

POLICLICK

DESARROLLO DE HERRAMIENTAS WEB 2.0 DE ANÁLISIS DE INFORMACIÓN A TRAVÉS DE FEEDS

MEMORIA DEL PROYECTO FINAL DE CARRERA

Director: Diego Álvarez Sánchez

Alumno: Alberto Pérez Fuster

Segundo cuatrimestre de 2010

1. INICIACIÓN

A modo de introducción se va a hablar de lo que son las herramientas web 2.0, del porqué de este proyecto y de las tecnologías que se han empleado.

1.1 ENUNCIADO DEL TRABAJO DEL PROYECTO



La información referente al enunciado del trabajo del proyecto ha sido proporcionada por el Observatorio de Medios.

El **Observatorio de medios de comunicación 2.0** nace con la necesidad de generar información para el análisis que no requiera una gran dedicación, en horas de trabajo efectivo, a las personas que integran esta iniciativa. El objetivo último del Observatorio no es tanto convertirse en un generador de documentos que recojan el análisis de la información generada, sino en generar información para el análisis y hacerlos públicos a través de un sitio Web de acceso libre a la sociedad en general.

Por tanto, la generación de información para el análisis debe automatizarse y sistematizarse al máximo. El sitio Web del Observatorio debe poner a disposición del equipo de trabajo del mismo, y del público en general, **información útil, veraz y contrastable**.

Las herramientas TIC seleccionadas para la implementación del sitio Web tienen que cumplir como mínimo los siguientes requisitos:

- Deben facilitar la **creación, búsqueda, sindicación y socialización**.
- Estar desarrolladas sobre Wordpress.
- Ser gratuitas y/o Open source.
- Fáciles de usar.
- No requerir instalación de software en el equipo del usuario.

El *Anexo 1* recoge los antecedentes de este proyecto, aportados por el Observatorio de Medios.

1.2 DESARROLLO HUMANO SOSTENIBLE

En esta sección se recopilan los medios analizados que han sido seleccionados por el Observatorio de Medios, con sus direcciones rss empleadas, y las palabras significativas que se han considerado para el filtrado de las noticias.

- Medios analizados y las direcciones rss.

Para la realización de las visualizaciones se han recogido las noticias de las secciones internacionales de los diarios españoles que se muestran a continuación:

	El Mundo http://estaticos.elmundo.es/elmundo/rss/internacional.xml Fecha de visita: 26 de Mayo de 2010
	El País http://www.elpais.com/rss/feed.html?feedId=1001 Fecha de visita: 26 de Mayo de 2010
	ABC http://www.abc.es/rss/feeds/abc_Internacional.xml Fecha de visita: 26 de Mayo de 2010
	La Vanguardia http://feeds.feedburner.com/lavanguardia/internacional Fecha de visita: 26 de Mayo de 2010
	Público http://www.publico.es/estaticos/rss/internacional Fecha de visita: 26 de Mayo de 2010
	Rebelión http://www.rebelion.org/rss_portada.php Fecha de visita: 26 de Mayo de 2010

Para la confección de mapas por países se han analizado las secciones internacionales de los siguientes diarios:



España:

	El Mundo http://estaticos.elmundo.es/elmundo/rss/internacional.xml Fecha de visita: 26 de Mayo de 2010
	El País http://www.elpais.com/rss/feed.html?feedId=1001 Fecha de visita: 26 de Mayo de 2010
	ABC http://www.abc.es/rss/feeds/abc_Internacional.xml Fecha de visita: 26 de Mayo de 2010
	La Vanguardia http://feeds.feedburner.com/lavanguardia/internacional Fecha de visita: 26 de Mayo de 2010




Francia:

	Le Monde http://www.lemonde.fr/rss/sequence/0,2-3210,1-0,0.xml Fecha de visita: 26 de Mayo de 2010
	Le Figaro http://rss.lefigaro.fr/figaro-international Fecha de visita: 26 de Mayo de 2010

Reino Unido:

	Times http://feeds.timesonline.co.uk/c/32313/f/440158/index.rss Fecha de visita: 26 de Mayo de 2010
	The Guardian http://feeds.guardian.co.uk/theguardian/world/rss Fecha de visita: 26 de Mayo de 2010

Estados Unidos:

	Cnn http://rss.cnn.com/rss/cnn_world.rss Fecha de visita: 26 de Mayo de 2010
	Usa Today http://rssfeeds.usatoday.com/UsatodaycomWorld-TopStories Fecha de visita: 26 de Mayo de 2010
	The Washington Post http://feeds.washingtonpost.com/wp-dyn/rss/world/index.xml Fecha de visita: 26 de Mayo de 2010

Las agencias internacionales que se han tenido en cuenta son las siguientes:

	The Associated Press http://hosted.ap.org/lineups/WORLDHEADS-rss_2.0.xml?SITE=AZPHG&SECTION=HOME Fecha de visita: 26 de Mayo de 2010
	Ansa http://www.ansa.it/web/notizie/rubriche/mondo/mondo_rss.xml Fecha de visita: 26 de Mayo de 2010
	Estadao http://www.estadao.com.br/rss/internacional.xml Fecha de visita: 26 de Mayo de 2010
	Reuters http://feeds.reuters.com/reuters/worldNews Fecha de visita: 26 de Mayo de 2010

- Palabras significativas.

Las palabras escogidas para el filtrado de noticias (por su título), en la sección de las gráficas son las siguientes:

“hambre”, “pobreza”, “guerra”, “enfermedades”, “medicamentos”, “muerte”, “áfrica”, “violencia”, “miseria”, “muertos”.

1.3 TECNOLOGÍAS INVOLUCRADAS

El termino **Web 2.0** está comúnmente asociado con un fenómeno social, basado en la interacción que se logra a partir de diferentes aplicaciones web, que facilitan el compartir información, la interoperabilidad, el diseño centrado en el usuario y la colaboración en la World Wide Web. Ejemplos de la Web 2.0 son las comunidades web, los servicios web, las aplicaciones Web, los servicios de red social, los servicios de alojamiento de videos, las wikis, blogs y mashups (mezcla de aplicaciones).

Un sitio Web 2.0 permite a sus usuarios **interactuar** con otros usuarios o cambiar contenido del sitio web, en contraste a sitios web no-interactivos donde los usuarios se limitan a la visualización pasiva de información que se les proporciona. La Web 2.0 pone a disposición de millones de personas herramientas y plataformas de fácil uso para la publicación de información en la red. De este modo, **los usuarios no sólo son consumidores sino también productores de la información.**

Las tecnologías web 2.0 en las que se va a basar este proyecto son las siguientes:

- **Wordpress.**

Wordpress es un sistema de gestión de contenido enfocado a la creación de blogs (sitios web periódicamente actualizados), desarrollado en PHP y MySQL, bajo licencia GPL. Wordpress permite la instalación de **plantillas**, ya que el contenido está separado del diseño, y ampliar sus funcionalidades mediante *widgets* (componentes) y *plugins* de todo tipo.

Hay dos modos de instalación. Uno es mediante el servicio gratuito limitado (o de pago) que ofrece Wordpress para alojar un blog. El otro es descargándose la aplicación para instalarla en un servidor externo. Esta segunda opción es la que permite una mayor configuración, y es la que se ha escogido para este proyecto.

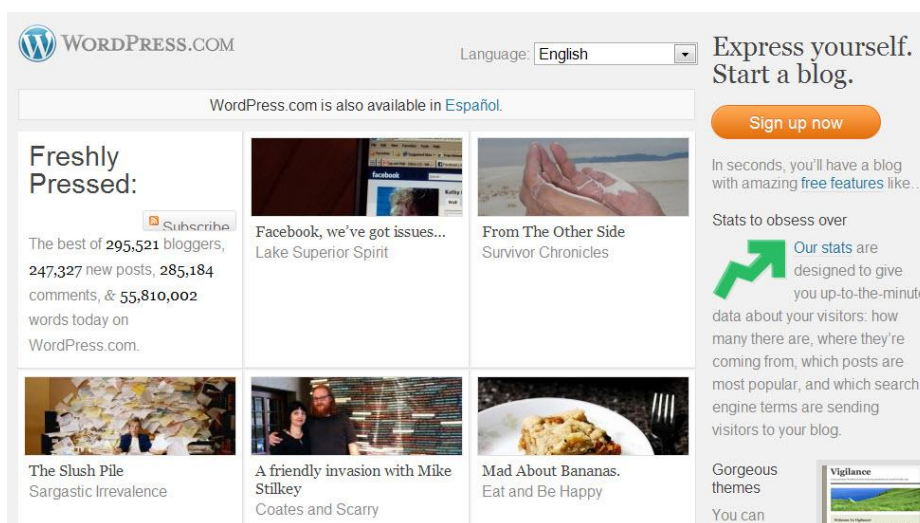


Figura 1. Wordpress.

- Yahoo Pipes.

Es una aplicación web gratuita que proporciona una interfaz gráfica de usuario para construir *mashups* (aplicaciones híbridas) capaces, por ejemplo, de leer y manipular *feeds*. Un *feed* es un archivo generado por algunos sitios web (ya sean diarios de noticias o weblogs) que contiene una versión específica de la información publicada en esa web.

En este proyecto se ha empleado Yahoo Pipes para hacer que el blog Wordpress tenga tres funciones básicas:

- 1) Generar entradas de gráficas y noticias de forma automática a partir de los feeds de diferentes diarios.
- 2) Recoger noticias de los diarios y situarlas en puntos del mapa del mundo (de acuerdo con el lugar del que están informando).

Para obtener el mapa hacemos uso del código que nos devuelve *Yahoo Badge*, una herramienta que emplea los *pipes* y javascript para utilizar las salidas de los *rss* y hacer representaciones.

- 3) Extraer los países de los que informan las noticias y su número de repeticiones, con el objetivo de generar una tabla mediante php que pueda ser exportada a la web *Many Eyes*.

The screenshot shows the Yahoo Pipes website interface. At the top, there's a navigation bar with the 'pipes' logo and links for Home, My Pipes, Browse, Discuss, Documentation, and Create a pipe. A user is logged in as 'observatorio10'. The main content area features a 'Featured Pipe: Upcoming.org Combined Feed' with a colorful abstract background. A tooltip over the pipe explains it's a combination of individual events, friends' events, site activity, and site news. Below this, there are links for 'Craigslist and Yahoo! Pipes - resolved', 'Connection refused and other Pipes issues', and 'Craigslist and Yahoo! Pipes'. The 'About Pipes' section describes it as a composition tool for aggregating and manipulating content. A 'Hot Pipes' section lists several popular pipes: 'Flickr: Comments on my photos by other people' by bitrot, 'Yahoo Finance Stock Quote Watch List Feed w/Chart' by Paul Donnelly, and 'US population by state' by Jonathan. A 'Learn More about Pipes' link points to the Pipes Blog.

Figura 2. Yahoo Pipes.

- Many Eyes.

Many Eyes (<http://manyeyes.alphaworks.ibm.com/manyeyes>) es una web patrocinada por IBM que sirve para analizar visualmente diversos *data sets* (conjuntos de datos) en forma de

visualizaciones (gráficos), de los cuales hay muchos disponibles en la web. Los usuarios pueden subir *data sets*, editarlos y compartirlos, además de utilizar los ya creados por otros usuarios y conversar sobre las informaciones visuales obtenidas.

Este proyecto intenta hacer de enlace con *Many Eyes*, creando tablas de datos mediante php (a partir de los resultados de los *pipes*), que sean fácilmente exportables a dicha web, para crear gráficos que pongan de manifiesto la cobertura que hacen las agencias internacionales sobre la pobreza y las desigualdades norte-sur.

1.4 METODOLOGÍA DE DESARROLLO

Durante la consecución de este proyecto se ha seguido una **metodología ágil**, por ser más idónea para el desarrollo de este producto que una secuencial o tradicional.

Las metodologías ágiles valoran:

- Al individuo y las interacciones en el equipo de desarrollo más que a las actividades y las herramientas.
- Desarrollar software que funciona más que conseguir una buena documentación, es decir, el minimalismo respecto de la documentación. Por este motivo, el tiempo de respuesta a los cambios será menor.
- La colaboración con el cliente más que la negociación de un contrato.
- Responder a los cambios más que seguir estrictamente una planificación.

Principios de las metodologías ágiles:

1. La prioridad principal es satisfacer al cliente mediante **tempranas y continuas entregas de software**.
2. Dar **bienvenida a los cambios**. Los procesos ágiles aplican los cambios para que el cliente sea competitivo.
3. **Buena comunicación con el cliente**, quien forma parte activamente del proceso de desarrollo.
4. La **simplicidad** es esencial.
5. La planificación es muy importante, pero no se siguen estrictamente unas pautas.

Conviene aclarar que los requisitos para este proyecto estaban bien claros y definidos desde el principio, por lo que en este aspecto no iban a haber cambios, pero sí, presumiblemente, en cómo iban a desarrollarse las funcionalidades.

2. PLANIFICACIÓN

En este segundo apartado se va a abordar, en primer lugar, la definición del proyecto, objetivos y el ámbito en el que se engloba; y en segundo lugar, el desarrollo del plan para lograr los objetivos.

2.1 Documento de definición del proyecto:

2.1.1 OBJETIVOS DEL PROYECTO

- Ofrecer una herramienta web 2.0 a las asociaciones sin ánimo de lucro, que dirigen campañas de sensibilización y movilización a la comunidad universitaria, que ofrezca información útil sobre la visión del mundo que tienen los medios de comunicación que utilizan para informarse.
- Cumplir los requisitos marcados por la convocatoria de ayudas del Centro de Cooperación al Desarrollo (CCD).

2.1.2 CONTEXTO DEL PROYECTO

El proyecto se enmarca en el contexto de la línea de actuación del CCD en materia de sensibilización y educación para el desarrollo.

- **OBJETO:**
Apoyar actividades realizadas por miembros de la comunidad universitaria cuyo objetivo sea **fomentar los valores de solidaridad en la universidad**, así como sensibilizar a la sociedad sobre las causas de la pobreza y las **desigualdades Norte-Sur**.
- **CONDICIONES:**
Con el objetivo de fortalecer el compromiso solidario de la Universidad en la erradicación de la pobreza y el conocimiento de las desigualdades entre los países enriquecidos y los empobrecidos, esta convocatoria de ayudas pretende apoyar **actividades de sensibilización, educación para el desarrollo** y promoción de la participación social.

Requisitos específicos:

1. Una persona sólo podrá presentar una solicitud.
2. El **importe máximo** por acción no podrá exceder un total de **2.000 euros**.
3. En caso de percibir otras ayudas para la misma acción se deberá informar a la mayor brevedad posible al CCD.

2.1.3 DEPENDENCIAS DEL PROYECTO

1. La solución ICT será alojada en un servidor propiedad de la UPV. Deben tenerse en cuenta las políticas de seguridad y restricciones impuestas por la UPV.
2. La solución se implementará en Wordpress: <http://es.wordpress.com>
3. Las herramientas de producción deben de estar basadas en software libre.
4. El importe máximo por acción no podrá exceder un total de 2.000 euros.
5. Este proyecto va dirigido a la comunidad universitaria de la UPV.
6. La solución deberá mostrar la visión que tienen los medios de la pobreza y las desigualdades norte-sur.

2.1.4 ESPECIFICACIÓN DEL PRODUCTO

a) Usuarios.

- Grupos de usuarios.

Responsables de la xarxa UPV de organizar y dirigir campañas de sensibilización y movilización hacia la comunidad universitaria a través de las asociaciones sin ánimo de lucro.

