



CCADET

CENTRO DE CIENCIAS APLICADAS Y DESARROLLO TECNOLÓGICO
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Informe Técnico

SISTEMA DE SELECCIÓN DE NOTICIAS PARA LA DIVULGACIÓN

Autores

Josefina Bárcenas López
José Antonio Domínguez Hernández
Alejandra Alvarado Zink

Fecha

Noviembre 2017

Tipo de Proyecto
CONACYT 208558

SISTEMA DE SELECCIÓN DE NOTICIAS PARA LA DIVULGACIÓN

Josefina Bárcenas López
José Antonio Domínguez Hernández
Alejandra Alvarado Zink*

Centro de Ciencias Aplicadas y Desarrollo Tecnológico
Dirección General de Divulgación de la Ciencia*

Resumen

El presente trabajo presenta el diseño y construcción de un sistema capaz de recabar información de las principales fuentes y blogs de investigadores que comparten noticias y descubrimientos en sus sitios de Internet. El sistema fue diseñado con la finalidad de tener un espacio electrónico que automatizará la publicación de noticias de ciencia y tecnología, uno de los objetivos del proyecto Temas de Ciencia Contemporánea. Se hace uso de los sistemas RSS (Really Simple Syndication) y redes sociales de los sitios web de instituciones, revistas y otros espacios que sean de interés científico y publican información validada por la comunidad científica. Se muestra el diseño y desarrollo del sistema, el cual se dividió en tres etapas

- Planeación
- Análisis de requerimientos
- Diseño

El proyecto está elaborado bajo el Proceso Unificado de Desarrollo de Software (RUP, Rational Unified Process), metodología que por su característica de desarrollo y dada la naturaleza del sistema, fue la que mejor se adaptó a este proyecto.

Índice +

Introducción	1
1. Los medios de difusión en Internet.....	2
1.1 Para qué sirve el RSS	2
1.2 Definición de RSS.....	2
1.3 Qué es realmente en RSS.....	2
1.4 Archivos RSS.....	2
1.5 Por qué RSS	2
2. Archivos de Sintaxis XML.....	2
2.1 Codificación de los caracteres	3
2.2 Elementos	3
2.3 Raíz de un documento	3
2.4 Utilización de XML	4
3. Un lenguaje de programación para la web: PHP	5
3.1 XML en PHP	5
3.2 Simple XML.....	5
3.3 Funciones y métodos	5
3.4 Implementando PHP	6
4. Plan de proyecto de software	6
4.1 Funciones principales del proyecto	6
Detección de los procesos.	6
4.2 Análisis de riesgos	7
4.3 Requerimientos del software.....	8
4.3.1 Características del sistema	8
4.3.2 Contenido del sistema	8
4.4 Administración de las fuentes de contenido	8
4.4.1 Descripción y prioridad.	8
4.5 Publicación manual.....	9
4.5.1 Descripción y prioridad.	9
4.6 Requerimientos funcionales.	9
4.7 Administración de las noticias.....	9
4.7.1 Descripción y prioridad	9
4.7.2 Secuencia de estímulo/respuesta	9
4.7.3 Requerimientos funcionales.....	9

4.8 Suscriptores	9
4.8.1 Descripción y prioridad	9
4.8.2 Requerimientos funcionales.....	10
5. Modelo Conceptual del SSND	10
6. Modelo de Diseño del SSDN	11
6.1 Diseño de datos	11
6.2 Descripción de las clases	11
6.3 Descripción de la base de datos utilizada.....	12
6.4 Diagrama entidad –relación	12
7. Diseño de software	14
7.1 Descripción del medio ambiente de desarrollo	14
7.2 Descripción de la interfaz	14
7.3 Prototipo de la interfaz	14
8.Requerimientos funcionales del sistema	17
8.1 Requerimientos de desempeño	17
8.2 Requerimientos de seguridad.	17
8.3 Atributos de calidad de software	17
Conclusiones y recomendaciones	17
Agradecimientos.....	18
Referencias bibliográficas	18

Introducción

Es de suma importancia el papel que juega en la sociedad y en el mundo la ciencia y la tecnología, a través de los años observamos nuevos conocimientos y descubrimientos que estas dos disciplinas nos ofrecen. En la actualidad contamos con nuevas formas de comunicación, modernos medios de transporte, modernas computadoras y sofisticados métodos y aparatos médicos. Pero *¿sabemos lo que implica el desarrollo de estos descubrimientos y el costo que tiene el mismo?*, Ante esta pregunta es necesario decir que es importante la divulgación de la ciencia. Es por ello que nacen ideas nuevas de divulgación como: canales de radio, canales de televisión, sitios de Internet, foros etc.

El Proyecto Temas de Ciencia Contemporánea tiene como objetivo principal la socialización del conocimiento a través de la divulgación de la ciencia y la tecnología. La necesidad de contar con un espacio destinado a la divulgación de las noticias más importantes en materia de ciencia y tecnología que las principales fuentes de información y divulgadores de la ciencia publican en Internet. Estar actualizados con la información que en los últimos treinta años se encuentra en todo el espectro de direcciones electrónicas en la web sería sumamente difícil, si así lo vemos, si no existieran en la actualidad las redes sociales o los medios de difusión electrónica como las listas de noticias o los sistemas RSS.

El sistema web SSND (Sistema de Selección de Noticias Para la Divulgación) surge como una propuesta del proyecto Temas de Ciencia Contemporánea (TCC) para concentrar aquella información que es de interés de grupos de expertos en diversos temas de ciencia así como para las comunidades académicas y escolares universitarias que requieren de información científica y noticias de desarrollo tecnológico actualizados.

