
- Urueña, Alberto, y otros. 2011. *Las redes sociales en Internet*. s.l. : Oficina Nacional de las Telecomunicaciones y de la Seguridad Informática, Fondo Europeo de Desarrollo Regional, 2011.
- Wilson, Deirdre y Sperber, Dan. 2004. *La teoría de la relevancia*. s.l. : Revista de Investigación Lingüística, 2004.