- Primary Persona.
 - Víctor, de veinticinco años de edad, es estudiante de 3º de Arquitectura en la Universidad Politécnica de Valencia, además de responsable de organizar campañas de sensibilización y movilización hacia la comunidad universitaria de una organización sin ánimo de lucro llamada Sociedad y Desigualdad.
 - Escribe y habla castellano y valenciano con fluidez y tiene un nivel de habla y escritura del inglés medio.
 - Su horario de conexión a internet suele ser tardes después de las clases y fines de semana en diferentes horarios dependiendo de cuando este en casa.
 - Esta ampliamente familiarizado con el uso de los sistemas informáticos entre los que se encuentran ordenadores, móviles, pda's...etc. Está muy interesado en las nuevas tecnologías, y las tecnologías de la información y la comunicación (TIC).
 - Sus objetivos son concienciar y sensibilizar a la gente mediante campañas de información, sobre diversos temas.
 - En cuanto a los objetivos que desea alcanzar con el sitio web, Víctor desea disponer de información contrastada, cuantificada, completa y de una forma fácil (sin tener que buscar en demasiados sitios) que le ayude a orientar las campañas de concienciación y sensibilización.
 - Víctor apreciará un diseño sencillo y estructurado de la información, además de una interfaz agradable a la vista y acorde con la temática, y le molestaría que la información no estuviese contrastada (es decir, que por ejemplo solo publicarían información de una sola fuente) y que los contenidos no se actualizaran con frecuencia.

b) **Objetivos.**

- **Objetivos de la solución Web:**
 - La información relativa a la visión del mundo que ofrecen los medios de comunicación se genere al 80% de forma sistemática y automática.
 - La información contenida en el sitio Web se actualice diariamente de forma automática o semiautomática.
 - Lograr en el plazo de 3 meses que el 100% de las asociaciones sin ánimo de lucro, que dirigen campañas de sensibilización y movilización a la comunidad universitaria, esté suscrita al Observatorio de medios de comunicación 2.0
 - Contrastar noticias de distintas fuentes y mostrar la información, para dar la visión que pretende ofrecer el observatorio.
- **Objetivos del CCD para el Observatorio de medios de comunicación 2.0:**

Que el 15% de los miembros de la comunidad universitaria de la UPV conozcan el proyecto en el plazo de 6 meses.

- **Objetivos de los usuarios:**

Obtener información concreta y contrastable sobre la visión que dan los medios de comunicación sobre la pobreza y las desigualdades.

c) **Contenidos y funcionalidades.**

El blog está dividido en cuatro secciones (páginas):

1. **Home.**

Esta sección contiene las entradas con las **noticias** filtradas por palabras de los diarios y su **gráfica de barras** correspondiente.

Cada día, a las 23:50 se genera una entrada por cada noticia de los diarios (filtrada por palabras tales como hambre, pobreza, muerte...) y otra entrada con la gráfica que representa el número de noticias que haya habido ese día para cada diario.

Al pinchar sobre la cabecera de cada noticia se puede acceder a la fuente original, para su consulta.

Al comienzo de la página hay una entrada fija en la que se explica brevemente las gráficas y su funcionamiento.

Para la realización de las gráficas de barras se han recogido las noticias de las secciones internacionales de los diarios españoles que se muestran a continuación:

	El Mundo http://estaticos.elmundo.es/elmundo/rss/internacional.xml Fecha de visita: 26 de Mayo de 2010
	El País http://www.elpais.com/rss/feed.html?feedId=1001 Fecha de visita: 26 de Mayo de 2010

	ABC http://www.abc.es/rss/feeds/abc_Internacional.xml Fecha de visita: 26 de Mayo de 2010
	La Vanguardia http://feeds.feedburner.com/lavanguardia/internacional Fecha de visita: 26 de Mayo de 2010
	Público http://www.publico.es/estaticos/rss/internacional Fecha de visita: 26 de Mayo de 2010
	Rebelión http://www.rebelion.org/rss_portada.php Fecha de visita: 26 de Mayo de 2010

2. Mapas de noticias.

Esta sección está compuesta por seis mapas. Cada uno de ellos corresponde a un país (España, Francia, Reino Unido, Alemania, Estados Unidos y Canadá), en donde se geolocalizan las noticias de la sección internacional de sus diarios. De este modo se puede ver fácilmente qué partes del mundo se están cubriendo y cuáles quedan invisibles.

También hay al principio de la página una explicación de los mapas y su funcionamiento.

Para la confección de mapas por países se han analizado las secciones internacionales de los siguientes diarios:

España:

	El Mundo http://estaticos.elmundo.es/elmundo/rss/internacional.xml Fecha de visita: 26 de Mayo de 2010
	El País http://www.elpais.com/rss/feed.html?feedId=1001 Fecha de visita: 26 de Mayo de 2010
	ABC http://www.abc.es/rss/feeds/abc_Internacional.xml Fecha de visita: 26 de Mayo de 2010
	La Vanguardia http://feeds.feedburner.com/lavanguardia/internacional Fecha de visita: 26 de Mayo de 2010




Francia:

	Le Monde http://www.lemonde.fr/rss/sequence/0,2-3210,1-0,0.xml Fecha de visita: 26 de Mayo de 2010
	Le Figaro http://rss.lefigaro.fr/figaro-international Fecha de visita: 26 de Mayo de 2010

Reino Unido:

	Times http://feeds.timesonline.co.uk/c/32313/f/440158/index.rss Fecha de visita: 26 de Mayo de 2010
	The Guardian http://feeds.guardian.co.uk/theguardian/world/rss Fecha de visita: 26 de Mayo de 2010

Estados Unidos:

	Cnn http://rss.cnn.com/rss/cnn_world.rss Fecha de visita: 26 de Mayo de 2010
	Usa Today http://rssfeeds.usatoday.com/UsatodaycomWorld-TopStories Fecha de visita: 26 de Mayo de 2010
	The Washington Post http://feeds.washingtonpost.com/wp-dyn/rss/world/index_xml Fecha de visita: 26 de Mayo de 2010

3. Tablas de datos

Esta sección genera unas tablas de datos que muestran para diferentes agencias internacionales de noticias el número de noticias por país. Esta información está pensada para ser exportada a la web Many Eyes para su representación.

<http://manyeyes.alphaworks.ibm.com/manyeyes>



Figura 3. Many Eyes, página de inicio.

Al principio de esta sección hay una breve explicación de la funcionalidad.

Las agencias internacionales que se han tenido en cuenta para esta sección son las siguientes:

	<p>The Associated Press http://hosted.ap.org/lineups/WORLDHEAD_S-rss_2.0.xml?SITE=AZPHG&SECTION=HOME Fecha de visita: 26 de Mayo de 2010</p>
	<p>Ansa http://www.ansa.it/web/notizie/rubriche/mondo/mondo_rss.xml Fecha de visita: 26 de Mayo de 2010</p>
	<p>Estadao http://www.estadao.com.br/rss/internacional.xml Fecha de visita: 26 de Mayo de 2010</p>
	<p>Reuters http://feeds.reuters.com/reuters/worldNews Fecha de visita: 26 de Mayo de 2010</p>

4. **Acerca de.**

Esta página informa sobre este proyecto y su propósito.

El blog además dispone de una barra lateral con el archivo de noticias para cada mes, las entradas recientes y las herramientas de administración.

La arquitectura de la información (parte pública) es la que se representa en el siguiente gráfico:

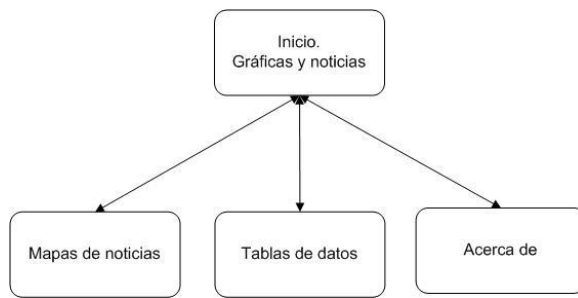


Figura 4. Secciones del blog.

2.2 Planificación:

A continuación se presentan las fases y las tareas necesarias para el desarrollo de la solución del proyecto.

FASE 1. PUESTA EN MARCHA

- 1.1 Creación de cuenta en Yahoo Pipes (0 días).**
Se necesita una cuenta en Yahoo Pipes para poder crear los pipes, pues estos son la base del funcionamiento de este proyecto.
- 1.2 Hacer pruebas con los Pipes (2 días).**
Se realizarán unas pruebas unitarias de los módulos principales que vamos a utilizar.
- 1.3 Búsqueda de alojamiento web y creación de cuenta de hosting (1 día).**
Hay que elegir un servidor gratuito y que sea adecuado para alojar el blog Wordpress.
- 1.4 Instalar Wordpress (0 días)**
Se incluirá los ficheros de Wordpress dentro de espacio ofrecido por el servidor de hosting y se creará una cuenta para su administración.
- 1.5 Búsqueda de plugins que hagan falta (2 días).**
Se necesitan plugins para el blog que permitan tanto su alimentación mediante *rss* como soporte *php*, entre otras cosas.
- 1.6 Instalación y pruebas de plugins en Wordpress (1 día).**
Una vez obtenidos los plugins, se instalan en Wordpress.
- 1.7 Configurar la parte visual de Wordpress (1 día).**
Instalación de temas, configuración de widgets y demás personalizaciones.
- 1.8 Crear cuenta en Many Eyes (0 días).**
Se necesita esta cuenta para exportar las tablas obtenidas en el blog y crear diferentes visualizaciones en esta web.
- 1.9 Redacción de la sección “Acerca de..” (1 día).**
En esta página se blog se hablará de la organización en la cual se enmarca este proyecto, y sus objetivos.

FASE 2. ELABORAR SECCIÓN GRÁFICAS

- 2.1 Configuración de plugin *FeedWordpress* (1 día).
Elegir el tipo de actualización, entre otras configuraciones.
- 2.2 Configuración de plugin *FWP Limit Size of Posts* (1 día).
Determinar el tamaño en palabras de las entradas.
- 2.3 Elaboración de los *Pipes* (2 días).
Crear en Yahoo Pipes los *pipes* necesarios para la obtención de las noticias de los diarios y la generación de las gráficas a partir de ellas.
- 2.4 Incluir entrada fija con explicación (1 día).
Una entrada al inicio de la página que explique las gráficas.
- 2.5 Configuración de *Cron* (1 día).
Crear una tarea *cron* en el servidor para que todos los días a la hora establecida actualice las entradas del blog con nuevas noticias y gráficas.

FASE 3. ELABORAR SECCIÓN MAPAS

- 3.1 Elaboración de los Pipes (2 días).
Creación de los pipes necesarios para geolocalizar las noticias de los diarios.
- 3.2 Incluir código Badge en página (1 día).
Obtener el código script de los mapas resultantes de ese pipe e integrarlos en la página.
- 3.3 Incluir explicación en la página (1 día).
Breve mensaje explicativo al comienzo de la página de los mapas.

FASE 4. ELABORAR SECCIÓN TABLAS

- 4.1 Elaboración de los Pipes (2 días).
Se creará un pipe para cada diario (agencia internacional) que extraerá los países y el número de repeticiones de cada uno de ellos, para confeccionar una tabla.
- 4.2 Configuración de plugin *Exec-PHP* (1 día).
Este plugin es necesario para dar soporte php a Wordpress.
- 4.3 Incluir código php y explicación en la página (1 día).
Se incluirá una breve explicación, además del código en la página “Tablas de datos del blog”.
- 4.4 Código php para recorrer los feeds y dibujar las tablas (1 día).
El contenido de los rss de las noticias tienen que ser leídos y escritos en la página del blog mediante código php para crear las tablas.

FASE 5. PRUEBAS

- 5.1 Pruebas de la sección de Gráficas (5 días).
Pasos a seguir para probar la actualización del blog con una nueva gráfica y sus noticias.
- 5.2 Pruebas de la sección Mapas (2 días).
Comprobar que se cargan los mapas.

5.3 Pruebas sobre *Many Eyes* (1 día).

Pasos a seguir para crear una nueva tabla de datos en la web *Many Eyes*, copiando una tabla del blog, y para obtener diferentes visualizaciones.

A continuación se presenta un gráfico que representa el proceso de desarrollo del proyecto:

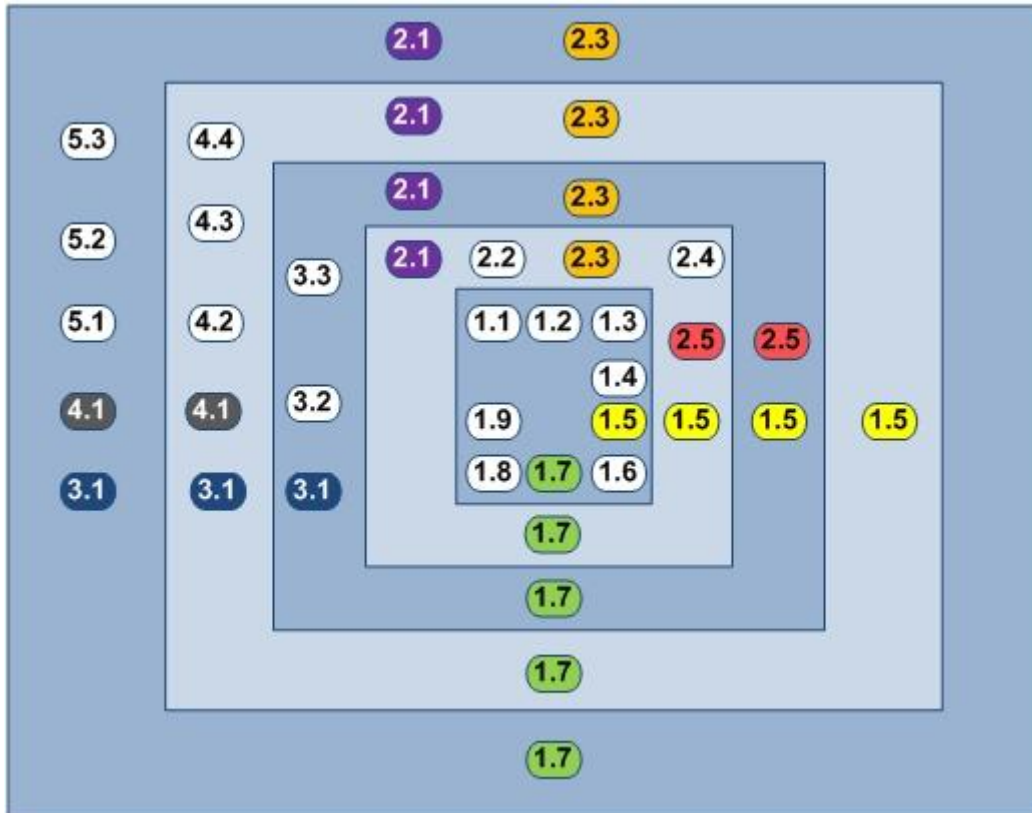


Figura 5. Proceso de desarrollo del proyecto.

La figura anterior muestra gráficamente el proceso de **desarrollo ágil** que se ha seguido para realizar este proyecto. Cada recuadro se corresponde con un nuevo **incremento** en el proceso, y dentro de cada uno las tareas que se añaden al mismo. Los números son los de las tareas que se encuentran en las cinco fases anteriormente expuestas en la planificación.

Como puede observarse, hay tareas que se van revisando en los diferentes incrementos del proceso de desarrollo. Estas tienen que ver principalmente con procesos de elaboración de pipes, pero también con la búsqueda y configuración de pipes y del Cron.

3. EJECUCIÓN

Los resultados obtenidos de la ejecución de las fases y tareas propuestas en la planificación son los siguientes:

FASE 1. PUESTA EN MARCHA

1.1 Creación de cuenta en Yahoo Pipes.

Por razones de seguridad y confidencialidad no se incluyen los datos de la cuenta.