Este sistema se encarga de recabar las noticias que fuentes de información publican en sus sitios para posteriormente seleccionar las notas más importantes y publicarlas en el sitio Temas de Ciencia Contemporánea. Requiere implementación de RSS por parte de las fuentes de información que comparten sus publicaciones ya que hasta el momento es la técnica más utilizada y apropiada para este tipo de sistemas.

1. Los medios de difusión en Internet

1.1 Para qué sirve el RSS

El RSS (*Really Simple Syndication*) facilita la gestión y publicación de información y noticias webs. Es una forma estandarizada de distribución de la información de las páginas en Internet a los lectores de la web. Esta información se distribuye a través de las fuentes RSS o Canales RSS. Gracias al RSS, los lectores pueden tener una herramienta útil para mantenerse informados sobre las noticias y páginas electrónicas que le resultan de interés, conservando y almacenando toda la información en un solo lugar que se actualiza de manera automática.

Por otro lado, el RSS brinda al lector un importante ahorro en tiempo en la lectura de información y noticias, ya que sólo con abrir el lector RSS (un programa, navegador o un lector web online), el usuario puede ver cuáles son las últimas actualizaciones y noticias que han publicado las diferentes páginas web a las que está suscrito.

1.2 Definición de RSS

RSS son las siglas de *Really Simple Syndication*, un formato XML para indicar o compartir contenido en la web. Se utiliza para difundir información actualizada frecuentemente a usuarios que se han suscrito a la fuente de contenidos.

1.3 Qué es realmente en RSS

Cuando hablamos de RSS nos referimos usualmente a la tecnología completa para distribución de contenidos de los sitios web. Pero un RSS es realmente un formato de archivo, basado en XML, que sirve para recoger contenidos publicados en páginas web. Los RSS tienen extensión .rss o bien .xml, pero en realidad son un simple archivo de texto donde aparecen referencias a contenidos publicados, en un formato específico, creado a partir de XML.

En el archivo RSS simplemente están los datos de las novedades del sitio, como el título, fecha de publicación o la descripción.

El programa que lee el RSS es el encargado de darle estilo o apariencia a los datos que se incluyan en el archivo y presentarlos de una manera atractiva al usuario y de fácil lectura. El hecho de que RSS sea un formato basado en XML significa que el archivo RSS se compone por una serie de etiquetas definidas que tendrán un formato dado, que respetará las reglas generales de XML.

1.4 Archivos RSS

Los archivos RSS son facilitados por algunos sitios web que generan información específica sobre lo que se haya publicado en su sitio. Cuando se produce alguna actualización en los contenidos de esa web, el archivo RSS se reescribe automáticamente. Así, si se accede el archivo RSS se puede saber que contenidos se han actualizado, pero sin necesidad de entrar en la web. Estos archivos son utilizados para publicar titulares de noticias o entradas de blogs; estos contenidos se pueden integrar en otras webs, estando permanentemente actualizados.

1.5 Por qué RSS

Se hizo la implementación de RSS para el sistema de selección de noticias para la divulgación, por ser la técnica más adecuada, además de ser la más utilizada por sitios y blogs para compartir el contenido, por esta razón el sistema está orientado a funcionar como un lector RSS, pero con la característica de poder escoger que publicaciones de la fuente nos interesan para posteriormente publicarlas directamente en el sitio Temas de Ciencia Contemporánea.

2. Archivos de Sintaxis XML

Un archivo XML (*eXtended Markup Language*) es un archivo de texto que contiene un determinado número de etiquetas estructuradas en árbol. Los elementos del árbol están todos vinculados entre ellos a través de las ramas que se hayan creado (Sebastien y Thierry, 2009).

El archivo XML se organiza en varios componentes que pueden presentarse en forma de árbol (Figura 1).

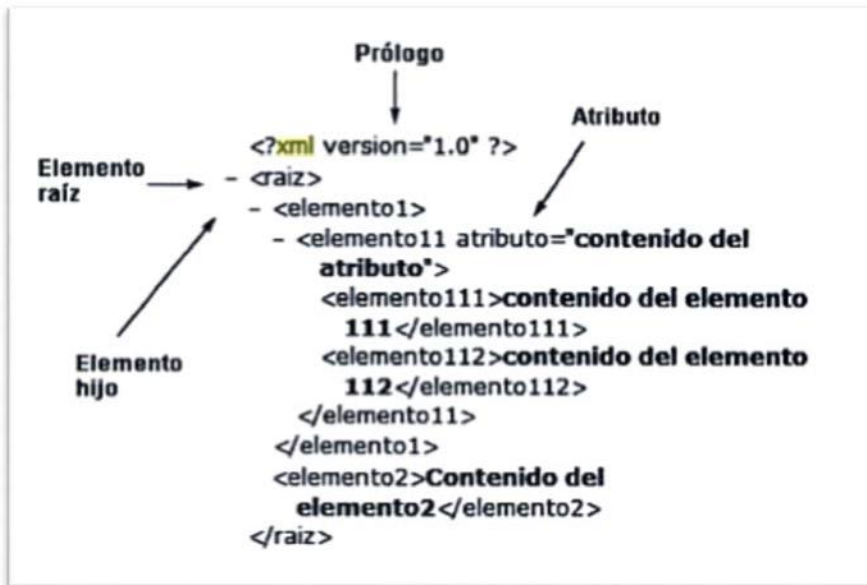


Figura 1. Texto que muestra los componentes de un documento XML

La primera línea de un documento XML se denomina prólogo y contiene, entre otras, la declaración de la versión XML empleada para la descripción de la información. Esta declaración tiene la siguiente forma:

```
<?xml versión="1.0" ?>
```

