

Universidad de las Ciencias Informáticas

Facultad 4



Título: Módulo Búsqueda para el Portal Educativo

**Trabajo de Diploma para optar por el título de
Ingeniero en Ciencias Informáticas**

Autor: Dayana Hernández Chi

Tutor: Ing. Orlando Felipe Salvador Broche

La Habana

“Año 55 de la Revolución.”

Junio, 2013

PENSAMIENTO

Pensamiento

“Internet facilita la información adecuada, en el momento adecuado, para el propósito adecuado.”

Bill Gates

DECLARACIÓN DE AUTORÍA

DECLARACIÓN DE AUTORÍA

Declaro que soy el único autor de este trabajo y autorizo a la facultad 4 de la Universidad de las Ciencias Informáticas a hacer uso del mismo en su beneficio.

Para que así conste firmo la presente a los ____ días del mes de _____ del año _____.

Dayana Hernández Chi

Ing. Orlando Felipe Salvador Broche

Firma del Autor

Firma del Tutor

DATOS DE CONTACTO

DATOS DE CONTACTO

AGRADECIMIENTOS

AGRADECIMIENTOS

Primeramente le doy gracias a mi Dios, porque a Él le debo todo lo que tengo y todo lo que soy y si hasta aquí he llegado es por ti Señor, infinitas gracias.

Agradezco a mis padres por darme la vida y porque siempre han estado a mi lado, gracias por su apoyo.

Gracias a "nana" y "nano" por ser mis hermanos y estar a mi lado.

Gracias a mi familia por su preocupación y amor hacia mí.

Gracias a Graviel, mi amor gracias por amarme con todas mis virtudes y defectos, y gracias porque con tu amor me has hecho una mejor persona.

Gracias a mis amistades: Dai, has sabido ser mi hermana y amiga; Alien, gracias por tu amistad en todo momento; Leandro, gracias por ser mi amigo en cualquier circunstancia; Carlos, gracias por compartir conmigo cuando necesitaba alguien a mi lado.

A todos mis hermanos de la UCI gracias por su apoyo y amistad, no me he sentido sola en esta escuela porque he tenido una gran familia a mi lado que me ha ayudado a crecer como persona y como cristiana; a todos ustedes mis más profundos agradecimientos y sepan que siempre existirá un pedazo importante en mi corazón donde ustedes estarán.

A cada uno de mis compañeros y compañeras de aula, apartamento y proyecto, por soportarme y aceptarme, doy gracias por cada uno de ustedes.

A todo los que me han ayudado con mi tesis: a mi tutor Orlando por su apoyo, al líder del proyecto Javier porque me ayudó cuando tenía dudas, a Delvis por revisarme el documento, al tribunal y la oponente por su apoyo y sugerencias que permitieron que esta tesis fuera mejor.

DEDICATORIA

DEDICATORIA

Al Espíritu Santo, por ser mi Dios, mi parakletos y mi mejor amigo. Toda la gloria sea a ti.

RESUMEN

Con el auge alcanzado en las Tecnologías de la Información y las Comunicaciones se ha hecho necesario la localización y recuperación del gran cúmulo de información existente. Se han creado diferentes herramientas para recuperar la información que han sido desarrolladas y utilizadas en los últimos tiempos, cada una con sus características específicas, para responder a necesidades concretas. Estas herramientas hacen uso de modelos y técnicas para recuperar la información relevante en sitios o documentos.

El proyecto Portal Educativo es un *software* que almacenará una enorme suma de información en los diferentes contenidos publicados, a la que los usuarios del sitio necesitan acceder de forma rápida y eficiente. Por ello se hace necesario crear un módulo que permita a los usuarios de este portal consultar los contenidos de su interés, usando técnicas avanzadas para la recuperación de la información. Para implementar dicho módulo y responder a las necesidades antes planteadas se realizó un estudio de las principales herramientas y técnicas actuales de recuperación de información, así como las últimas tendencias en el desarrollo de portales educativos.

PALABRAS CLAVE

Búsqueda, portal educativo, recuperación de la información.