1.2 Hacer pruebas con los *Pipes*.

Creamos un *pipe* para probar los módulos más importantes que vamos a emplear en el proyecto. A continuación se propone un ejemplo:

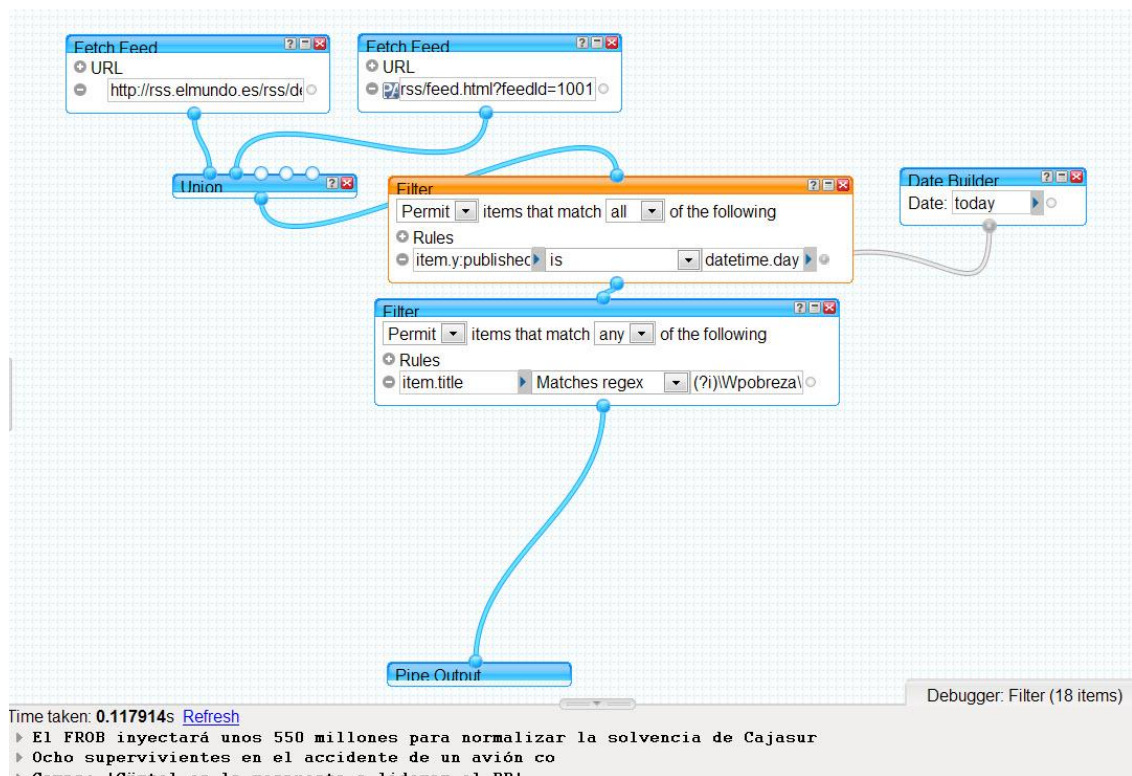


Figura 6. Ejemplo de pipe.

Pasos para la creación de este pipe de ejemplo:

1. Se elige el módulo *Fetch Feed* del panel de la izquierda (en *Sources*) y se arrastra al área de diseño del pipe. En su campo URL se incluye el rss de un feed.
2. Repetimos nuevamente el paso 1 con otro *Fetch Feed*.
3. Se crea un módulo *Union* (panel de la izquierda, dentro de *Operators*), y se conecta la salida de los dos módulos anteriores con la entrada de éste.

4. Se seleccionan un módulo *Filter*, a cuya entrada le conectamos la salida del módulo *Union*. Se seleccionan los valores que aparecen en la imagen.
5. Se crea un módulo *Date Builder*, y en su campo “Date” le asignamos el valor *today*. La salida de este módulo se conecta con el último campo del módulo *Filter* anterior. En las opciones que se despliegan en dicho campo se elige *datetime.day*.
6. Se crea otro módulo *Filter*, conectado con el anterior y se le incluyen los valores que aparecen en la imagen. En el último campo se escribe lo siguiente: *(?i)\Wpobreza\W*, para que busque esa palabra completa.

Descripción de los módulos que aparecen (y de otros que no, pero que también se usarán en el proyecto):

- *Fetch Feed*: módulo que recibe la url de un *feed rss* y devuelve los valores que contiene.
- *Filter*: módulo para filtrar los ítems de acuerdo a unas reglas como puede ser que deje pasar noticias que contengan en el título determinadas palabras, o cuya fecha de publicación coincida con una dada.
- *Auto-Geo*: Pipe que geolocaliza las noticias de un feed.
- *Google Bar Charts*. Pipe que crea una gráfica dados unos valores numéricos.
- *Date Builder*: Módulo para construir un elemento fecha. Se le pueden pasar valores como *today*, *yesterday*, *-1days*...
- *Union*. Módulo que combina hasta cinco entradas.
- *Split*. Módulo contrario a *Union*. Hace dos copias de una misma entrada.

Para ver los resultados que devuelve cada módulo basta con pulsar sobre ellos y en la ventana de abajo se mostrarán (en este ejemplo se ha pulsado sobre el primer módulo *Filter*).

```

Time taken: 0.05398s Refresh
▶ Netanyahu: 'No pediremos disculpas'
▶ UGT y CC00 negociarían la Ley de Huelga 'en frío'
▶ Videla otra vez en el banquillo
▶ CiU denuncia posible financiación irregular tripar
▶ El Gobierno sienta las bases para cerrar las líneas de tren no rentables
▶ Primeras retenciones en las carreteras
▶ El buen tiempo se toma unas vacaciones
▶ El Estado tendrá tendrá 10.600 empleos menos en 2013
+9 more...

```

Figura 7. Resultados de un módulo del pipe.

1.3 Búsqueda de alojamiento web y creación de cuenta de hosting.

Para la búsqueda del servicio de hosting se han aplicado los siguientes criterios:

1. Acceso ftp.
2. Buen espacio para alojar datos.
3. MySQL (necesario para alojar el blog Wordpress)
4. Servicio de Cron Jobs
5. Que sea gratuito.

El hosting seleccionado es 000webhost.com, por la experiencia de proyectos pasados (en la asignatura de IPM) y porque cumple con los criterios mencionados.

Por razones de seguridad y confidencialidad no se incluyen los datos de la cuenta ftp ni del “cPanel” del hosting.

000webhost.com (\$0.00 webhost), is an industry leader in providing top class free web hosting services without advertising! There are no hidden costs, no adverts, and no restrictive terms. Lightning fast speeds, maximum reliability and fanatical user support are just a few of the features you'll receive with our service.

We are probably **the only** free web hosting company that has **proof** for "99% uptime guarantee". Check our uptime stats. We managed to reach even 99.9% uptime for most of the servers. Can your current hosting provider show you any proof for that "99.9% uptime" they guarantee?

Free or paid hosting? Compare our plans:

- Home
- View All Features
- Free Domain Hosting
- Free PHP Hosting
- Free Hosting with MySQL
- Free cPanel Hosting
- Website Builder
- Fantastico Autoinstaller
- Frequent Questions
- Web Hosting Reviews
- Website Templates
- Affiliate Program

Figura 8. Servidor de alojamiento 000webhost.com

1.4 Instalar Wordpress.

Para la instalación de Wordpress se ha seguido la siguiente guía que explica el proceso paso a paso: <http://ayudawordpress.com/como-instalar-wordpress-paso-a-paso/>

Fecha de visita de la web: 1 de Julio de 2010

1.5 Búsqueda de plugins que hagan falta.

Los plugins elegidos para el proyecto son los siguientes:

- *FeedWordPress*, para la actualización del blog a través de feeds.
- *Exec-Php*, para dar soporte php a las páginas de Wordpress.
- *FWP Limit Size of Posts*, añadido de *FeedWordPress* para limitar el tamaño de las entradas.

1.6 Instalación y pruebas de plugins en Wordpress.

En los apartados 1.1, 2.2 y 4.2 de la fases de Ejecución se detallan los pasos para instalar y configurar cada uno de los plugins.

1.7 Configurar la parte visual de Wordpress

Para el aspecto visual del blog se ha optado por un tema llamado “9ths current” (<http://wordpress.org/extend/themes/9ths-current>), con una sola barra lateral derecha. Los widgets usados son: Categorías, Archivos, Entradas recientes y Meta.

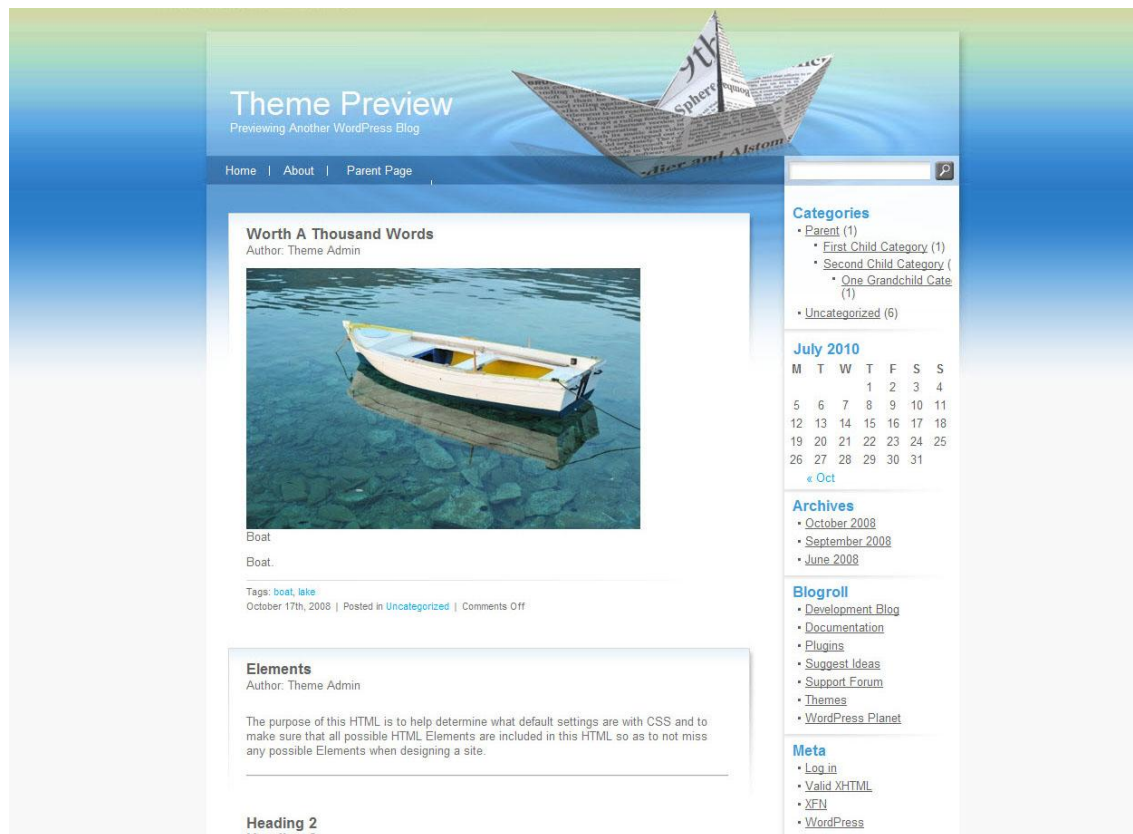


Figura 9. 9ths current Preview

1.8 Crear cuenta en *Many Eyes*.

Por razones de seguridad y confidencialidad no se incluyen los datos de la cuenta.

1.9 Redacción de la sección “Acerca de..”

Texto:

Este blog forma parte de un Proyecto Final de Carrera que está orientado a la sensibilización y educación para el desarrollo del Centro de Cooperación al Desarrollo. Mediante el seguimiento de los medios de comunicación se pretende ofrecer a las asociaciones sin ánimo de lucro, que dirigen campañas de sensibilización y movilización a la comunidad universitaria, información útil sobre la visión del mundo.

FASE 2. ELABORAR SECCIÓN GRÁFICAS

2.1 Configuración de plugin *FeedWordPress*.

Explicación:

FeedWordPress es un plugin que permite que el blog Wordpress recoja el contenido de las direcciones rss (o feeds) que se le pasen para publicar entradas automáticamente o semi-automáticamente con dicha información. Es decir, para “alimentar” el blog con noticias de otros sitios web.

Para configurar este plugin se deben seguir los siguientes pasos:


Instalación:

1. Descomprimir el archivo y subirlo a la cuenta ftp.
2. Sustituir los archivos rss.php y rss-functions.php que se encuentran en el directorio wp-includes por los que vienen en la carpeta MagpieRSS de FeedWordPress.
3. Activar el plugin FeedWordPress dentro de la sección Plugins – Instalados del Escritorio de Wordpress.

Configuración:

Dentro de Syndication. Sección Syndication:

En “Add new Source” hay que pegar la dirección rss del pipe “gráfica+noticias” y pulsar sobre el botón “Syndication”. De este modo el plugin actualizará, cuando toque, el blog con entradas que se corresponderán con los elementos que obtenga del pipe.

 *Syndicated Sites Settings*

Update feeds now

Check currently scheduled feeds for new and updated posts.

Note: Automatic updates are currently turned **off**. New posts from your feeds will not be syndicated until you manually check for them here. You can turn on automatic updates under [Feed & Update Settings](#).

[Update](#)

Keep FeedWordPress improving

FeedWordPress makes syndication simple and empowers to stream content from all over the web into your WordPress hub. That's got to be worth a few lattes. If you're finding FWP useful, [a modest gift](#) is the best way to support steady progress on development, enhancements, support, and documentation.

[Donate](#)

You can make a gift online (or [set up an automatic regular donation](#)) using an existing PayPal account or any major credit card.

[Maybe Later](#) [Dismiss](#)

Syndicated sources

[Update Checked](#) [Unsubscribe](#) Add new source: [Syndicate](#)


<input type="checkbox"/>	Nombre	Feed	Updated
<input type="checkbox"/>	grafica+noticias	pipes.yahoo.com/pipes/pipe.run?_...	18 horas ago Ready for next update 17 ago

Figura 10. FeedWordPress, sección Syndication.

Sección Feeds & Updates:

En “Updates” hay que seleccionar la opción “cron job or manual updates”, porque la opción que trae el plugin de actualizaciones automáticas no nos interesa, ya que esto se encargará de hacerlo el Cron del servidor, que llamará, mediante un fichero, al update del plugin cada día a las 23:50.

El campo “Wait” de “Update scheduling” lo dejamos como está, en 60 minutos, que es el tiempo mínimo recomendado, para que no haya exceso de actividad en el servidor.

 **Feed and Update Settings**

These settings affect posts syndicated from any feed unless they are overridden by settings for that specific feed. - defaults for all feeds - ▾

Feed Information

Syndicated Link category: ▾
 FeedWordPress will syndicate the links placed under this link category.

Link Names: Update contributor titles automatically when the feed title changes

Short descriptions: Update contributor descriptions automatically when the feed tagline changes

Homepages: Update contributor homepages automatically when the feed link changes

Update Scheduling

Updates: ▾

If you want to use a cron job, you can perform scheduled updates by sending regularly-scheduled requests to http://observamedios.hostoi.com?update_feedwordpress=1. For example, inserting the following line in your crontab:

```
*/10 * * * * /usr/bin/curl --silent http://observamedios.hostoi.com?update_feedwordpress=1
```

will check in every 10 minutes and check for updates on any feeds that are ready to be polled for updates.

Update scheduling: How long should FeedWordPress wait between updates before it considers a feed ready to be polled for updates again?

Wait minutes between polling.

Recommendation. Unless you are positive that you have the webmaster's permission, you generally should not set FeedWordPress to poll feeds more frequently than once every 60 minutes. Many webmasters consider more frequent automated polling to be abusive, and may

Figura 11. FeedWordPress, sección *Feeds & Updates*.

En el campo “Time limit on updates” 86340 segundos, que son 24 horas menos 1 minuto. Esta es una función del plugin que descarta las noticias que cuya fecha de publicación sea anterior a un tiempo dado, pero en realidad esto no nos hace falta, ya que en el pipe ya controlamos que no se devuelvan noticias anteriores a las últimas 24 horas.

En “Updated posts” marcamos la opción “No, leave the syndicated copy unmodified”, ya que no queremos que el plugin actualice entradas que ya fueron publicadas.

complain to your web host, or ban your IP address, as retaliation for hammering their servers too hard.

Note. This is a default setting that FeedWordPress uses to schedule updates when the feed does not provide any scheduling requests. If a feed does provide update scheduling information (through elements such as `<nss:ttl>` or `<sy:updateFrequency>`), FeedWordPress will respect the feed's request.

Time limit on updates: limit updates to no more than...
86340 seconds

Updated Posts

Updated posts: When a syndicated feed includes updated content for a post that was already syndicated, should the syndicated copy of the post be updated to match the revised version?

Yes, update the syndicated copy to match

No, leave the syndicated copy unmodified

Figura 12. FeedWordPress, Sección Feeds & Updates.

Sección Posts & Links:

En el campo “Permalinks” escogemos la opción “point to the copy on the original website”, para que el pinchar sobre el título de la noticia se acceda a la web original de la noticia. Esto se hace para cumplir uno de los objetivos del proyecto, que es el de contrastar la información.

Figura 13. FeedWordPress, sección Posts & Links.



Figura 14. FeedWordpress, sección Posts & Links 2.

2.2 Configuración de plugin *FWP Limit Size of Posts*.

Explicación:

Este plugin se utiliza para limitar el número de palabras o de caracteres (según se prefiera) que van a tener las entradas del blog generadas por el *FeedWordpress*.

Se deben seguir los siguientes pasos para configurar este plugin:

Instalación:

1. Descomprimir el archivo y subirlo a la cuenta ftp.
2. Activar el plugin dentro de la sección Plugins – Instalados del Escritorio de Wordpress.