2.1 Codificación de los caracteres

El prólogo permite igualmente especificar el estándar de la codificación de caracteres utilizados en el documento. La escritura de caracteres XML está conforme a la norma UNICODE, desarrollada para facilitar la interpretación de los archivos. El conjunto de archivos XML presentados en este documento están conforme a la norma ISO 8859-1 correspondientes a la codificación de caracteres para lenguas de Europa occidental (Latin-1). Esta codificación permite utilizar la mayoría de las acentuaciones, razón por la cual fue su elección.

2.2 Elementos

Los elementos constituyen la arquitectura en el árbol de un documento XML. El nombre de los elementos se encuentra en el interior de las etiquetas y su valor se sitúa entre una etiqueta de apertura y una etiqueta de cierre:

```
<NombreElemento> Valor del Elemento </NombreElemento>
```

2.3 Raíz de un documento

El árbol de los elementos que permiten construir el documento XML es el siguiente, todos los documentos XML deben contener una raíz: la raíz es el elemento padre del resto de elementos (Figura 2).



Figura 2. Imagen que ilustra la raíz de un documento como elemento padre del resto de elementos

El siguiente documento XML aplica el anterior esquema. Al elemento padre se le conoce con el nombre de directorio. Da lugar a los elementos de tipo entrada.

```

<?xml version "1.0"?>

<directorio>

<entrada>
  <apellido>BENITEZ</apellido>
  <nombre>Javier</nombre>
  <fechanacimiento />
  <direccion />
</entrada>

<entrada>
  <apellido>MENÉNDEZ</apellido>
  <nombre>Miguel</nombre>
  <fechanacimiento />
  <direccion />
</entrada>

<entrada>
  <apellido>HENDRIX</apellido>
  <nombre>Jimi</nombre>
  <fechanacimiento />
  <direccion />
</entrada>

</directorio>

```

2.4 Utilización de XML

La comprensión de la interpretación de este tipo de archivos fue de suma importancia durante la elaboración de este proyecto ya que la estructura de las fuentes RSS que los sitios y blogs comparten está basada en este formato. De esta manera se implementó un algoritmo hecho en el lenguaje de programación PHP, que extrae la información que contiene la fuente RSS a través de las etiquetas que conforman al archivo (Rossainz, 2017).

3. Un lenguaje de programación para la web: PHP

PHP es el acrónimo de *Hipertext Preprocessor* es un lenguaje de programación de código abierto especialmente utilizado para desarrollo web y que puede ser incrustado en HTML. Una respuesta clara y visual es el siguiente ejemplo.(Figura 3)

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<html>
  <head>
    <title>Example</title>
  </head>
  <body>

    <?php
      echo "Hola, ¡soy un script PHP!";
    ?>

  </body>
</html>
```

Figura 3. Ejemplo de programación en código PHP

3.1 XML en PHP

El lenguaje de programación PHP nos ofrece, en las últimas versiones de su distribución base, una serie de extensiones para manipular documentos en formato XML. Las más populares son: *SimpleXML* y *DOM*.

3.2 Simple XML

Esta extensión se caracteriza por su simplicidad y nos permite, en pocas líneas, generar aplicaciones funcionales. Requiere de PHP versión 5 o superior. Está pensado para ser utilizado en proyectos no demasiado complejos, en los cuales las necesidades tengan que ver con acceder a valores contenidos en los elementos y atributos del documento, de la manera más fácil posible.

3.3 Funciones y métodos

La forma de trabajo consiste en leer el contenido XML, ya sea desde un archivo o desde una variable, y asignarlo a un objeto **simplexml**, el cual nos permitirá un acceso intuitivo a sus partes. SimpleXML provee una serie de métodos predeterminados para este tipo de objetos.

Las funciones disponibles son las siguientes:

- **Simplexml_load_file**
- **Simplexml_load_string**
- **Simple_import_dom**

Y los métodos disponibles:

- **asXML**
- **attributes**
- **children**
- **addAttribute**
- **addChild**

3.4 Implementando PHP

La función de PHP utilizada para este proyecto es **simplexml_load_file** la cual nos permite recuperar el contenido desde un documento en disco o desde una URL de un servidor, el cual es el caso para el sistema de noticias. En el siguiente texto se muestra la declaración del script PHP con el uso de esta función.

```
<?php

$xml = simplexml_load_file ('archivo.xml');

?>
```

En la Figura 4 se muestra el código del sistema donde se aprecia la utilización de dicha función.