Índice

INTRODUCCIÓN	5
CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA	9
1.1 INTRODUCCIÓN.....	9
1.2 LA RECUPERACIÓN DE LA INFORMACIÓN EN SISTEMAS DIGITALES.....	9
1.2.1 <i>Técnicas y modelos de recuperación de la información</i>	11
1.2.2 <i>Bibliotecas y herramientas de búsqueda</i>	14
1.2.3 <i>Lenguajes y tipos de búsqueda</i>	18
1.3 PORTALES EDUCATIVOS.....	19
1.3.1 <i>Tendencias actuales. Comunidades virtuales</i>	20
1.3.2 <i>Tecnologías para desarrollar portales educativos. CMS</i>	21
1.4 METODOLOGÍA DE DESARROLLO.....	22
1.4.1 <i>RUP</i>	22
1.5 TECNOLOGÍAS Y HERRAMIENTAS DE DESARROLLO	22
1.5.1 <i>Tecnologías</i>	22
1.5.2 <i>Herramientas de Desarrollo</i>	24
1.6 CONCLUSIONES PARCIALES	25
CAPÍTULO 2: CARACTERÍSTICAS DEL SISTEMA.....	26
2.1 INTRODUCCIÓN.....	26
2.2 PROPUESTA DE SISTEMA	26
2.3 MODELO DE DOMINIO	26
2.4 ESPECIFICACIÓN DE LOS REQUISITOS DE SOFTWARE.....	27
2.4.1 <i>Requerimientos Funcionales</i>	27
2.4.2 <i>Requerimientos No Funcionales</i>	28
2.5 DEFINICIÓN DE LOS CASOS DE USO	29
2.5.1 <i>Definición de actores</i>	30
2.5.2 <i>Diagrama de casos de uso</i>	31
2.5.3 <i>Descripción de casos de uso</i>	32
2.6 CONCLUSIONES PARCIALES	36
CAPÍTULO 3: ANÁLISIS Y DISEÑO DEL SISTEMA	37
3.1 INTRODUCCIÓN.....	37
3.2 ARQUITECTURA	37
3.3 MODELO DE ANÁLISIS	38
3.3.1 <i>Diagramas de Clases del Análisis</i>	39
3.4 MODELO DE DISEÑO	39
3.4.1 <i>Diagramas de Interacción. Diagramas de Colaboración</i>	39
3.4.2 <i>Diagramas de Clases de Diseño</i>	40
3.5 DISEÑO DE LA BASE DE DATOS	42
3.5.1 <i>Modelos de Datos</i>	42
3.6 PATRONES DE DISEÑO	44
3.6.1 <i>Patrones GOF</i>	44
3.7 CONCLUSIONES PARCIALES	45
CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA	46