Configuración:

3. Ir a la sección *Syndication - Posts & Links* y al final de la página:
Marcamos la opción “Limit the size of syndicated posts”, y ponemos 20 words. Así limitamos el contenido de cada entrada a 20 palabras.

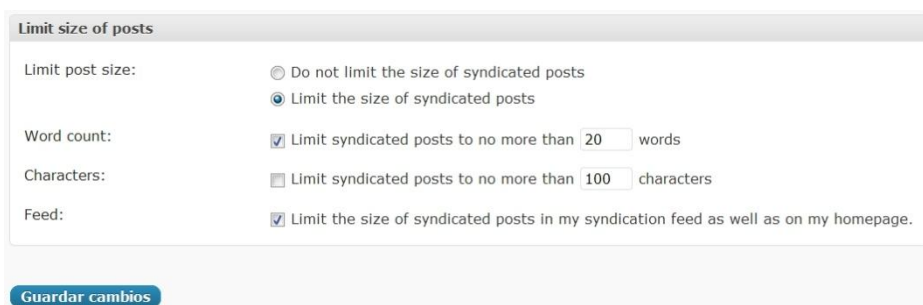


Figura 15. FWP Limit Size of Posts, configuración.

2.3 Elaboración de los Pipes.

Rss Filtrados.

Este Pipe se encarga de coger el Feed de un diario y filtrar sus noticias. En primer lugar escoge las noticias cuya fecha de publicación coincide con la del día actual. En segundo lugar, deja pasar a las noticias cuyo título contenga alguna de las siguientes palabras completas: hambre, pobreza, guerra, enfermedades, medicamentos, muerte, África, violencia, miseria y muertos.

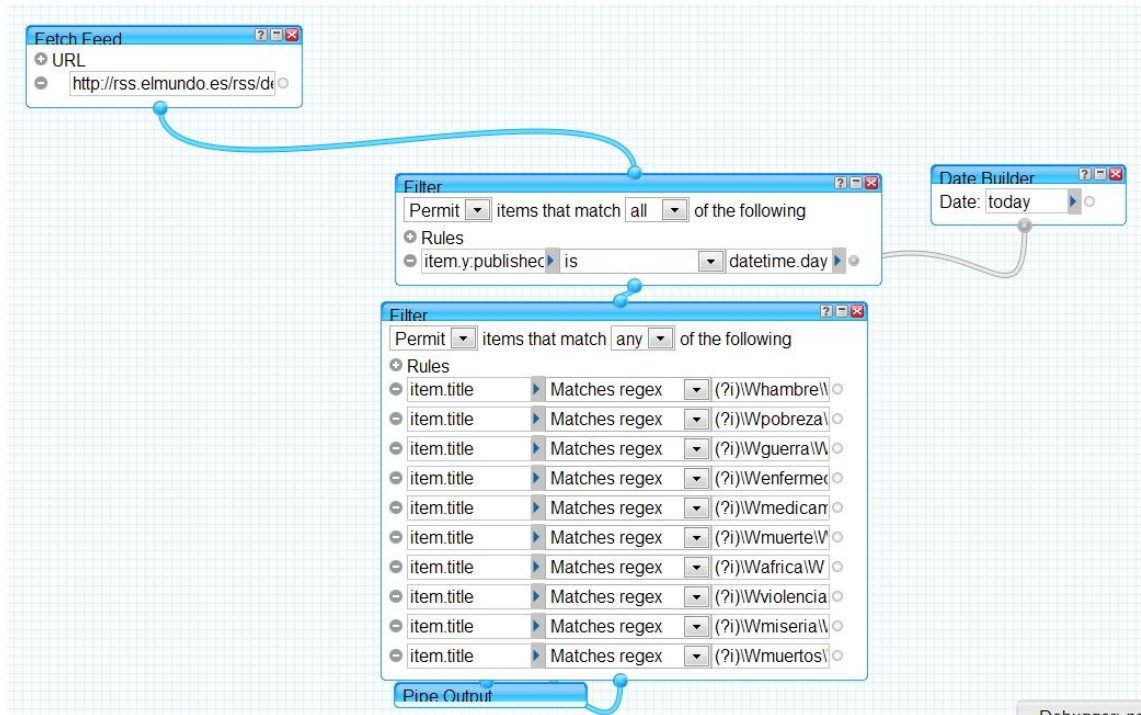


Figura 16. Pipe Rss Filtrados. Filtra las noticias por palabras.

Nota: Se hacen seis pipes como éste, uno por cada diario de España.

Pasos para su creación:

1. Se elije el módulo *Fetch Feed*, arrastrándolo al área de creación de pipes y se incluye en su campo la dirección *rss* del diario.
2. Se selecciona el módulo *Filter* y se conecta con el módulo anterior, y a continuación se rellena con los datos que aparecen en la imagen.

Importante: Estos datos se seleccionan desplegando las opciones de cada campo, que dependen de los ítems que recibe por la entrada, procedentes del módulo anterior. Por ejemplo, si en el *feed rss* del mundo no hubiera elementos con el campo *published* no se podría elegir el valor *item.y:published* del módulo *Filter*.

3. Se elije el módulo *DateBuilder* y en su campo escribimos “today”, y lo conectamos con el campo que tiene en gris el módulo anterior (al desplegar los posibles valores para este campo del módulo *Filter* se selecciona *datetime.day*).

- Se elije el módulo *Filter*, se conecta su entrada con la salida del anterior *Filter* y rellenamos sus valores con los de la imagen. Cada una de las palabras del tercer campo van incluidas entre $(?)\backslash W$ y $\backslash W$. Por ejemplo $(?)\backslash Whambre\backslash W$
- La salida de este módulo *Filter* se conecta con el *Pipe Output*, que dará los resultados del pipe.

Gráfica+noticias.

Este pipe hace uso del pipe anterior para filtrar las noticias de los feeds de seis diarios españoles y con ellas hace dos operaciones, tomando dos caminos mediante el módulo *Split*. En una, simplemente engloba todas las noticias para que sean devueltas en la salida del pipe. El plugin *FeedWordpress* hará una entrada de cada una de estas noticias. En la otra, se hará un recuento mediante el módulo *Count* de las noticias de cada uno de los diarios, y se pasará estos datos al módulo *Google Bar Charts*, que se encarga de dibujar la gráfica comparativa. Además, añadirá mediante los módulos *Date Builder* y *Date Formatter* la fecha al título de la gráfica.

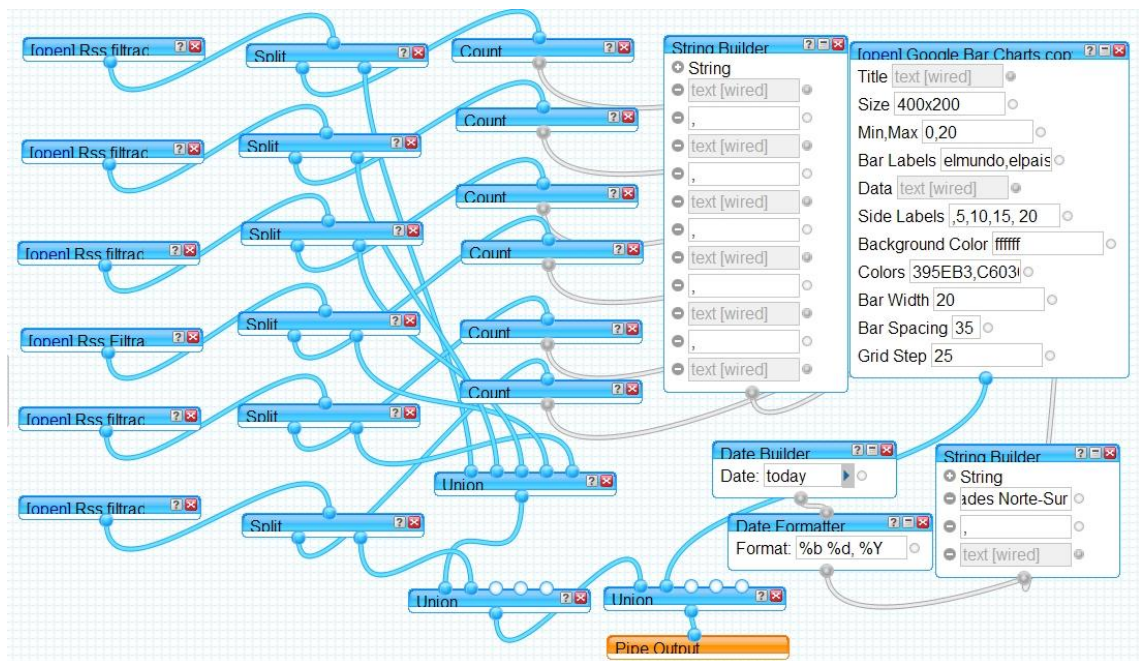


Figura 17. Pipe Gráfica+noticias

Pasos para su creación:

- Se seleccionan y arrastran al área de creación del pipe los seis pipes anteriormente creados: *Rss filtrados elmundo*, *Rss filtrados elpais*, *Rss filtrados rebelión*, *Rss filtrados ABC copy*, *Rss filtrados lavanguardia copy* y *Rss filtrados público*.
- Se crean seis módulos *Split* cuyas entradas se conectarán a cada uno de los seis módulos anteriores.
- El módulo *Split* tiene dos salidas. La primera se conectará con un módulo *Count*, y la segunda con un módulo *Union*, tal y como aparece en la imagen.

4. Se crea un módulo *String Builder* y se conectan sus campos con cada una de las salidas de los módulos *Count*. Entre cada uno de esos campos irá otro con la coma (,) como valor.
5. Se elije el módulo *Google Bar Charts*, y se conecta su campo "Data" con la salida del módulo *String Builder* antes creado.
6. Se crean los módulos *Date Builder*, *Date Formatter* y *String Builder* y se rellenan y conectan tal y como aparece en la imagen. La salida del módulo *Date Formatter* está conectada con el campo "text[wired]" del módulo *String Builder*. Y la salida de éste está conectado con el campo "Title" del módulo *Google Bar Charts*. Esto es para que en los títulos de cada gráfica aparezca su fecha.
7. Se rellenan los campos que quedan del módulo *Google Bar Charts* tal y como aparece en la imagen. El valor del campo "Colors" (colores de cada una de las barras de datos) es 395EB3,C60303,00BC16,F26000,663399,FFCC00,FF0099
8. Por último se conectan los módulos *Union* entre sí, como aparece en la imagen, y con el *Pipe Output*.

2.4 Incluir entrada fija con explicación.

Crear una nueva entrada y a la hora de editarla, marcar en el apartado Visibilidad (en la barra de la derecha) el atributo "Fija". Esto es para que dicha entrada siempre esté la primera de todas, al comienzo del blog.



Figura 18. Entrada fija

Texto:

La gráfica de barras representa el número de noticias relacionadas con el Desarrollo Humano Sostenible que han aparecido en los diarios El Mundo, El País, Rebelión, ABC, La Vanguardia y Público en las últimas veinticuatro horas.

El blog se alimenta de forma automática cada día a las 23:50 horas, pasándole al plugin FeedWordPress el rss resultante de un pipe de Yahoo, que se encarga de obtener mediante rss las noticias de la sección internacional de diferentes diarios.

Este pipe filtra las noticias por palabras tales como "hambre", "pobreza", "guerra" o "muerte" (que aparezcan en el título), y representa en la gráfica el número de noticias de cada diario.

2.5 Configuración de Cron.

Explicación:

Un *Cron Job* es una tarea que se programa para ser ejecutada cada cierto tiempo, gracias al reloj interno del servidor o del sistema operativo. En nuestro caso haremos uso del acceso gratuito al Cron que ofrece el servidor 000webhost, donde está alojado el blog. Como es gratuito, tiene un límite de sentencias php que podemos invocar. Por ello, haremos una llamada simple a un fichero php, que es el que hará lo necesario para que se ponga en marcha *Feedwordpress*.

Contenido del fichero `linea3.php`:


```
<?php
    $curl = curl_init();
    curl_setopt ($curl, CURLOPT_URL, "http://observamedios.hostoi.com?update_feedwordpress=1");
    //return the transfer as a string
    curl_setopt($curl, CURLOPT_RETURNTRANSFER, true);
    if(curl_exec($curl) === false)
    {
        echo 'Curl error: ' . curl_error($curl);
    }
    else
    {
        echo 'Operation completed without any errors';
    }
    curl_close ($curl);
?>
```

Este fichero php será ejecutado por el *Cron* cada día a las 23:50 horas, para actualizar el blog, llamando al update del plugin *Feedwordpress*.

Pasos a seguir:

1. Antes de nada, hay que incluir el archivo **linea3.php** dentro del directorio `public_html` del servidor.
2. Entrar en el área de miembros de la web del servidor 000webhost e introducir el email y la clave.
3. Entrar en **Go to CPanel** (bajo de Action)
4. Dentro de *Advanced*, entrar en la sección **Cron Jobs**.
5. Una vez en el manager, introducir lo siguiente:
 - en el segundo recuadro de “Command to run”: `public_html/linea3.php`
 - 59 minutos, 14 horas (son 9 horas menos, y lo que queremos es que se actualice a las 23:59 de nuestra hora).
 - Every Day, Every Month, Every Week Day

La siguiente figura recoge el resultado en el panel de creación de una nueva tarea *Cron Job*:

 **Add new Cron Job**

Command to run:


Minute (s):

Hour (s):

Day(s):

Months(s):

Weekday(s):

 **List of current cron jobs:**


» Time	» Command to run	» Action
50 14 ***	php -f /home/a6031473/public_html/linea3.php	

Figura 19. Nueva tarea *Cron Job*.

Nota: El proceso para actualizar el blog desde el *Cron Job* que se ha explicado se debe a que el servidor de hosting es gratuito y no permite modificar los comandos que aparecen en el primer parámetro de la Figura 19. Es por ello que se ha usado un fichero php en el que se incluye la instrucción *curl* para realizar el *update* de *FeedWordPress*.

En caso de disponer de otro servidor que sí lo permita no haría falta usar un fichero php, pues bastaría con poner en la línea de comando del Cron Job lo siguiente:

`*/10 * * * * /usr/bin/curl --silent http://observamedios.hostoi.com?update_feedwordpress=1`
 (tal y como pone en el panel de configuración *Feeds & Updates* del plugin *FeedWordPress*).

FASE 3. ELABORAR SECCIÓN MAPAS

3.1 Elaboración de los Pipes.

Auto-Geo.

Este pipe recibe la dirección de un feed rss y geolocaliza sus noticias haciendo uso de la base de datos de la web de *geonames* (<http://www.geonames.org>), que devuelve la latitud y longitud de cada ciudad. Después se escogerán sólo las noticias cuya fecha de publicación coincida con la actual.

El resultado que devuelve el pipe es el conjunto de noticias con su geolocalización, para poder ser ubicadas en un mapa.

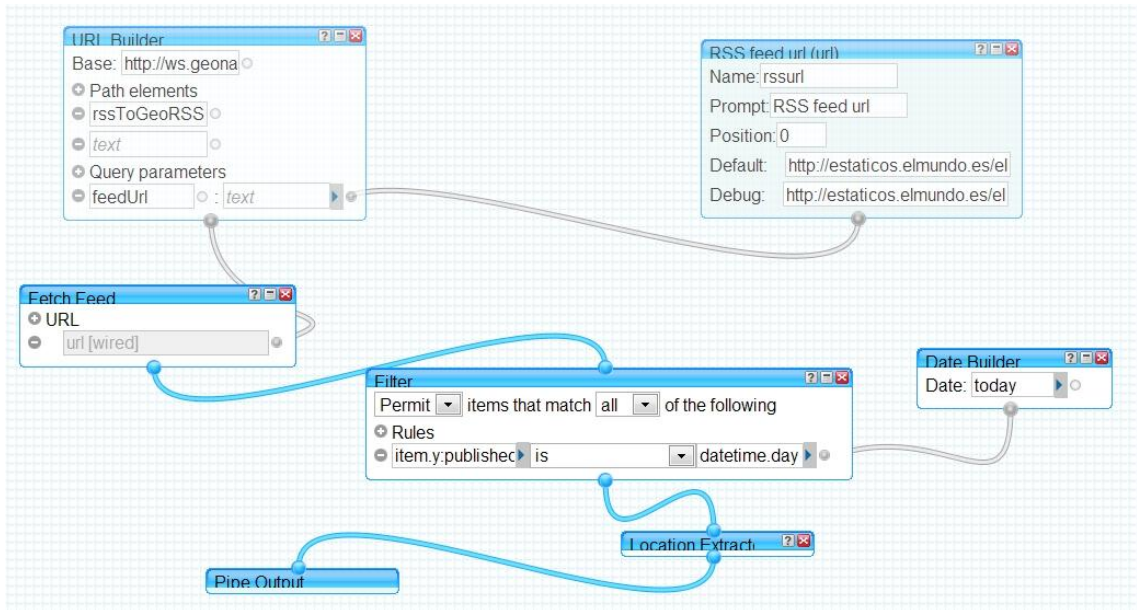


Figura 20. Pipe Auto-Geo. Para geo-localizar las noticias en el mapa.