```
//contenido de la fuente RSS
$mispost = simplexml_load_file($row2['xml']); //declaramos la ubicacion del XML
echo "<ul>"; // empieza la lista
foreach ($mispost->channel->item as $post):
    $title=$post->title; // declaracion de variables
    $link=$post->link;
    $description=$post->description;
```

Figura 4. Código del sistema que muestra la función simplexml_load

4. Plan de proyecto de software

4.1 Funciones principales del proyecto

El sistema a desarrollar realizará las siguientes funciones principales, las cuales se irán detallando conforme avance la etapa de requerimientos del sistema.

Detección de los procesos.

1. Configuración de fuentes de contenido RSS.
2. Automatización de la búsqueda de noticias de ciencia y tecnología.
3. Selección de las noticias de interés.
4. Publicación de las noticias seleccionadas en el sitio web Temas de Ciencia
 - a. Contemporánea.
5. Suscripción de usuarios.

Configuración de fuentes de Contenido RSS: El Responsable podrá asignar o eliminar las fuentes de contenido que dese que alimenten al sistema.

Automatización de la búsqueda de noticias de Ciencia y Tecnología: El responsable podrá seleccionar de una lista de fuentes, las deseadas para la extracción de las noticias.

Selección de las Noticias de interés: El responsable podrá seleccionar de la información recabada a través de las fuentes de información asignadas, las noticias que le parezcan importantes.

Publicación de las noticias seleccionadas en el sitio de internet Temas de Ciencia Contemporánea: El sistema podrá publicar las noticias que el responsable ha seleccionado como importantes para su divulgación en el sitio web Temas de Ciencia Contemporánea.

Suscripción de usuarios: El sistema podrá suscribir usuarios que deseen recibir las noticias más relevantes que se publican en el sitio web Temas de Ciencia Contemporánea, a través de su correo electrónico, cuenta de Facebook o Twitter.

4.2 Análisis de riesgos

Se muestran los riesgos posibles del proyecto a desarrollar. Se determina una categoría para cada uno de ellos, la probabilidad de que ocurra, determinada de forma empírica, y el impacto del riesgo: (1: catastrófico, 2: crítico, 3: marginal, 4: despreciable). Tabla I.

Riesgos	Categoría	Impacto
El usuario final tiene expectativas a las que tiene el equipo de desarrollo	CC	1
Fracaso del proyecto debido a la falta de experiencia del equipo de desarrollo	EP	1
El cliente cambia requisitos	CC	2
Dificultad en la adaptación del nuevo software por parte del usuario final	CC	3
Falta de planeación y definición en el proceso de ingeniería de software	EP	1
Falta de fondos para completar el proyecto	EP	1
Falta de compromiso del personal hacia el proyecto	EP	1
La agenda nos es realista	EP	2
No hay tiempo para hacerlo "bien"	EP	1
Los requerimientos no son estables ni entendibles	CC	1
Los tiempos de desarrollo son inflexibles y limitados	EP	2
Depende de la integración de herramientas de desarrollo	EP	3

Tabla I. Análisis de riesgos del proyecto

Para la tabla de riesgos, las categorías corresponden:

EP: categoría genérica tamaño y experiencia de la plantilla del personal.

CC: corresponde a características del cliente.

4.3 Requerimientos del software

4.3.1 Características del sistema

En esta sección se enumeran las características que tendrá el sistema obtenidos de la ingeniería de requerimientos. A cada característica se le asigna una prioridad, la cual va en función de que tan necesaria es la característica descrita para el usuario final del sistema. Los valores de esta prioridad serán alta, media y baja; alta si la característica es de suma importancia para el funcionamiento del sistema y la resolución del problema planteado. Se describe una secuencia de estímulo/respuesta a grosso modo. Así también para cada característica se listan los requerimientos funcionales necesarios para cubrirla.

4.3.2 Contenido del sistema

Descripción y prioridad. El sistema será capaz de almacenar y dar de baja las fuentes que proveen el contenido RSS al sitio web. La prioridad es Alta.

Secuencia de estímulo/respuesta. El responsable solicita al sistema realizar un ingreso o una eliminación de un contenido, proveniente de una fuente ya sea un sitio web o blog personal. El sistema deberá de almacenar la dirección del contenido RSS del sitio o blog personal.

Requerimientos funcionales. El sistema permitirá ingresar contenido RSS.- Cuando el responsable dese ingresar información publicada en un sitio web o blog personal al sistema, lo podrá hacer ingresando la URL del contenido RSS proveniente de dicho sitio o blog personal.