4.1	INTRODUCCIÓN.....	46
4.2.	IMPLEMENTACIÓN	46
4.2.1	<i>Diagrama de Despliegue</i>	46
4.2.2	<i>Diagrama de Componentes</i>	47
4.3	PRUEBAS DE SOFTWARE.....	47
4.3.1	<i>Pruebas Unitarias</i>	48
4.3.2	<i>Modelo de prueba. Pruebas de Funcionalidad</i>	49
4.4	CONCLUSIONES PARCIALES	58
CONCLUSIONES		59
RECOMENDACIONES.....		61
REFERENCIAS BIBLIOGRÁFICAS		62
BIBLIOGRAFÍA.....		65
ANEXOS.....		66
ANEXO 1: DEFINICIÓN DE CASOS DE USO DEL SISTEMA		66
	<i>Tabla 1: Definición de CU Filtrar por boletín.</i>	66
	<i>Tabla 2: Definición de CU Filtrar por avisos.</i>	66
	<i>Tabla 3: Definición de CU Filtrar por eventos.</i>	66
	<i>Tabla 4: Definición de CU Filtrar por recursos educativos.</i>	66
	<i>Tabla 5: Definición de CU Filtrar por clasificación de eventos.</i>	67
	<i>Tabla 6: Definición de CU Filtrar por clasificación de avisos.</i>	67
	<i>Tabla 7: Definición de CU Filtrar por clasificación de boletín.</i>	67
	<i>Tabla 8: Definición de CU Filtrar por clasificación de recurso educativo.</i>	67
ANEXO 2: DESCRIPCIÓN DE CASOS DE USO DEL SISTEMA.....		67
	<i>Tabla 1: Descripción del Caso de Uso Filtrar por boletín.</i>	68
	<i>Tabla 2: Descripción del Caso de Uso Filtrar por avisos.</i>	68
	<i>Tabla 3: Descripción del Caso de Uso Filtrar por eventos.</i>	69
	<i>Tabla 4: Descripción del Caso de Uso Filtrar por recursos educativos.</i>	70
	<i>Tabla 5: Descripción del Caso de Uso Filtrar por clasificación de eventos.</i>	70
	<i>Tabla 6: Descripción del Caso de Uso Filtrar por clasificación de avisos.</i>	71
	<i>Tabla 7: Descripción del Caso de Uso Filtrar por clasificación de boletín.</i>	72
	<i>Tabla 8: Descripción del Caso de Uso Filtrar por clasificación de recurso educativo.</i>	73
ANEXO 3: DIAGRAMAS DE CLASES DE ANÁLISIS		73
	<i>Figura 1: Diagrama de Clases de Análisis del CU Realizar búsqueda avanzada.</i>	73
	<i>Figura 2: Diagrama de Clases de Análisis del CU Filtrar por avisos.</i>	73
	<i>Figura 3: Diagrama de Clases de Análisis del CU Filtrar por boletín.</i>	74
	<i>Figura 4: Diagrama de Clases de Análisis del CU Filtrar por eventos.</i>	74
	<i>Figura 5: Diagrama de Clases de Análisis del CU Filtrar por recurso educativo.</i>	74
	<i>Figura 6: Diagrama de Clases de Análisis del CU Filtrar por clasificación de eventos.</i>	75
	<i>Figura 7: Diagrama de Clases de Análisis del CU Filtrar por clasificación de avisos.</i>	75
	<i>Figura 8: Diagrama de Clases de Análisis del CU Filtrar por clasificación de comunidades virtuales.</i>	75
	<i>Figura 9: Diagrama de Clases de Análisis del CU Filtrar por clasificación de recurso educativo.</i>	76
	<i>Figura 10: Diagrama de Clases de Análisis del CU Filtrar por clasificación de boletín.</i>	76
ANEXO 4: DIAGRAMAS COLABORACIÓN		76
	<i>Figura 1: Diagrama de Colaboración del CU Realizar búsqueda avanzada.</i>	76
	<i>Figura 2: Diagrama de Colaboración del CU Filtrar por avisos.</i>	76

ÍNDICE

<i>Figura 3: Diagrama de Colaboración del CU Filtrar por boletín.....</i>	<i>77</i>
<i>Figura 4: Diagrama de Colaboración del CU Filtrar por eventos.</i>	<i>77</i>
<i>Figura 5: Diagrama de Colaboración del CU Filtrar por recurso educativo.</i>	<i>77</i>
<i>Figura 6: Diagrama de Colaboración del CU Filtrar por clasificación de eventos.</i>	<i>78</i>
<i>Figura 7: Diagrama de Colaboración del CU Filtrar por clasificación de avisos.</i>	<i>78</i>
<i>Figura 8: Diagrama de Colaboración del CU Filtrar por clasificación de comunidades virtuales.</i>	<i>78</i>
<i>Figura 9: Diagrama de Colaboración del CU Filtrar por clasificación de recurso educativo.</i>	<i>78</i>
<i>Figura 10: Diagrama de Colaboración del CU Filtrar por clasificación de boletín.</i>	<i>79</i>
ANEXO 5: DIAGRAMAS DE CLASES DE DISEÑO WEB	79
<i>Figura 1: Diagrama de Clases del CU Filtrar por avisos.</i>	<i>79</i>
<i>Figura 2: Diagrama de Clases del CU Filtrar por eventos.....</i>	<i>80</i>
<i>Figura 3: Diagrama de Clases del CU Filtrar por boletín.....</i>	<i>81</i>
<i>Figura 4: Diagrama de Clases del CU Filtrar por recurso educativo.....</i>	<i>81</i>
<i>Figura 5: Diagrama de Clases del CU Filtrar por clasificación de comunidades virtuales.</i>	<i>82</i>
<i>Figura 6: Diagrama de Clases del CU Filtrar por clasificación de avisos.</i>	<i>82</i>
<i>Figura 7: Diagrama de Clases del CU Filtrar por clasificación de eventos.</i>	<i>83</i>
<i>Figura 8: Diagrama de Clases del CU Filtrar por clasificación de recurso educativo.</i>	<i>83</i>
<i>Figura 9: Diagrama de Clases del CU Filtrar por clasificación de boletín.....</i>	<i>84</i>
<i>Figura 10: Diagrama de Clases del CU Realizar búsqueda avanzada.....</i>	<i>84</i>
GLOSARIO.....	85

INTRODUCCIÓN

El auge de las Tecnologías de la Información y las Comunicaciones (TIC) en las últimas décadas ha propiciado que se generen grandes volúmenes de información, lo cual trajo como consecuencia la necesidad de crear nuevas técnicas para su localización y recuperación.