Este módulo para geo localizar noticias en un mapa es facilitado por otro desarrollador de Yahoo Pipes. Tan sólo se ha cambiado la dirección *rss* de los campos “Default” y “Debug” del módulo *RSS feed url* y se han añadido los módulos *Filter* y *Date Builder* tal y como aparecen en la imagen, para dejar pasar sólo las noticias cuyas fechas coincidan con la actual.

Pasos para su creación:

1. Hacer un clon del pipe *RSS Auto-GeoPostedto Google Maps*.
2. Cambiar la dirección *rss* de los campos “Default” y “Debug” del módulo *RSS feed url*.
3. Añadir los módulos *Filter* y *Date Builder* tal y como aparecen en la imagen, conectando *Filter* con *Location Extractor*.

Mapa diarios España.

Este pipe hace uso del anterior para cuatro diarios. El mismo pipe se usa para los otros cuatro países restantes, de los cuales mostrará un mapa de noticias.

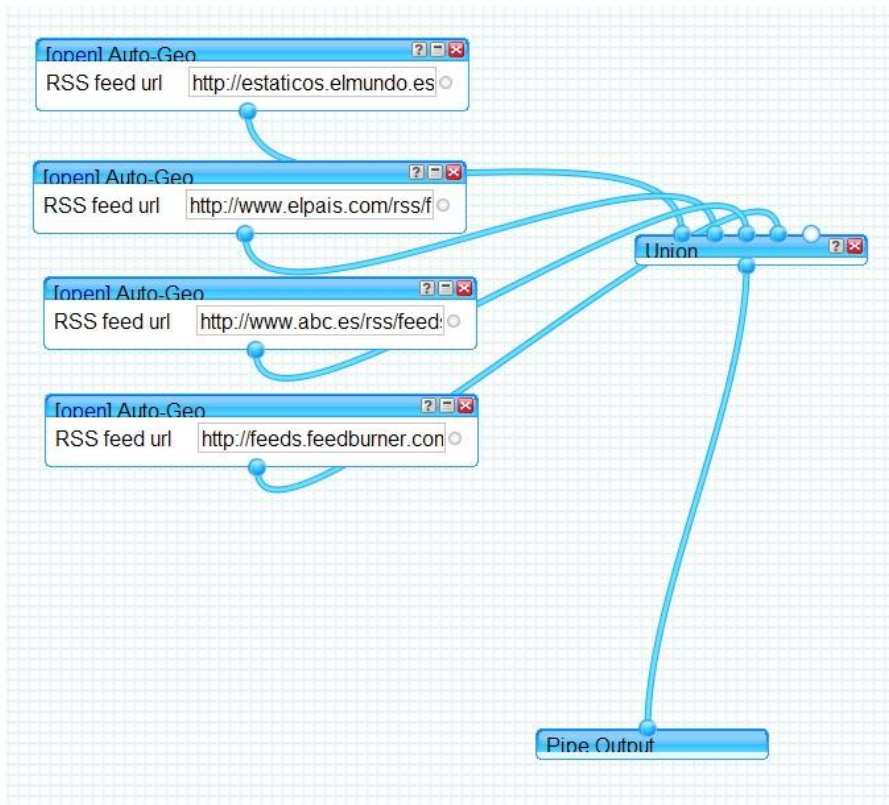


Figura 21. Pipe Mapa diarios de España.

Pasos para su creación:

1. Se elige cuatro veces el módulo *Auto-Geo*, antes creado, e incluimos en ellos las direcciones rss de los cuatro diarios que vamos a utilizar para los mapas.
2. Se crea un módulo *Union* para unir estos cuatro módulos.
3. La salida del módulo *Union* se conecta con el *Pipe Output*.

3.2 Incluir código *Badge* en página.

Yahoo Pipes ofrece una herramienta *Badge* que en este caso es muy útil, puesto que localiza las noticias devueltas por el pipe anterior y las representa en un mapa de Yahoo. Simplemente tenemos que copiar el código script que nos da (entrando en “Get as a Badge options” y luego eligiendo Wordpress) y pegarlo en la página del blog en la que queremos mostrar los mapas.

Mapa de noticias de diarios de España:

```
<script
src="http://l.yimg.com/a/i/us/pps/mapbadge_1.3.js">{"pipe_id":"f69fd0fe3208a956119230d3264eb124", "_btype":"map", "width":"550", "height":"450"}</script>
```

Mapa de noticias de diarios de Francia:

```
<script
src="http://l.yimg.com/a/i/us/pps/mapbadge_1.3.js">{"pipe_id":"fde07bfb9fe86d4d4230358f5bb6af92", "_btype":"map", "width":"550", "height":"450"}</script>
```

Mapa de noticias de diarios de Reino Unido:


```
<script
src="http://l.yimg.com/a/i/us/ppp/mapbadge_1.3.js">{"pipe_id":"b9f51adb7b5530d972abc940a4b9a782","_bty
pe":"map","width":"550","height":"450"}</script>
```

Mapa de noticias de diarios de Estados Unidos:

```
<script
src="http://l.yimg.com/a/i/us/ppp/mapbadge_1.3.js">{"pipe_id":"34220ee3a5f384a6e656e2bd16dd37f1","_bty
pe":"map","width":"550","height":"450"}</script>
```

Mapa de noticias de diarios de Canadá:

```
<script
src="http://l.yimg.com/a/i/us/ppp/mapbadge_1.3.js">{"pipe_id":"74c468c10868f50f23eb82488aba22b5","_bty
pe":"map","width":"550","height":"450"}</script>
```

Las siguientes figuras recogen el resultado de integrar este código en el blog:



Figura 22. Mapa diarios de España. Resultado capturado del blog.

Mapa de noticias de diarios de Reino Unido:
Mapa diarios Reino Unido



Figura 23. Mapa diarios Reino Unido. Resultado capturado del blog.

Como se puede ver en la imagen, este modelo de representación ubica en el mapa las noticias, por lo que cumple con el objetivo del proyecto de mostrar la visión que se tiene en los medios del mundo. Además, permite contrastar las noticias ya que pinchando sobre cada punto marcado en el mapa se puede acceder a la noticia misma.

3.3 Incluir explicación en la página.

Texto:

Los siguientes mapas de Yahoo Pipes geolocalizan las noticias de la sección internacional de los diarios de diferentes países en el presente día. De este modo se puede ver fácilmente qué partes del mundo se están cubriendo y cuáles quedan invisibles.

FASE 4. ELABORAR SECCIÓN TABLAS

4.1 Elaboración de los Pipes.

Países.

Este pipe recoge la url de un feed dado, cuyas noticias son filtradas, igual que en anteriores ocasiones, por la fecha de publicación, dejando pasar sólo las que correspondan con el día actual. Después emplea el módulo *Location Extractor* que devolverá la localización de

cada una de ellas, y concretamente el atributo *item.y:location.country*, el cual nos interesa para obtener el país de cada noticia, al que acompañará el número de repeticiones. Mediante el módulo *Sub-element* elegimos dicho atributo *country*, que será el título de un nuevo Rss.

Para obtener el número de repeticiones de cada país en las noticias usamos el módulo *Unique*, del cual nos interesa el atributo *item.y:repeatcount* (número de repeticiones). El resultado del pipe serán los países y el número de veces que aparece cada uno.

Este rss devuelto por el pipe tendrá que ser leído mediante php para escribirlo a modo de tabla en la página del blog. La información tiene que estar tabulada para que al pegar el texto en la web *Many Eyes*, ésta pueda reconocerla como una tabla de la cual podamos obtener diferentes visualizaciones y gráficos.

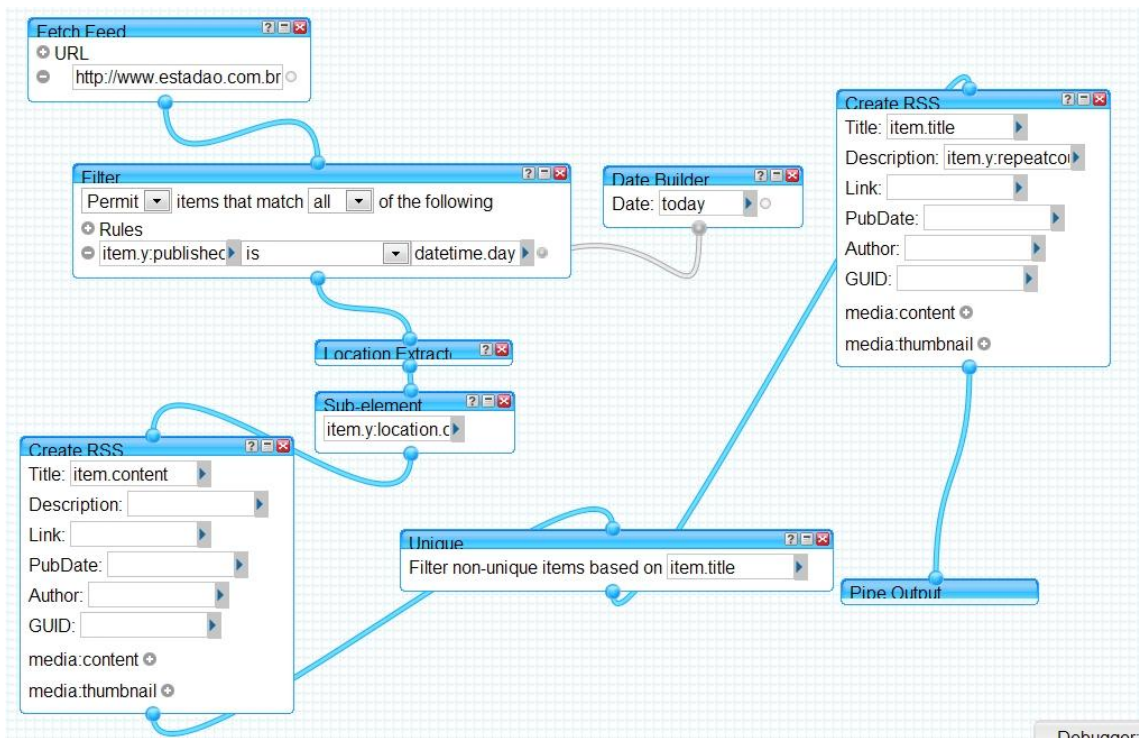


Figura 24. Pipe de Países. Devuelve los países de las noticias y su recuento.

Nota: Se hacen cuatro pipes como éste, uno por cada una de las agencias.

Pasos para su creación:

1. Se elije el módulo *Fetch Feed*, al que se le incluye la dirección rss de una de las agencias.
2. Se elijen los módulos *Filter* (cuya entrada se conecta con la salida del anterior módulo) y *Date Builder* y se combinan como aparece en la imagen.
3. La salida del *Filter* se conecta con la entrada del *Location Extractor*, y la salida de éste con un módulo *Sub-element*, que se crea para seleccionar elementos dado un valor. Se elije como valor de su campo el *item.y:location.country*. Así se escogen sólo los nombres de los países de las noticias.

4. Se crea un módulo *Create RSS*, al que le pasamos como entrada los ítems del módulo *Sub-element*. De este módulo sólo rellenamos el campo “Title” con el valor `item.content`.
5. Se selecciona el módulo *Unique*, conectando su entrada con la salida del anterior y poniendo su campo como `item.title`. Este módulo se usa para filtrar los ítems repetidos y contar el número de repeticiones.
6. Se crea otro módulo *Create RSS*, al que se le pasa la información resultante del módulo anterior, y se rellenan únicamente sus campos “Title” y “Description” con los valores `item.title` e `item.y:repeatcount`, respectivamente.
7. El módulo *Pipe Output* devolverá los nombres de los países que aparecen en las noticias y el número de veces que aparece cada uno.

4.2 Configuración de plugin *Exec-PHP*.

Explicación:

Este plugin es necesario para poder usar Php en las páginas de Wordpress. Para su configuración se deben seguir los siguientes pasos.

Instalación:

1. Descomprimir el archivo y subirlo a la cuenta ftp.
2. Activar el plugin *Exec-PHP* dentro de la sección *Plugins – Instalados* del Escritorio de Wordpress.

Configuración:

Ir a la sección *Usuarios – Tu perfil* del Escritorio de Wordpress y activar la casilla “Disable WYSIWYG Conversion Warning”.

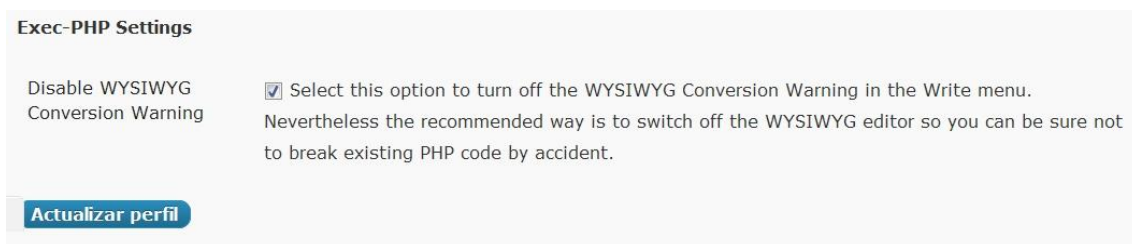


Figura 25. Exec-PHP, configuración.

4.3 Incluir código php y explicación en la página Tablas de datos.

Texto explicativo (en código html también, debido a las tablas):

<p>A continuación se muestra para diferentes agencias internacionales de noticias el número de noticias por país a día de hoy.</p>

<p>Esta información está pensada para que se seleccione con el puntero la correspondiente tabla y se pegue en la web Many Eyes, que permite generar muchos tipos de gráficos de datos.</p>

<p>Nota: En caso de que salga "PHP Error Message", darle a actualizar página.</p>

4.4 Código php para recorrer los feeds y dibujar las tablas.

El siguiente código recorre los ítems del *feed rss* del pipe, y escribe sus valores (País y Veces) tabulados como una tabla. Con la inclusión del código php en la página del blog, las tablas se cargarán automáticamente cada vez que se entre en dicha sección.

Explicación:

Como se puede observar en el código, para cada una de las tablas se carga en memoria (*load*) el contenido de la dirección del pipe que se creó para devolver los países y el número de repeticiones de cada uno, para una agencia.

La dirección del pipe es a su vez un fichero XML, por lo que se recorre mediante php (usando DOM, una herramienta de acceso a documentos XML) cada uno de los elementos (ítems) del *rss*. Y se escribe en una columna de la tabla el nombre del país, y en la otra el número de repeticiones.