El sistema permitirá eliminar contenido RSS.- Cuando el responsable dese eliminar la información proveniente de un sitio web o blog personal, este lo podrá realizar seleccionando el nombre del sitio o blog de la lista que provee de información al sistema.

4.4 Administración de las fuentes de contenido

4.4.1 Descripción y prioridad.

El sistema será capaz de permitir o no la publicación de noticias, de forma temporal o permanente. La prioridad es Alta.

- Secuencia de estímulo/respuesta
- El responsable solicita al sistema quitar una fuente de contenido ya sea de forma temporal o permanente. El sistema quita la fuente de contenido, para dejar de recibir información al sistema.
- Requerimientos funcionales
- Selección de fuentes de contenido.- El sistema permitirá seleccionar que fuentes de contenido desea visualizar, las fuentes seleccionadas podrán alimentar al sistema de las noticias más relevantes del sitio origen.
- Quitar fuente de contenido.- El sistema podrá quitar de forma temporal las fuentes de contenido que en un momento dado no deseemos que publique noticias en el sistema.
- Eliminación de fuente de contenido.- El sistema podrá eliminar de forma permanente la fuente de contenido del sistema.

4.5 Publicación manual

4.5.1 Descripción y prioridad.

El sistema será capaz de publicar noticias que carecen de conexión RSS, mediante el llenado de un formulario. La prioridad es Alta.

- Secuencia de estímulo/respuesta
- El *responsable* solicita al sistema la publicación manual de una noticia, el sistema muestra el formato de llenado para la publicación.

4.6 Requerimientos funcionales.

Publicación de Noticia Manual. - Se podrán publicar noticias de forma manual, a través del llenado de los campos correspondientes al formato; Título, contenido, fuente original, imagen.

4.7 Administración de las noticias.

4.7.1 Descripción y prioridad

El sistema tendrá la característica de seleccionar solo las noticias que desea publicar en el sitio web Temas de Ciencia Contemporánea, así como posteriormente las podrá quitar de las publicaciones. La prioridad es Alta.

4.7.2 Secuencia de estímulo/respuesta

El responsable solicita al sistema de selección de noticias la publicación o el retiro de una noticia, esto lo realiza seleccionando la nota de interés mediante la opción correspondiente publicar o quitar. El sistema manda o retira la publicación al sitio web de Temas de Ciencia Contemporánea/noticiarios.

4.7.3 Requerimientos funcionales

- *El sistema permitirá la selección de noticias.* El responsable podrá seleccionar la o las noticias provenientes del contenido RSS del sitio web o blogs de investigadores, para su publicación en el sitio web Temas de Ciencia Contemporánea.
- *El sistema permitirá quitar las noticias publicadas.* El sistema podrá quitar las noticias que posteriormente han sido publicadas, en la página de noticiarios del sitio web Temas de Ciencia Contemporánea.

4.8 Suscriptores

4.8.1 Descripción y prioridad

El sistema contará con suscriptores que posean una cuenta de Facebook o Twitter. La prioridad es media

- *Secuencia de estímulo/respuesta.* El sistema recibe la petición de suscripción mediante el botón representativo del servicio ya sea Facebook o Twitter. El usuario se firma con el nombre de usuario y contraseña del servicio al que pertenece.

4.8.2 Requerimientos funcionales

El sistema permitirá la subscripción de usuarios.- El sistema podrá subscribir usuarios para recibir a través de su cuenta de Facebook o Twitter, las noticias más importantes publicadas en el sitio web Temas de Ciencia Contemporánea.

5. Modelo Conceptual del SSND

El diagrama de clases de la Figura 5 muestra la estructura estática (diagrama de clases) del Sistema de Selección de Noticias para su Divulgación. La notación utilizada en este diagrama se describe en la Figura 6.

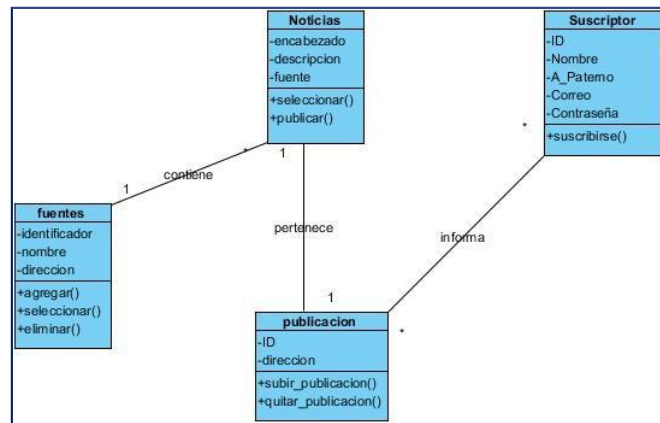


Figura 5. Diagrama de Clases del Sistema de Selección para la Divulgación de Noticias

ClaseB -ListaAtributos +ListaMetodos()	Clase. Se representa por medio de un rectángulo dividido en tres partes. La primera almacena el nombre de la clase, la segunda parte enlista los atributos que forman parte de la clase, estos atributos tienen asignada una visibilidad, la cual indica si el atributo puede ser accedido por otras clases (+) o solo por la clase que los define ().