Entre los *software* más comunes en Internet para almacenar información se encuentran los portales. Se pueden hallar muchas definiciones para portales informáticos entre ellas que: *“un portal es un sitio Web cuyo objetivo es ofrecer al usuario, de forma fácil e integrada, el acceso a una serie de recursos y de servicios, entre los que suelen encontrarse buscadores, foros, compra electrónica, etc.”* (1).

Existen diferentes tipos de portales, la categorización más difundida sobre estos es la que atiende al tipo de información que albergan. Según esto se clasifican en dos tipos: portales generales (horizontales o megaportales) y portales especializados (verticales o corporativos). (2)

Los portales generales son aquellos cuyo contenido es de índole genérico, es decir que abarcan casi todos los temas posibles de Internet. Estos están dirigidos a un usuario con conocimientos básicos de Internet.

Los portales especializados o temáticos pretenden satisfacer las necesidades de información de una comunidad de usuarios específica. Los portales verticales son sitios *web* que proporcionan información y servicios a una industria en particular. Los portales corporativos o institucionales son sitios *web* que proporcionan información de la empresa o institución a la que pertenecen, fundamentalmente a los empleados de la misma. (2)

Los portales educativos pertenecen a la categoría de portales especializados. Estos se definen como *“un sitio web en el que se incluyen tanto contenidos como servicios y que se brindan a la comunidad educativa en su conjunto. La mayoría de los portales educativos se caracterizan por su desinterés económico siendo patrocinados por instituciones, ministerios y empresas del sector educativo. La naturaleza “En línea (Online)” de estos almacenes de información, permite al usuario acceder a sus contenidos de forma ágil, sencilla e instantánea lo que sumado a la gran calidad del producto final permite que este tipo de estructuras de información constituya una de las mejores soluciones en la construcción del conocimiento.”* (3)

Estos portales persiguen objetivos específicos, como servir a todos los niveles del mundo educativo, especialmente a profesores y estudiantes; constituyen el punto de encuentro de todos los niveles educativos del país de referencia; reúnen el material educativo, servicios y recursos disponibles para la enseñanza. (2) Además en todos estos sitios se puede ver la presencia de un buscador para facilitar la recuperación de la información. Los buscadores son herramientas que localizan información, tanto en un sitio como en la red, en dependencia del fin con el que se hayan desarrollado, para ello usan un concepto denominado recuperación

INTRODUCCIÓN

de la información (RI) definido por Ricardo Baeza-Yates¹ como, “*dada una necesidad de información (consulta) y un conjunto de documentos, ordenar los documentos de más a menos relevantes para esa necesidad y presentar un conjunto de aquellos con mayor relevancia*”. (4)

En el proyecto Portal Educativo del Centro de Tecnologías para la Formación (FORTES) de la facultad 4 de la Universidad de las Ciencias Informáticas (UCI) se almacenará un gran cúmulo de datos en diferentes contenidos como: recursos educativos, comunidades virtuales, eventos, avisos y boletines. Los usuarios que interactuarán con este portal educativo pertenecen a una comunidad virtual donde los miembros tienen intereses profesionales similares, por tanto, la localización de contenidos publicados de un tema determinado es una prioridad para que estos puedan ser consultados mediante una búsqueda eficiente, filtrada por fecha, tipo de contenido y otros criterios que permitan acotar el número de resultados.

A pesar de la necesidad de recuperar los contenidos de interés para los usuarios del portal no se cuenta con ninguna herramienta que facilite la localización de contenidos de interés para los usuarios. Por lo anterior la búsqueda de información sobre un tema específico puede tornarse engorrosa para los que acceden al portal, ya que deben revisar manualmente cada contenido publicado hasta encontrar la información que les interese.