Código PHP de las cuatro tablas de datos de la página del blog correspondiente:

```
<p><strong>AP International News</strong>:
<a href="http://www.ap.org/">The Associated Press</a></p>

<table>
<tr><td>País</td><td>Veces</td></tr>
<?php
$xml = new DOMDocument();
$xml->load('http://pipes.yahoo.com/pipes/pipe.run?_id=b7712718cd488befa65c32906af6af7a&_render=rss');
$raiz = $xml->documentElement;
$entradas = $raiz->getElementsByTagName('item');
$count = $entradas->length;
for ($i=0; $i<$count; $i++) {
    $titulo = $entradas->item($i)->getElementsByTagName('title')->item(0)->nodeValue;
    $valor = $entradas->item($i)->getElementsByTagName('description')->item(0)->nodeValue;
    echo "<tr><td>$titulo</td><td>$valor</td></tr>"; }
?>
</table>

<BR>
<p><strong>Ansa Mondo</strong>:
<a href="http://www.ansa.it/">ANSA</a></p>

<table>
<tr><td>País</td><td>Numero</td></tr>
<?php
$xml = new DOMDocument();
```

```

$xml->load('http://pipes.yahoo.com/pipes/pipe.run?_id=23b8cd2959803e7faee135fba55ff882&_render=rss');
$raiz = $xml->documentElement;
$entradas = $raiz->getElementsByTagName('item');
$count = $entradas->length;
for ($i=0; $i<$count; $i++) {
    $titulo = $entradas->item($i)->getElementsByTagName('title')->item(0)->nodeValue;
    $valor = $entradas->item($i)->getElementsByTagName('description')->item(0)->nodeValue;
    echo "<tr><td>$titulo</td><td>$valor</td></tr>"; }
?>
</table>

```

```

<BR>
<p><strong>Estadao Internacional</strong>:
<a href="http://economia.estadao.com.br/">Estadao</a></p>

```

```

<table>
<tr><td>País</td><td>Numero</td></tr>
<?php
$xml = new DOMDocument();
$xml->load('http://pipes.yahoo.com/pipes/pipe.run?_id=6d3a6fa7f4c0f846acf66e126f3b1ba3&_render=rss');
$raiz = $xml->documentElement;
$entradas = $raiz->getElementsByTagName('item');
$count = $entradas->length;
for ($i=0; $i<$count; $i++) {
    $titulo = $entradas->item($i)->getElementsByTagName('title')->item(0)->nodeValue;
    $valor = $entradas->item($i)->getElementsByTagName('description')->item(0)->nodeValue;
    echo "<tr><td>$titulo</td><td>$valor</td></tr>"; }
?>
</table>

```

```

<BR>
<p><strong>Reuters Internacional</strong>:
<a href="http://www.reuters.com/">Reuters</a></p>

```

```

<table>
<tr><td>País</td><td>Numero</td></tr>
<?php
$xml = new DOMDocument();
$xml->load('http://pipes.yahoo.com/pipes/pipe.run?_id=084f0c7b8d3f3e8b80812454ad19897b&_render=rss');
$raiz = $xml->documentElement;
$entradas = $raiz->getElementsByTagName('item');
$count = $entradas->length;
for ($i=0; $i<$count; $i++) {
    $titulo = $entradas->item($i)->getElementsByTagName('title')->item(0)->nodeValue;
    $valor = $entradas->item($i)->getElementsByTagName('description')->item(0)->nodeValue;
    echo "<tr><td>$titulo</td><td>$valor</td></tr>"; }
?>
</table>

```

La siguiente figura muestra el resultado de la integración del código anterior en el blog:

AP International News:
[The Associated Press](#)

País	Veces
Spain	1
El Salvador	1
Iraq	1
Thailand	1
Belgium	1
Eesti	1
Greece	1

Ansa Mondo:
[ANSA](#)

País	Veces
United States	5
Iraq	3
Russia	1
Brasile	1
India	1
Spagna	1
Thailand	3
Al Iraq	2
Corea	1
Israele	1
Pakistan	2
Grecia	1

Estadao Internacional:
[Estadao](#)

País	Veces
United States	1
Argentina	1

Figura 26. Tabla de recuento de países. Resultado capturado del blog.

Al entrar en esta página del blog, las tablas se cargan automáticamente. Pero hay casos en los que en varias aparece un mensaje que dice “PHP Message Error”. En este caso sólo hay que actualizar la página desde el navegador (F5) para que se visualicen correctamente.

FASE 5. PRUEBAS

5.1 Pruebas de la sección de Gráficas.

Para probar el funcionamiento de las gráficas, si no se desea esperar a las 23:50 horas se crea un *Cron* programado para una determinada hora (como se explica en el apartado 2.5), aunque para ello habría que hacer una copia del fichero **linea3.php**, ya que no es posible usar un mismo fichero para dos *Crons* distintos; y ver si se genera la gráfica con las noticias correspondientes.

Para hacer una prueba rápida hay otro modo de realizarlo, más simple, sin hacer uso de los *Crons*. Sería entrar dentro del tablero de administración de Wordpress y en la sección *Syndication* seleccionar el feed que se ha creado y darle al botón “Update”. Esto actualizará el blog con las nuevas entradas para ese día.

Nota importante: Si se hacen diversas pruebas consecutivas, será necesario eliminar cada vez las nuevas entradas generadas en las pruebas, vaciando incluso la papelera. Si no se hiciera, FeedWordpress detectaría que dichas entradas ya se encuentran publicadas y por tanto no volvería a actualizarlas de nuevo.

La figura siguiente muestra el resultado en las entradas:



Figura 27. Ejemplo de gráfica.

5.2 Pruebas de la sección de Mapas.

En esta sección realmente no hay muchas pruebas que hacer, salvo en lo que es la implementación del pipe. Una vez los códigos script están insertados en la página del blog, se cargarán los mapas correspondientes gracias a la aplicación *Badge* de Yahoo.

Hay veces en las que tarda mucho en cargar, o simplemente no llega a hacerlo. Pero esto es culpa del Yahoo Pipes, que no siempre está operativo.

5.3 Pruebas sobre *Many Eyes*.

Seleccionamos el texto de una de las tablas de la página “Tablas de datos” del blog y lo copiamos al portapapeles. Después vamos a la web de *Many Eyes* (<http://manyeyes.alphaworks.ibm.com/manyeyes>), en la cual tenemos que estar registrados y logeados para subir *data sets*, y entramos en el apartado “upload data set”. Una vez allí pegamos el texto anteriormente copiado en el apartado “2 Paste the data”. Comprobamos que la aplicación reconoce las columnas de datos. Rellenamos el campo “Title” requerido y pulsamos sobre el botón “Create”.

El resultado será una tabla como la siguiente:

Data Sets : Ansa Mundo [Edit | Delete]

Uploaded by: **Observatorio** Created at: Monday May 10 2010, 02:34 AM
Data Source: none provided
Description:

[view as text](#) [edit data set](#)

	País	Veces
1	United States	5
2	Iraq	3
3	Russia	1
4	Brasile	1
5	India	1
6	Spagna	1
7	Thailand	3
8	Al Iraq	2
9	Corea	1
10	Israele	1
11	Pakistan	2
12	Grecia	1

[watch this](#) [add to topic center](#) [visualize](#) [rate this](#)

Versions (1)
1. Original Data Set by **Observatorio** Today at 02:34 AM

Figura 28. Tabla en Many Eyes.

Para crear visualizaciones tan sólo tenemos que darle al botón “visualize” que aparece en la imagen, y seguidamente escoger entre las visualizaciones disponibles.

Estos son algunos ejemplos:

Bar Chart:

El gráfico de barras (*Bar Chart*) es un método clásico para comparaciones numéricas. El gráfico de barras de *Many Eyes* puede mostrar uno o más conjuntos de variables.

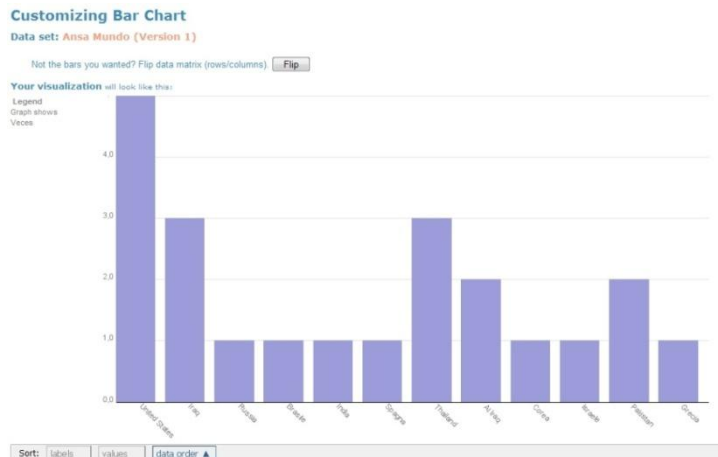


Figura 29. Many Eyes, Bar Chart.

Tag Cloud:

La nube de etiquetas (*Tag Cloud*) es una visualización de frecuencias de palabras. Permite saber la frecuencia con que aparecen las palabras en un texto dado, o ver la relación entre una columna de palabras y una columna de números. El tamaño de la palabra corresponde a la cantidad asociada a esa palabra.



Figura 30. Many Eyes, Tag Cloud.

Pie Chart:

El gráfico circular (Pie Chart) es una forma familiar de mostrar las proporciones, en porcentaje, entre diferentes datos.

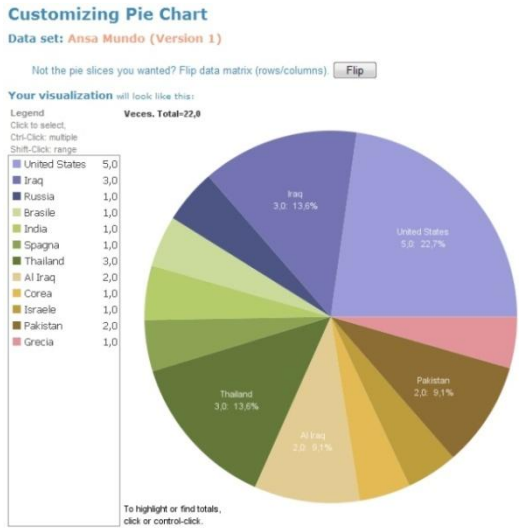
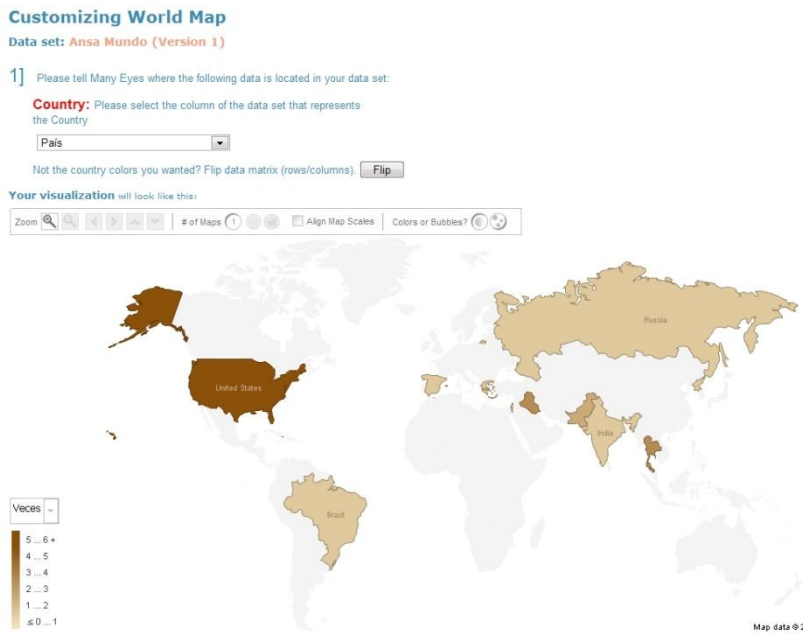


Figura 31. Many Eyes, Pie Chart.

World Map (o mapa caliente).

De todos los mapas, éste es quizá el más interesante, ya que extrae los nombres de los países y les da un grado de color de acuerdo con su columna Veces. Debido a que los nombres de los países que extrae el pipe de las noticias están en el idioma correspondiente a la nacionalidad del diario, la aplicación de *Many Eyes* nos mostrará el nombre en inglés que se ajuste a cada país.



4. PROBLEMAS

Dada la gran cantidad de problemas que hubo durante el transcurso del proyecto, y de la realización de las pruebas, este tema merecía un apartado propio. Algunos de los problemas vinieron, entre otros, por las propias limitaciones del *FeedWordPress* y por los escasos servicios de *Cron Job* gratuito que ofrecía el hosting. Pero sin duda, el que se llevó la palma fue **Yahoo Pipes**, la base de este proyecto y a la vez **una fuente inagotable de problemas**.

4.1 PROBLEMAS CON LAS GRÁFICAS

- Filtrado de palabras.
Dado que los feeds de los diarios escogidos sólo contenían los títulos de las noticias, la fecha y poco más, el filtrado de noticias sólo se podía hacer por las palabras contenidas en los títulos de las mismas. Con esta técnica de filtrado por términos son muchas las noticias que se “cuelan” sin que tengan que ver con el tema buscado. Un ejemplo: la palabra “guerra” puede darnos noticias como “La guerra de Microsoft contra Google”. Este fue un problema sin solución.
- Filtrar palabras completas.
Otro problema era que las noticias no se estaban filtrando por palabras completas. Es decir, si buscábamos la palabra “sur”, cualquier noticia que contuviera en su título esta corta palabra dentro de otra, por ejemplo “censura”, era publicada.
Esto tuvo fácil solución en el Pipes, usando “Matches regex” en lugar de “Contains”, dentro del módulo *Filter*. Y en el campo del término poniendo, por ejemplo: `(?i)\Wsur\W`.
- Actualización diaria.
Uno de los objetivos básicos de este proyecto es generar gráficas con frecuencia diaria. Pero el plugin *FeedWordPress* sin ayuda externa la única opción que ofrece al respecto es, aparte de la actualización manual, otro tipo de actualización semi-manual que consiste en que cada X minutos busca actualizaciones en los feeds, pero sólo cuando alguien accede al blog. De modo que si el blog no tiene ninguna visita, no hay actualizaciones. Y haciendo pruebas, ni aun así funcionaba bien.
La solución fue emplear *Crons*, que son tareas que se programan en el servidor para que a una hora y un minuto dados haga una función (como llamar al update del *FeedWordPress*). El servidor de hosting ofrecía manejo de *Cron Jobs*, aunque limitados por ser un alojamiento gratuito, pero con las opciones suficientes para lo que se pedía en el proyecto.
- Zona horaria.
Las primeras pruebas con el Cron no funcionaban, pero era simplemente porque la hora del servidor no coincidía con la de aquí. Eran nueve horas menos.
- Gráficas vacías y el módulo *Sort*.
En un principio no se publicaban las gráficas vacías (las que no tenían noticias) debido a

que en el pipe había un módulo *Sort* que ordenaba las noticias por orden alfabético. Al haber cero noticias el módulo daba un warning y no se creaba la gráfica. Este problema se solucionó simplemente quitando el módulo *Sort*, que no era necesario.

- Noticias del día.

El plugin cogía todas las noticias contenidas en el feed del pipes. Para limitar el filtrado a sólo las del día presente, la solución fue comparar en el módulo *Filter* la fecha de la noticia (atributo *item.y:publish.day*) con (*is*) la fecha actual (módulo *Date Builder* con atributo *today*).

- Varias gráficas vacías.

Cuando en un día no había noticias, la gráfica no se publicaba en el blog si ya existía antes otra gráfica vacía publicada. Esto era debido a que *FeedWordPress* no publica una entrada si el contenido del feed, es decir su lista de item títulos, es exactamente el mismo que el de uno que ya se publicó anteriormente (y en este caso era el mismo, ya que todas las gráficas tenían de título "Análisis de noticias de Desigualdades Norte-Sur". La única opción que daba el plugin en este aspecto era la de actualizar el contenido antiguo por el nuevo, pero eso no valía para el propósito actual.

La solución estaba en el Pipes, mediante al uso de los módulos *Date Builder* (valor *today*) y *Date Formater* (para darle formato a la fecha). Esa información se pasaba a un *String Builder* para crear una cadena que fuera del estilo "Análisis de noticias de Desigualdades Norte-Sur, 24-04- 10". Y se le pasaba al título del módulo que crea la gráfica. De este modo los títulos de cada gráfica eran únicos.

- Fallo del *Pipes* a la hora de actualizar.

Al aumentar el tamaño y la complejidad del pipe (siendo notoria la inclusión de la fecha en los títulos de las gráficas), y teniendo en cuenta que hay momentos en los que Yahoo Pipes no está operativo, no siempre se actualizaban las entradas del blog.

Esto llevó a que se tomase la decisión de usar varios *Crons* consecutivos , separados por intervalos de 5 minutos (uno a las 23:50, otro a las 23:55 y el último a las 23:59), para que en el caso de que fallase el Pipes en el momento de la actualización, tuviese posibilidades de funcionar en el siguiente Cron. Pero fue una solución que no servía, puesto que el plugin *Feedwordpress* estaba puesto para que buscara actualizaciones cada 60 minutos. De modo que si intentaba actualizar a las 23:50, daba igual que hubiese otros *Crons* inmediatos, ya que el plugin no buscaría nuevas actualizaciones hasta una hora después. Y según ponía en el plugin, no es aconsejable poner el intervalo de actualización a menos de 60 minutos, porque el servidor podría banear la ip.

Este fue otro de los problemas sin solución.