	Asociación. Indica que existe una relación estática entre clases. La flecha y la etiqueta establecen la dirección y el nombre de la asociación. Los símbolos al inicio y al final de la asociación indican la multiplicidad. La multiplicidad establece el número de objetos que tienen que ver en la asociación.

Figura 6. Notación del Diagrama de clases de SSDN

6. Modelo de Diseño del SSDN

6.1 Diseño de datos

La parte medular del sistema, la forma, la estructura que almacenará los datos (la base de datos) y la estructura que definirá la funcionalidad del sistema (las clases, sus datos y métodos) (Figura 7).

6.2 Descripción de las clases

La mayoría de las clases que usará el sistema fueron extraídas de la etapa de requerimientos y análisis además existen otras clases que fueron incluidas para un buen funcionamiento del sistema. Algunas son utilitarias a las clases existentes o funcionan como una clase unificadora de las funcionalidades que otras clases ofrecen (Tabla II y Tabla III).

Clase administración_del_sistema

Datos miembro	
Fuente	Almacena el nombre de la fuente que comparte la información.
XML	Almacena la dirección del archivo XML que contiene la información que comparte dicha fuente.
Contenidonoticia	Almacena la descripción de la noticia que contiene el archivo XML de la fuente original.
Titulo	Almacena el título de la noticia que contiene el archivo XML de la fuente original.
Link	Almacena el link de la fuente original
Imagen	Almacena la URL de la imagen representativa de la noticia.
Métodos miembro	
maneja_ContenidoRSS	Ingresar el nombre y la URL del archivo XML de la fuente indicada.
maneja_SeleccionRSS	Activa para la fuente especificada mediante un campo a TRUE.
maneja_QuitarRSS	Desactiva la fuente especificada mediante un campo a FALSE.
maneja_EliminaFuenteRSS	Elimina permanentemente la fuente del sistema.
maneja_publicacion	Ingresar al sistema la noticia seleccionada llenando los campos correspondientes (título, descripción, link, imagen), del archivo XML de la fuente original.
Conexión	Realiza la conexión a la bases de datos

Clase Selección_Noticias

Datos miembro	
Titulo	Almacena el título de la noticia contenido del archivo XML de la fuente original.
contenidonoticia	Almacena la descripción de la noticia contenida en el archivo XML de la fuente original.
Link	URL de la noticia original
Imagen	URL de la imagen de la noticia.
Métodos miembro	
InsertarNoticia	Ingresar la noticia especificada al sistema con los datos correspondientes (título, descripción, link, imagen) al archivo XML de la fuente original, en la tabla de publicaciones.

Tabla II. Descripción de las clases

Clase Quitar_Noticia	
Datos miembro	
Título	Almacena el título correspondiente a la noticia indicada.
Métodos miembro	
quitarNoticia	Quita la noticia de la tabla publicaciones de forma permanente.
Clase Noticias_Inicio	
Métodos miembro	
Fuente	Llama el nombre de la fuente que comparte el contenido RSS
XML	Llama la dirección web del archivo XML de la fuente original.
Métodos miembro	
Impresión_noticias	Muestra las noticias de las fuentes de contenido activadas en el sistema.
Clase Conexión	
Datos miembro	
Servidor	Almacena el nombre del servidor de base de datos.
db	Almacena el nombre de la bases de datos.
Usuario	Almacena le nombre de usuarios de la bases de datos.
Contraseña	Almacena la contraseña para el usuario correspondiente.
Métodos miembro	
Conexión	Ejecuta la conexión con el sistema manejador de bases de datos.
Select_db	Selecciona la base de datos indicada para la conexión.
Clase login	
Datos miembro	
nombre	Almacena le nombre de usuario.
Clave	Almacena la clave del usuario.
Métodos miembro	
validación	Ejecuta la consulta a la bases de datos con los datos indicados para la validación de usuario.

Tabla III. Descripción de las clases

6.3 Descripción de la base de datos utilizada

El sistema de selección de noticias para la divulgación deberá almacenar las publicaciones de las fuentes que lo alimentan para poder publicarlas posteriormente en el sitio web Temas de Ciencia Contemporánea para su divulgación, es por ello que es necesario la implementación de una bases de datos. Así mismo las direcciones de los archivos que comparten las noticias de las fuentes originales son almacenados en la bases de datos para poder compartir el contenido. (Tabla III,IV,V)

6.4 Diagrama entidad –relación

En las Tablas IV, V y VI, podemos observar la relación entre las tablas y los campos que la conforman.

Tabla administrador

Nombre del campo	Tipo	Descripción
Usuario	Cadena variable	Es el nombre del usuario administrador
Clave	Cadena variable	Es la contraseña para al usuario administrador.

Tabla IV. Descripción de contenidos de tablas

Tabla contenido RSS

Nombre del campo	Tipo	Descripción
Id	Entero	Identificador único para la fuente de contenido RSS
Fuente	Cadena variable	Es el nombre de la fuente que comparte el contenido
xml	Cadena variable	Es la dirección del archivo XML de la fuente original donde se encuentran las noticias que comparte.
estado	Cadena variable	Indica al sistema el estado actual para la fuente true= activada y false= desactivada.

Tabla V. Descripción de contenidos de tablas

Tabla de publicaciones

Nombre del campo	Tipo	Descripción
Id	Entero	Identificador único para la noticia seleccionada.
Descripción	Cadena variable de longitud larga	Almacena la descripción de la noticia que contiene el archivo XML de la fuente original.
Título	Cadena variable	Almacena el título de la noticia que contiene el archivo XML de la fuente original.
Link	Cadena	Almacena la dirección web de la fuente original de la noticia.
Imagen	Cadena variable	Almacena la dirección de la imagen representativa de la noticia.
Estado	Cadena variable	Indica el estado de la noticia (activada, desactivada).

Tabla VI. Tablas y campos de la bases de datos.

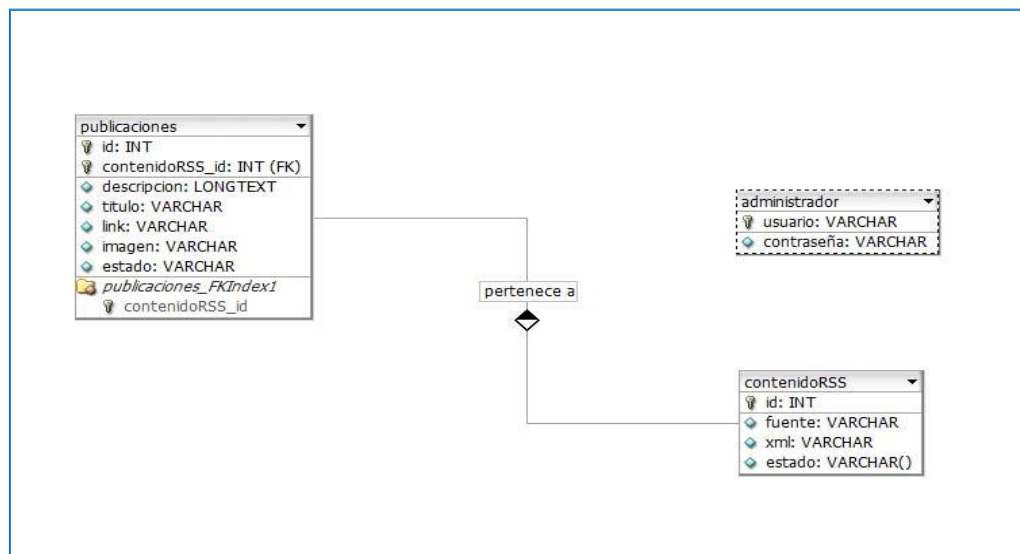


Figura 7 Diagrama entidad-relación de la base de datos del sistema

7. Diseño de software

7.1 Descripción del medio ambiente de desarrollo

La base de datos que soportara el sistema estará diseñada a través de MySQL. Es una herramienta que sirve para generar bases de datos. Para desarrollar los componentes del sistema (las clases, métodos, etc.) se utilizará el lenguaje de programación php. Sobre este mismo software se realizará la comunicación con la base de datos de MySQL.

El sistema de selección de noticias para la divulgación trabajara bajo el servicio apache que se instalara en una plataforma Linux Centos 6.2. Utilizando CSS, HTML y JavaScript se realizara la interfaz gráfica de usuario.

A continuación se describe brevemente el software necesario para desarrollar el sistema de selección de noticias.

- *Sistema operativo Linux Centos 6.2:* Es una distribución Linux que pretende proporcionar una conexión de clase empresarial plataforma informática que tiene 100% de compatibilidad binaria con su ascendente fuente, Red Hat Enterprise Linux.
- *Servidor web apache versión 2.2.15.* El servidor Apache es un software (libre) HTTP de código abierto para plataformas Unix (BSD, GNU/Linux, etc.), Windows, Macintosh y otras, que implementan el protocolo HTTP/1.1.
- *Lenguaje de programación PHP versión 5.2.6:* Es un lenguaje de programación interpretador, diseñado originalmente para la creación de páginas web dinámicas. Es usado principalmente en interpretación del lado del servidor.
- *Navegador Web Mozilla Firefox versión 3.0.1 o superior:* Es un navegador de internet desarrollado por la Corporación Mozilla y un gran número de voluntarios externos. Firefox, oficialmente abreviado como Fx o fx y comúnmente como FF.
- *Editor de textos Gedit:* Es un completo editor de textos libre que se distribuye junto al gestor de escritorio GNOME para sistemas tipo Unix. Este editor se caracteriza principalmente por su facilidad de uso, conseguida en gran parte gracias a una interfaz gráfica clara y limpia
- Software necesario, en la computadora servidor, para ejecutar el sistema.

Sistema operativo Linux Centos 6.2.
Servidor Web Apache versión 2.2.15.
Manejador de Bases de datos MySQL versión 5.0
Lenguaje de programación PHP versión 5.2.6.
Navegado web Mozilla Firefox versión 3.0.1 o superior.

7.2 Descripción de la interfaz

De acuerdo al análisis de los requerimientos del usuario y del sistema se puede construir un primer prototipo de la interfaz del sistema a desarrollar, el cual muestra las características principales que el sistema debe poseer.

7.3 Prototipo de la interfaz

A continuación se muestran algunas de las pantallas del prototipo de la interfaz del sistema a desarrollar. Las pantallas que no se incluyeron son muy similares.

Pantalla principal del Sistema. Muestra la pantalla principal del sistema sin contenido de noticias cargado. Se puede apreciar la barra de menú del sistema en la parte superior derecha (Figura 8).

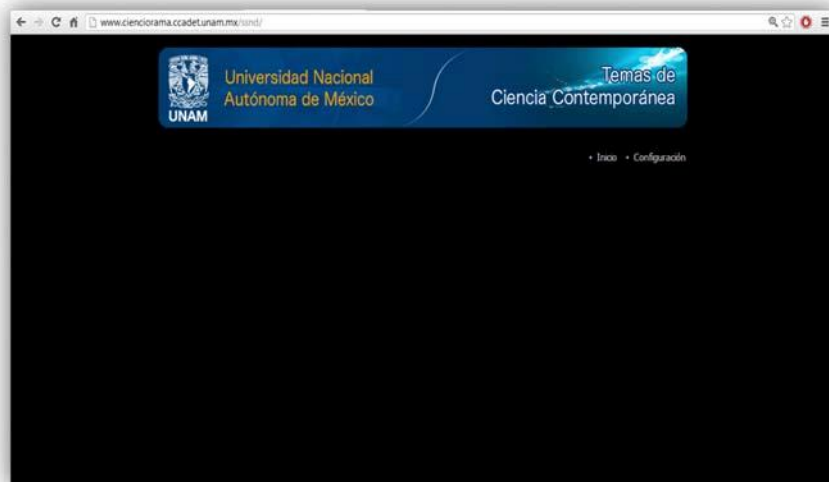


Figura 8. Pantalla principal del sistema sin información

Pantalla de acceso a configuración el sistema. Muestra la pantalla de acceso para la configuración del sistema, contiene dos campos de texto correspondientes a usuario y contraseña. (Figura 9)



Figura 9. Configuración del SSDN9