4.2 PROBLEMAS CON LOS MAPAS

- Pérdida de noticias a la hora de localizarlas.
También en la sección de Mapas, Yahoo Pipes era el protagonista de los problemas. En primer lugar porque había un porcentaje destacado de noticias (en proporción de, por ejemplo, 15 de 25) que no se muestran en el mapa debido a problemas en el módulo *Location Extractor*. Y cuantas más noticias se le pasaban, más se perdían. En un principio, esta sección sólo contenía un mapa en el que se mostraban todas las noticias juntas de muchos países diferentes, pero comparando los resultados del listado de noticias resultantes (unas 400) con las que se mostraban en el mapa (unas 70), se optó por hacer diferentes mapas separados por países. De este modo, en conjunto, se mostraban más noticias.
Se informó de esta pérdida de noticias en el *Pipe Suggestion Board* y en el foro de ayuda:

"This news can be displayed on the map Yahoo Badge, using the pipe "Auto-Geo RSS." In this regard, I have also seen a lost on a large percentage of news that are not shown on the map. For example, news only shows 16 (of 26) on the map."

No hubo respuesta por parte del equipo de desarrollo, por lo que fue un problema sin solución.

- No son pocas las veces en las que los mapas tardaban mucho en cargar, incluso simplemente no llegando a hacerlo en algunas de ellas. Dependía de lo operable que estuviera en ese momento Yahoo Pipes. Como nuevamente dependía del funcionamiento de Yahoo Pipes fue un problema sin solución.
- Mala ubicación, en ocasiones, de las noticias.
La ubicación de las noticias en el mapa presentaba problemas en muchas ocasiones, puesto que no siempre se localizaban en los puntos donde debería. Este problema sin solución estaba relacionado con la base de datos de *geonames*, que confundía ciertos lugares cuando el nombre de la localidad que extrae de la noticia está en castellano o en otra lengua diferente al inglés. Eso sin contar con que para lenguas minoritarias ni siquiera era capaz de ubicar las noticias en el mapa.

4.3 PROBLEMAS CON LAS TABLAS

- Incompatibilidad entre los plugins *Exec-Php* e *Inline-js*.
Para poder recorrer los RSS se optó por usar php. Para que Wordpress soportara php en sus páginas hubo que instalar *Exec-Php*. Al comenzar a hacer pruebas parecía no funcionar, pero era porque estaba instalado también el *Inline-js* (que daba soporte javascript a Wordpress), con el cual parecía ser incompatible. Así pues, la solución fue desinstalar el plugin *Inline-js*.

- La tabulación.

Con php se recorría sin problemas el RSS, pasándole la url del pipe, y se escribía en la página del blog su contenido. Pero había un problema, ya que la web *Many Eyes* sólo interpretaba los valores como una tabla si la información que se le pegaba estaba tabulada en dos columnas. La solución fue escribir dentro del php los valores en filas y columnas de html.

- Las tablas dejaron de funcionar por culpa de Yahoo Pipes.

Un nuevo fallo en Yahoo Pipes, concretamente en el módulo *Location Extractor*, hizo que las tablas, que iban sin problemas, dejaran de hacerlo.

El módulo *Location Extractor* dejó, de la noche a la mañana, de devolver los valores "location", entre ellos el atributo que se escogía con él módulo *Sub-element: item.y:location.country*. Este valor daba el país de la noticia.

Por ello se escribió un mensaje a Yahoo Pipes informando de este hecho:

"Hi,

I'm a student from Spain and I'm using Yahoo Pipes for a university project that deals with news filtering. I am writing you because I see that there are problems the past few days with the module "Location Extractor", in particular in combination with the module "Sub-element."

It used to work well, but since last week (not having made any change) I see the Location Extractor no returns items such as latitude, longitude, country ... And so It does not appear in any of the pipes that had built and operated, the item "item.y: location.country" to deploy the module sub-element. Thing before I could.

Here is the pipe, with counts the number of repetitions of each country:

http://pipes.yahoo.com/pipes/pipe.edit?_id=60b5f615b5b65bbcde76c1902eed8824

This news can be displayed on the map Yahoo Badge, using the pipe "Auto-Geo RSS." In this regard, I have also seen a lost on a large percentage of news that are not shown on the map. For example, news only shows 16 (of 26) on the map.

Thank you very much.

Greetings,

Alberto"

Varios usuarios respondieron afirmando haberse dado cuenta de este fallo, y que *Location Extractor* nunca había sido del todo fiable. Sin embargo, no hubo respuesta por parte del equipo de desarrollo, por lo que fue un problema sin solución.

5. CIERRE

Una vez vistos los apartados de “Ejecución” y “Problemas”, pasamos al documento de cierre.

5.1 DESCRIPCIÓN DE LA SOLUCIÓN

Como ya se dijo en “Contenidos y funcionalidades”, el blog consta de cuatro secciones. A continuación se muestra una captura de pantalla de cada una de ellas:

- Sección principal de gráficas (“Home”).

The screenshot shows the homepage of the 'Observatorio de Medios de Comunicación 2.0' website. The header features the site title and a navigation menu with links for 'Home', 'Mapas de noticias', 'Tablas de datos', and 'Acerca de'. A search bar is located in the top right corner. The main content area displays three news analysis charts, each with a title, author, and a bar chart showing the number of news items from various sources. The right sidebar contains sections for 'Categories', 'Archivos', 'Entradas recientes', and 'Meta'.

Observatorio de Medios de Comunicación 2.0
Un proyecto de Policlick

Home | Mapas de noticias | Tablas de datos | Acerca de

Explicación de la gráfica
Author: admin

La gráfica de barras representa el número de noticias relacionadas con el Desarrollo Humano Sostenible que han aparecido en los diarios El Mundo, El País, Rebelión, ABC, La Vanguardia y Público en las últimas veinticuatro horas.

El blog se alimenta de forma automática cada día a las 23:50 horas, pasándole al plugin FeedWordPress el rss resultante de un pipe de Yahoo, que se encarga de obtener mediante rss las noticias de la sección internacional de diferentes diarios. Este pipe filtra las noticias por palabras tales como "hambre", "pobreza", "guerra" o "muerte" (que aparezcan en el título), y representa en la gráfica el número de noticias de cada diario.

Marzo 1st, 2010 | Posted in [Internacional](#) | [No Comments](#) »

Análisis de noticias de Desigualdades Norte-Sur, 16-05-10
Author: grafica+noticias

Análisis de noticias de Desigualdades Norte-Sur, 16-05-10

Diario	Número de noticias
elmundo	0
elpais	0
rebelion	0
abc	0
lavanguardia	1
publico	1

Mayo 16th, 2010 | Posted in [Internacional](#) | [No Comments](#) »

Al menos 25 muertos durante una sesión del Parlamento en Somalia
Author: grafica+noticias

Otras 60 personas han resultado heridas después de que insurgentes islámicos de Alshabab dispararan morteros contra el edificio

Mayo 16th, 2010 | Posted in [Internacional](#) | [No Comments](#) »

Análisis de noticias de Desigualdades Norte-Sur, 15-05-10
Author: grafica+noticias

Análisis de noticias de Desigualdades Norte-Sur, 15-05-10

Diario	Número de noticias
elmundo	1
elpais	0
rebelion	0
abc	1
lavanguardia	1
publico	1

Mayo 15th, 2010 | Posted in [Internacional](#) | [No Comments](#) »

Categories
• [Internacional](#) (47)

Archivos
Seleccionar mes

Entradas recientes

- [Análisis de noticias de Desigualdades Norte-Sur, 16-05-10](#)
- [Al menos 25 muertos durante una sesión del Parlamento en Somalia](#)
- [Análisis de noticias de Desigualdades Norte-Sur, 15-05-10](#)
- [Dos muertos al estrellarse una avioneta en Tarragona](#)

Meta

- [Iniciar sesión](#)
- [RSS de Entradas](#)
- [RSS de los comentarios](#)
- [WordPress.org](#)

Figura 33. Sección Gráficas.

- Sección “Mapas de noticias”.



Figura 34. Sección Mapas.

- Sección “Tablas de datos”.

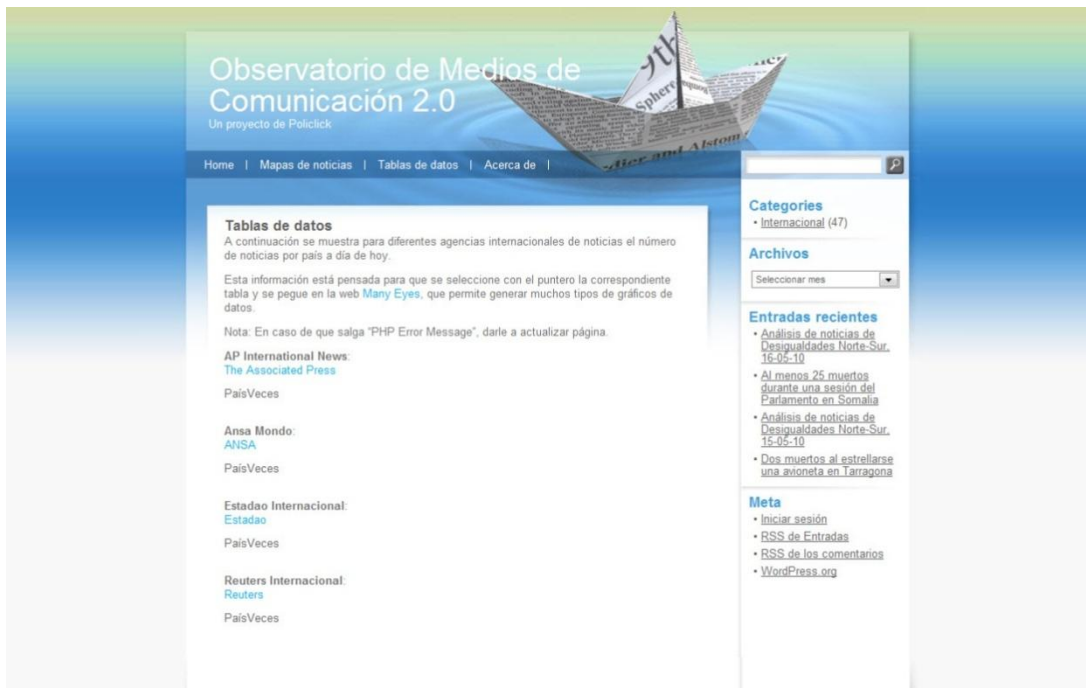


Figura 35. Sección Tablas de datos.

- Sección “Acerca de...”.

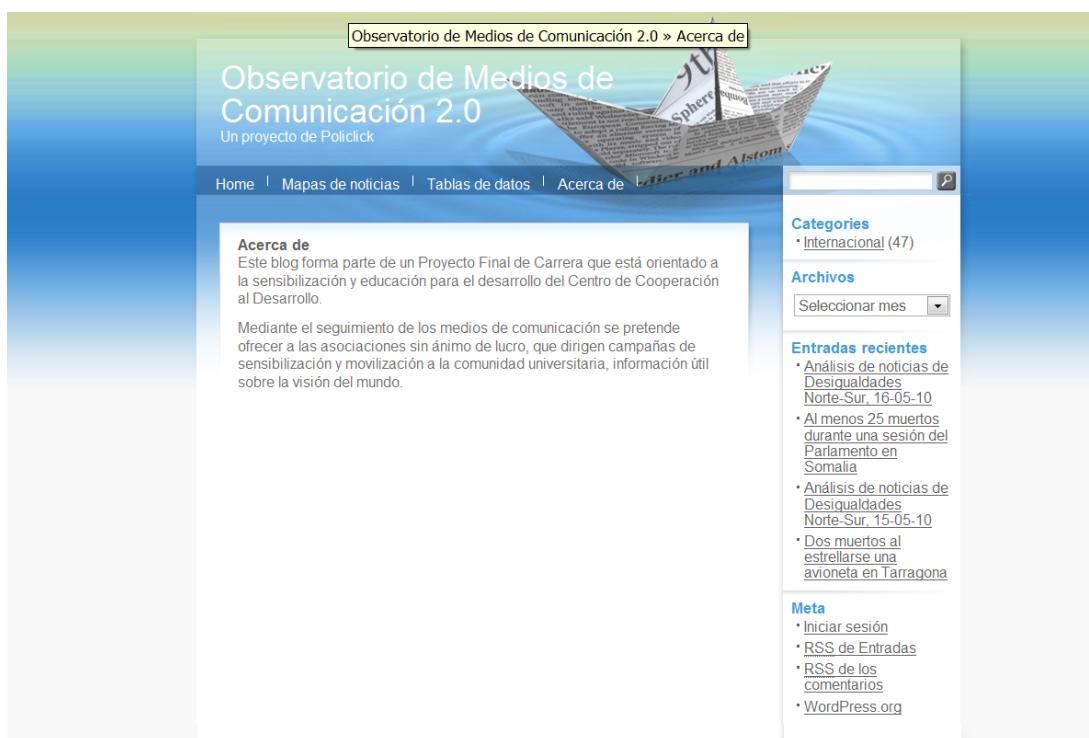


Figura 36. Sección Acerca de...

5.2 OBJETIVOS CONSEGUIDOS

Uno de los requisitos de partida de este proyecto es que tenía que desarrollarse mediante software libre. En este sentido Yahoo Pipes ofrece muchísimas posibilidades de hacer casi cualquier cosa pero, por el contrario, es una herramienta en la que uno no puede confiar la efectividad de su aplicación, ya que no son pocas las ocasiones en las que se muestra inoperativo. En este proyecto, no sólo ha dado problemas con la actualización del blog, sino que además el apartado de **Tablas de datos** se convirtió en inservible de un día para otro, cuando dejó de funcionar correctamente el módulo encargado de extraer la localización de las noticias.

El objetivo principal del Observatorio es mostrar la visión que tienen los medios sobre las desigualdades norte-sur y la pobreza en el mundo. En este sentido, dadas las limitaciones de las herramientas web 2.0 empleadas sólo se pudo filtrar las noticias por palabras significativas en su título. Y de este modo, es muy **difícil filtrar las noticias** de este tema en concreto.

Creamos un cuadro resumen con el grado de consecución de los objetivos marcados al inicio del proyecto:

Tabla 1. Objetivos y porcentaje conseguido

Objetivos	Porcentaje conseguido
La información relativa a la visión del mundo que ofrecen los medios de comunicación se genere al 80% de forma sistemática y automática.	100%
La información contenida en el sitio Web se actualice diariamente de forma automática o semiautomática.	100%
Contrastar noticias de distintas fuentes y mostrar la información, para dar la visión que pretende ofrecer el observatorio.	80%
Mostrar de forma clara y sencilla el contraste de todas las noticias filtradas.	80%
Valoración global del proyecto.	90%

5.3 LECCIONES APRENDIDAS

Las herramientas web 2.0 ofrecen una amplia gama de posibilidades para el desarrollo de aplicaciones web, ya que podemos combinar unas con otras, como se ha hecho en este proyecto enlazando Wordpress con Yahoo Pipes; e incluso, generando información con la intención de ser exportada a *Many Eyes*. Esto es muy interesante, aunque también hay que tener en cuenta que cuanto mayor es el número de piezas que interconectamos, mayores son las posibilidades de que falle por algún sitio. O que, en el caso del software libre alguna de las piezas deje de funcionar y afecte a toda la solución.

A la hora de afrontar un proyecto de este tipo es muy importante estar “despierto”, conocer las herramientas que se encuentran en la red, a nuestro alcance, y ver lo que pueden ofrecernos; buscar la reutilización. Muchas veces la solución más compleja no es siempre la mejor. Un ejemplo claro en este proyecto es la generación de Heat Maps (o mapas calientes), que permiten visualizar con diferentes tonos de color de qué países del mundo ha habido más noticias. En un principio se pensó en utilizar unas librerías de Google para hacer esto, pero aparte de que era bastante complejo, no teníamos permisos de instalación en el servidor. La solución más sencilla y efectiva fue enlazar este proyecto con el de *Many Eyes*, que permitía generar todo tipo de visualizaciones.

6. CONCLUSIONES

Cuando comencé este proyecto ya partía con una base en cuanto a los *pipes* y a su integración con Wordpress, tras haber cursado la asignatura de Introducción a la Producción Multimedia (IPM). Gracias a este proyecto he ampliado bastante mis conocimientos sobre la materia, y me he dado cuenta de la dimensión de las herramientas web 2.0, con las cuales se puede construir aplicaciones muy interesantes.