Menú de configuración del sistema. El menú principal de configuración del sistema, en ella podemos ingresar a cada una de las acciones que puede realizar el sistema. La página de configuración del sistema, en ella se carga el contenido, así como la selección de las fuentes que podrán publicar noticias. Para el caso de las fuentes que no cuentan con RSS se tiene la publicación manual. (Figura 10)

Figura 10. Menú de Configuración SSDN

Menú Selección de noticias. En la Figura 11 se observa la noticia proveniente de una fuente RSS, con la característica de edición, en esta opción el usuario podrá editar las noticias (solo para cuestiones de formato) y seleccionara las noticias de interés para su publicación.

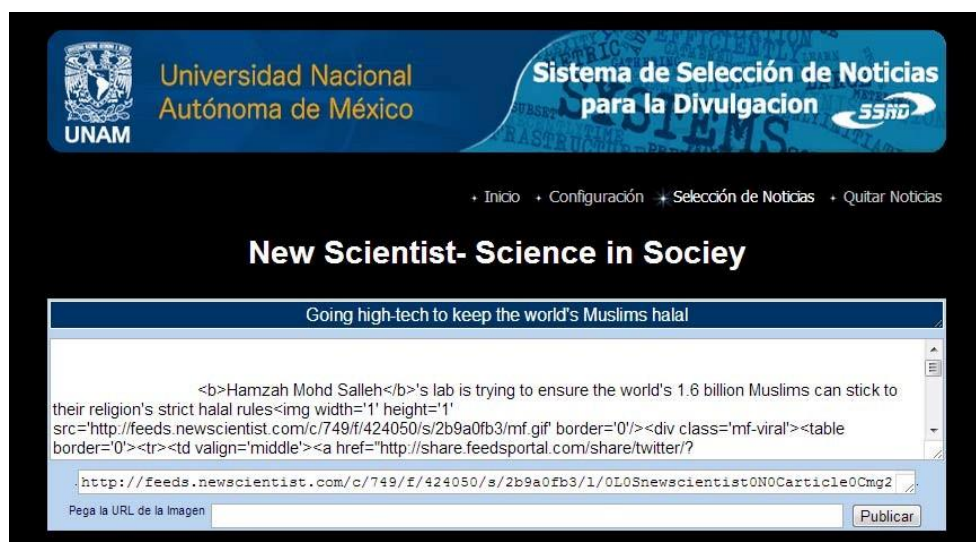


Figura 11. Página de selección de noticias

Menú Quitar Noticias. Las noticias que ya están publicadas en la página de noticiarios del sitio web Temas de Ciencia Contemporánea. Se podrán retirar de las publicaciones mediante el botón correspondiente (botón quitar). (Figura 12)



Figura 12. Eliminar noticias del SSDN

8. Requerimientos funcionales del sistema

8.1 Requerimientos de desempeño

Se utilizan para verificar el desempeño del sistema con respecto al tiempo de ejecución dentro del contexto de un sistema integrado.

El sistema de selección de noticias para la divulgación, no requiere de tiempos de respuesta crítica para la presentación de resultados. El tiempo de ejecución del sistema será aceptable para los requerimientos funcionales solicitados por el cliente. Por lo que no se implementará ningún mecanismo para la medición del desempeño.

8.2 Requerimientos de seguridad.

Estos requerimientos deciden la implementación de mecanismos de protección integrados en el sistema para que realmente lo protejan de irrupciones inapropiadas.

En el sistema se manejaran dos tipos de acceso. El tipo “usuario” y el tipo “administrador o responsable”, el cual contara con usuario y contraseña. El primero solo podrá ver las publicaciones de las noticias que el responsable ha seleccionado. El segundo tendrá acceso a todas las funciones implementadas en el sistema.

8.3 Atributos de calidad de software

Facilidad de uso.- El sistema será diseñado con el fin de facilitar su manejo por parte del usuario. Es por ello que no se implementan funciones redundantes o campos de datos que solo obstaculicen la agilidad en la funcionalidad del sistema.

Facilidad de acceso a los datos.- El usuario podrá acceder de una forma clara y rápida a toda la información deseada para el manejo del sistema.

Portabilidad.- El sistema podrá ser instalado en cualquier computadora con las mismas características en hardware.

Facilidad de entender al sistema.- El sistema estará organizado de manera que le usuario pueda interpretar rápidamente cada función del mismo.

Conclusiones

La implementación de RSS en los sitios web y Blog en Internet, son de suma importancia, ya que este medio permite de una manera fácil y sencilla compartir información. Cabe mencionar que existen diferentes formas de poder hacer uso de este tipo de tecnologías, sin embargo la forma en que se ha implementado este sistema de selección de noticias puede ser orientado a otro tipo de divulgación de la información quedando como una herramienta útil para cualquier tipo de negocio o empresa, que requiera consultar información de diferentes sitios web y que desee recabarla en una sola página.

El SSND es un desarrollo realizado para la operación y gestión de noticias científicas para el proyecto Temas de Ciencia Contemporánea. Principalmente su funcionamiento se ha orientado para mantener al grupo de académicos que participan en el proyecto, informados sobre temas actuales y de interés científico.

Agradecimientos

Agradecemos al CONACYT el apoyo para el desarrollo del proyecto 208558 y al Pas. Uriel Aguirre Monroy quien participó en el proyecto Temas de Ciencia Contemporánea y colaboró en el CCADET en el desarrollo informático de este software.

REFERENCIAS BIBLIOGRÁFICAS.

Rossainz, L.M. (2017) *Ingeniería Web. Metodologías para el desarrollo de software basado en web*. [Internet] México. Benemérita Universidad de Puebla, Fac. de Ciencias de la Computación. Disponible desde http://rossainz.cs.buap.mx/ingWeb/slides/2_IngWeb_Metodologias.pdf [Noviembre 2017]

RSS (2016) *RSS Explicado* [Internet] España. Disponible desde <http://www.rss.nom.es/> [19 de Enero de 2016]

Sebastien L., Thierry B. (2009) *XML practico Bases esenciales conceptos y casos prácticos*. París, Ed. ENI