Además de esto, también he aprendido a buscar por mí mismo soluciones a los problemas que iba encontrando. La consecución del proyecto podría decirse que fue exponencial. Después de una primera etapa de investigación de tecnologías le siguió otra aún más larga de cortos avances y de bloqueos con los problemas, en la que las soluciones que probaba resultaban poco efectivas. Fue entonces cuando ya teniendo los conocimientos necesarios, comenzaron a surgir mejores ideas y los objetivos se fueron logrando de forma mucho más rápida, pese a la gran cantidad de problemas que se escapaban a mi control (especialmente de Yahoo Pipes). Al tener una visión más global del proyecto, sabía mejor qué herramientas necesitaba y cómo podía abordar la solución. Esto me hizo también adquirir ciertos conocimientos básicos de *php*, tanto para recorrer *feeds rss* como la ejecución de tareas en el servidor mediante *Cron Jobs*.

Siempre es satisfactorio cuando se van logrando las metas, pero lo es más teniendo en cuenta el objetivo de este proyecto: poner de manifiesto la visión que tienen los medios sobre nuestro mundo, y concretamente sobre las desigualdades y la pobreza. Y en este sentido me siento verdaderamente realizado, y orgulloso de haber formado parte de un proyecto que se desmarca de las aplicaciones habituales de la informática, contribuyendo, humildemente, a poner la informática al servicio de la cooperación al desarrollo y de la información.

ANEXO 1: ANTECEDENTES DEL PROYECTO

A continuación se detallan los antecedentes del proyecto, aportados por el **Observatorio de Medios de comunicación.**

Introducción

Los Medios de Comunicación como Televisión, Prensa, Radio e Internet se han convertido en la principal fuente de información de los ciudadanos. A través de ellos, reciben continua información sobre lo que acontece en el mundo, tanto a nivel local como global. La realidad mediatizada, creada por los Medios de Comunicación, establece prioridades en el imaginario colectivo sobre lo que es, y no es, importante.

Los últimos datos del Estudio General de Medios indican que la penetración de la Televisión en el 2008 en el conjunto de la población (38.261.000) era del 88.5%, la de los Diarios del 42.1%, la de la Radio del 53.1% y la de Internet del 45.4%¹. Por otro lado, la cuota de pantalla informativa de Prime Time puede verse en la siguiente figura 1. Estos datos traducido a número de espectadores ofrecen cifras superiores a los 9.5 millones de personas informándose con los telediarios de máxima audiencia.

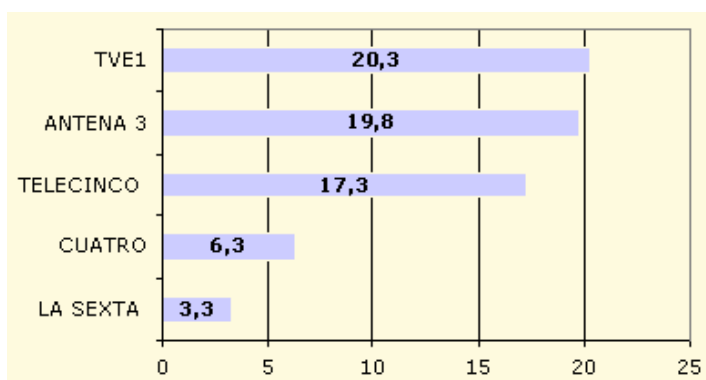


Figura 37. Cuota de pantalla de informativos de Prime Time.

En relación a la audiencia de Internet, en datos generales, el EGM muestra que cerca de 18.000.000 millones de personas han utilizado Internet en el último mes. En general, tanto este dato como el de uso de Internet el día anterior a la encuesta, muestra una tendencia al alza desde el año 1996.

Audiencia de Internet

Datos generales

	Usan Ordenador	Usan Ordenador habitualmente	Usan Internet último mes	Usan Internet ayer	Conexión a Internet en el hogar
1996	6.568 (19.4%)	4.449 (13.2%)	342 (1.0%)		
1997	7.374 (21.7%)	5.105 (15.0%)	931 (2.7%)	321 (0.9%)	
1998	8.433 (24.7%)	5.960 (17.5%)	1.579 (4.6%)	627 (1.8%)	
1999	9.280 (26.9%)	6.624 (19.2%)	2.429 (7.0%)	1.003 (2.9%)	
2000	10.350 (29.8%)	7.567 (21.8%)	4.363 (12.6%)	1.943 (5.6%)	
2001	11.512 (33.1%)	8.551 (24.6%)	7.120 (20.4%)	3.143 (9.0%)	6.338 (18.2%)
2002	11.731 (33.7%)	8.457 (24.3%)	7.828 (22.5%)	3.693 (10.6%)	7.626 (21.9%)
2003	13.082 (37.1%)	9.715 (27.6%)	9.472 (26.9%)	4.797 (13.6%)	8.908 (25.3%)
2004	14.805 (40.7%)	10.518 (28.9%)	11.812 (32.4%)	6.132 (16.8%)	10.778 (29.6%)
2005	15.667 (42.2%)	11.102 (29.9%)	12.740 (34.4%)	7.292 (19.7%)	12.118 (32.7%)
2006	16.245 (43.4%)	11.812 (31.6%)	14.025 (37.5%)	8.317 (22.2%)	13.345 (35.6%)
2007	17.291 (45.6%)	12.995 (34.3%)	15.563 (41.1%)	9.944 (26.2%)	15.860 (41.8%)
2008	18.949 (49.5%)	14.567 (38.1%)	17.373 (45.4%)	11.443 (29.9%)	17.949 (46.9%)
Variedad 2008 Vs. 2007	+9,6%	+12,1%	+11,6%	+15,1%	+13,2%

Nota: La primera cifra se expresa en miles de individuos y la cifra que aparece entre paréntesis representa el porcentaje de penetración sobre la población de 14 años o más.

Figura 38. Audiencia de Internet. Datos generales.

Por otro lado, según declara Mike Read, Senior Vice President and Managing Director, comScore Europe, casi la mitad de los internautas españoles visitó periódicos online durante el mes de febrero de 2009. El medio online es, sin duda, un mercado muy atractivo para los editores de diarios. Tal y como se aprecia un porcentaje elevado del consumo global de noticias está en Internet.

Otra Información destacada:

- De los 8,4 millones de usuarios únicos a periódicos online por parte de los internautas españoles, el 60% fueron hombres y el 69% eran menores de 45 años.
- Los "Heavy Users" de los periódicos online (defined as the top 20 percent of visitors by time spent in the category), dedicaron una media de 3 horas por usuario a un periódico online durante el mes de febrero '09, periodo de tiempo que supera en más de 4 veces al tiempo dedicado por el usuario de Internet en su conjunto.

Total Unique Visitors to a Selection of Leading Online Spanish Newspaper Titles* Total Spain – Age 15+, Home & Work Locations** February 2009 Source: comScore World Metrix			
Property	Total Unique Visitors (000)	% Reach	Total Pages Viewed (MM)
Total Spanish Internet Audience	18,282	100.0%	40,310
Newspapers	8,393	45.9%	402
Elmundo.es	3,508	19.2%	86
ElPais.com	3,289	18.0%	56
20Minutos.es	1,906	10.4%	37
ABC.es	1,248	6.8%	10
La Vanguardia	825	4.5%	11
ElPeriodico.com	600	3.3%	8
Laverdad.es	469	2.6%	5
Clarín.com	458	2.5%	4
Diariosur.es	374	2.0%	9
LaSprovincias.es	361	2.0%	6

*Based on a selection of individual online newspaper titles; excludes group newspaper properties.

** Excludes traffic from public computers such as Internet cafes or access from mobile phones or PDAs.

Figura 39. Total visitantes únicos de periódicos online españoles.

A la vista de estos datos consideramos que para poder implementar desde las ONG campañas de sensibilización eficaces, es necesario partir de la información –desinformación, en algunos casos– que los ciudadanos reciben. Por este motivo, se propone poner en marcha un **Observatorio de los Medios de Comunicación 2.0** que nos permita conocer **su visión del mundo. Su principal objetivo consistirá en monitorizar y analizar los diarios online nacionales, y crear informes de divulgación con los resultados obtenidos.**

Antecedentes

Durante Diciembre de 2006 y el primer trimestre del año 2007 se realizó la monitorización de una muestra representativa de telediarios de las cadenas nacionales (La2, Cuatro, La Sexta y Tele 5). Para cada uno de los telediarios se extrajeron las noticias relacionadas con el Desarrollo Humano Sostenible y se publicaron en un canal de Youtube.

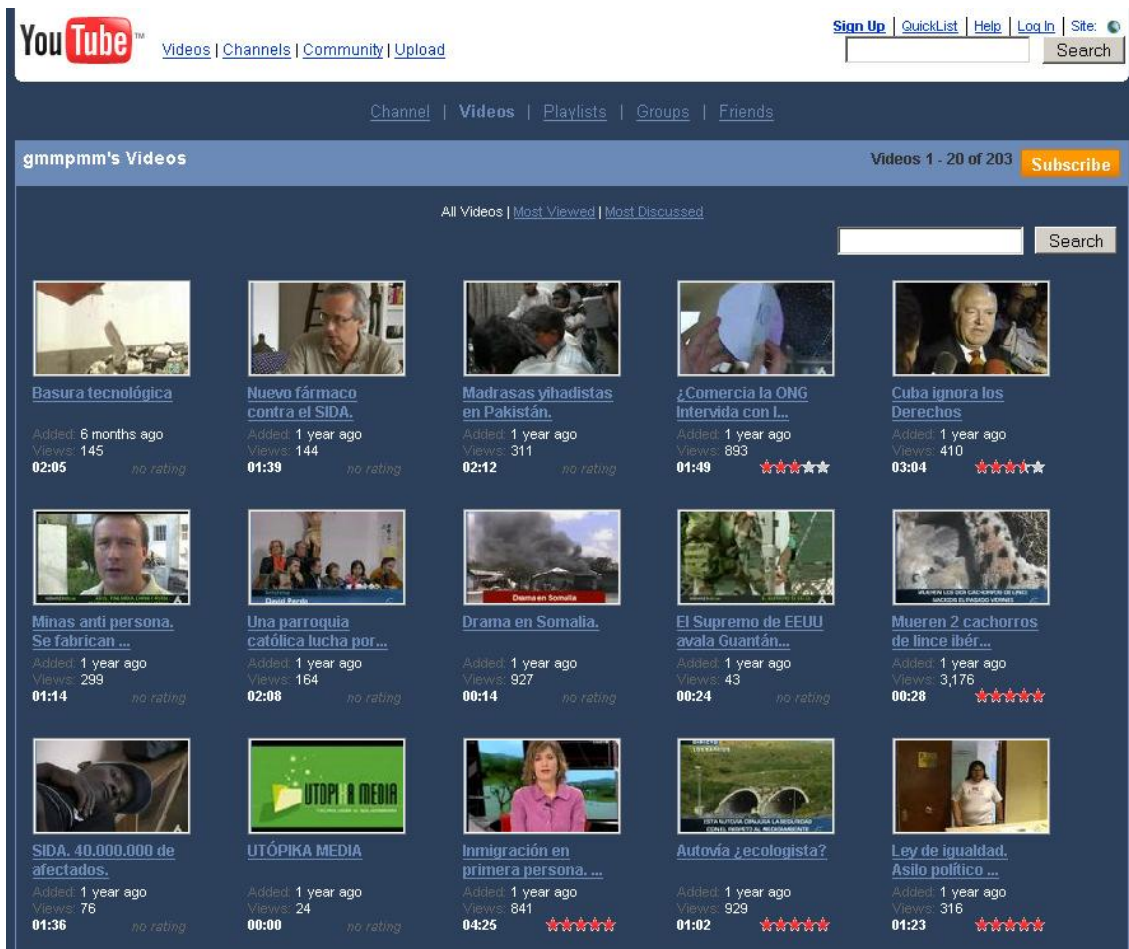


Figura 40. Canal GMMPPM en Youtube (<http://www.youtube.com/gmmpmm>).

Las estadísticas del canal muestran que en total se publicaron 203 noticias. El número total de veces que las noticias fueron visionadas asciende a 243.176 veces.

Las cifras de visualización de las noticias van desde las 16 veces que fue vista la noticia “El drama de Sudán” (La 2, 22:00 horas, 13 de marzo de 2007), hasta las 35.179 veces de la noticia “Cambio climático. Deshielo en el Polo Norte en 34 años” (Noticias Cuatro, 21:00 horas, 13 de diciembre de 2006).

ÍNDICE DE CONTENIDOS

CONTENIDO

Memoria del Proyecto Final de Carrera	0
1. Iniciación.....	1
1.1 Enunciado del trabajo del proyecto	1
1.2 Desarrollo Humano Sostenible	1
1.3 Tecnologías involucradas	4
1.4 Metodología de desarrollo.....	6
2. Planificación.....	7
2.1.1 Objetivos del proyecto.....	7
2.1.2 Contexto del proyecto.....	7
2.1.3 Dependencias del proyecto	8
2.1.4 Especificación del producto.....	8
Fase 1. Puesta en marcha.....	13
Fase 2. Elaborar sección Gráficas	14
Fase 3. Elaborar sección Mapas	14
Fase 4. Elaborar sección Tablas	14
Fase 5. Pruebas.....	14
3. Ejecución.....	16
Fase 1. Puesta en marcha.....	16
Fase 2. Elaborar sección Gráficas	20
Fase 3. Elaborar sección Mapas	29
Fase 4. Elaborar sección Tablas	33
	56

Fase 5. Pruebas.....	38
4. Problemas.....	43
4.1 Problemas con las Gráficas	43
4.2 Problemas con los Mapas	45
4.3 Problemas con las Tablas	45
5. Cierre.....	47
5.1 Descripción de la solución	47
5.2 Objetivos conseguidos.....	49
5.3 Lecciones aprendidas	50
6. Conclusiones	51
Anexo 1: Antecedentes del proyecto.....	52
Índice de contenidos.....	56
Índice de figuras:.....	57

ÍNDICE DE FIGURAS:

Figura 1. Wordpress.....	4
Figura 2. Yahoo Pipes.....	5
Figura 3. Many Eyes, página de inicio.....	12
Figura 4. Secciones del blog.	13
Figura 5. Proceso de desarrollo del proyecto.	15
Figura 6. Ejemplo de pipe.	16
Figura 7. Resultados de un módulo del pipe.	17
Figura 8. Servidor de alojamiento 000webhost.com.....	18

Figura 9. 9ths current Preview	19
Figura 10. FeedWordpress, sección Syndication.	21
Figura 11. FeedWordpress, sección <i>Feeds & Updates</i>	22
Figura 12. FeedWordpress, Sección Feeds & Updates.	23
Figura 13. FeedWordpress, sección Posts & Links.....	23
Figura 14. FeedWordpress, sección Posts & Links 2.....	24
Figura 15. FWP Limit Size of Posts, configuración.	24
Figura 16. Pipe Rss Filtrados. Filtra las noticias por palabras.	25
Figura 17. Pipe Gráfica+noticias.....	26
Figura 18. Entrada fija.....	27
Figura 19. Nueva tarea <i>Cron Job</i>	29
Figura 20. Pipe Auto-Geo. Para geo-localizar las noticias en el mapa.	30
Figura 21. Pipe Mapa diarios de España.....	31
Figura 22. Mapa diarios de España. Resultado capturado del blog.	32
Figura 23. Mapa diarios Reino Unido. Resultado capturado del blog.	33
Figura 24. Pipe de Países. Devuelve los países de las noticias y su recuento.....	34
Figura 25. Exec-PHP, configuración.....	35
Figura 26. Tabla de recuento de países. Resultado capturado del blog.....	38
Figura 27. Ejemplo de gráfica.	39
Figura 28. Tabla en Many Eyes.	40
Figura 29. Many Eyes, Bar Chart.....	41
Figura 30. Many Eyes, Tag Cloud.	41
Figura 31. Many Eyes, Pie Chart.	42

Figura 32	Many Eyes, World Map.....	42
Figura 33.	Sección Gráficas.....	47
Figura 34.	Sección Mapas.....	48
Figura 35.	Sección Tablas de datos.....	48
Figura 36.	Sección Acerca de.....	49
Figura 37.	Cuota de pantalla de informativos de Prime Time.....	52
Figura 38.	Audiencia de Internet. Datos generales.	53
Figura 39.	Total visitantes únicos de periódicos online españoles.....	54
Figura 40.	Canal GMMPPM en Youtube (http://www.youtube.com/gmmpmm).	55