Teniendo en cuenta lo antes planteado se define como **problema de investigación**: ¿Cómo lograr que los usuarios del Portal Educativo puedan de manera rápida y precisa localizar los contenidos de su interés?

Para dar solución al problema antes expuesto se plantea el siguiente **objetivo general**: Desarrollar un módulo que les brinde a los usuarios del Portal Educativo la posibilidad de consultar contenidos de su interés, mediante técnicas avanzadas de recuperación de la información.

Para dar cumplimiento al objetivo general, se definen como **objetivos específicos**:

- Elaborar el marco teórico-conceptual de las herramientas y técnicas más actuales utilizadas en la recuperación de información.
- Identificar las herramientas o bibliotecas informáticas adecuadas para dar solución al problema planteado, teniendo en cuenta la compatibilidad o capacidad de integración con la arquitectura propuesta por el proyecto Portal Educativo.
- Generar los artefactos ingenieriles para el sistema propuesto, según la metodología de desarrollo de *software* Proceso Unificado de Desarrollo (RUP).
- Implementar los requerimientos definidos.

¹ Vicepresidente de Investigación para Europa y Latinoamérica, lidera los laboratorios de Yahoo! Research en Barcelona, España y Santiago, Chile, supervisa además el laboratorio de Haifa, Israel.

INTRODUCCIÓN

- Realizar pruebas al sistema.

Se define como **objeto de estudio**: el proceso de búsqueda y recuperación de información en sistemas informáticos.

Teniendo como **campo de acción**: el proceso de búsqueda y recuperación de información en portales educativos.

Se plantea como **hipótesis** que un módulo de búsqueda para el *software* Portal Educativo, permite que los usuarios del mismo, puedan de manera rápida y precisa, localizar los contenidos de su interés.

Para cumplir con los objetivos de esta investigación se desarrollaron las siguientes **tareas de investigación**:

1. Revisión de la bibliografía actualizada para generar el marco teórico-conceptual.
2. Análisis de las herramientas y bibliotecas actuales utilizadas para la recuperación de información.
3. Identificación de la herramienta y la biblioteca más adecuadas para ser integradas a la arquitectura propuesta por el proyecto.
4. Estudio de las técnicas y algoritmos de indización, con el fin de poder comprender mejor los procesos que ocurren en la herramienta o biblioteca identificada.
5. Estudio de los lenguajes y tipos de búsqueda que existen y especialmente de la herramienta o biblioteca seleccionada.
6. Estudio de los portales educativos, sus tendencias modernas y las tecnologías para su desarrollo actual.
7. Elaboración de los artefactos ingenieriles según la metodología RUP.
8. Implementación de los requerimientos definidos.
9. Realización de pruebas al sistema.

Con el propósito de desarrollar las tareas que se plantearon anteriormente fueron empleados métodos de investigación que permitieron la obtención de información. A continuación se exponen con qué fines fueron utilizados estos métodos:

Métodos Teóricos:

Histórico-lógico: Se utiliza para realizar el estado del arte de los Sistemas de Recuperación de la Información (SRI) y las herramientas y bibliotecas actuales utilizadas para la RI; además se estudiaron diferentes tecnologías y herramientas para el desarrollo del módulo.

Modelación: La modelación es el método mediante el cual se crean abstracciones con el objetivo de explicar la realidad. (5) Se utilizará para el modelado de los casos de uso que respondan a los requisitos funcionales del sistema.

Métodos Empíricos:

INTRODUCCIÓN

Observación externa, no incluida y cerrada: Se observa cómo otros sistemas realizan la RI y qué herramientas, técnicas y bibliotecas utilizan.

El presente trabajo de diploma está compuesto por cuatro capítulos, de los cuales se muestra un resumen a continuación:

Capítulo 1. Fundamentación teórica: Incluye un marco teórico-conceptual del tema tratado, de las tendencias, técnicas, tecnologías, metodologías y *software* usados en la actualidad.

Capítulo 2. Características del sistema: En este capítulo se da una propuesta del sistema, se describe el modelo de dominio, se especifican los requisitos que debe cumplir el módulo, se elabora el diagrama de casos de uso y se especifican los casos de uso.

Capítulo 3. Análisis y diseño del sistema: En este capítulo se elaboran los diagramas de clases de análisis, el diagrama de paquetes, los de interacción, los de clases de diseño y el modelo de datos según propone la metodología empleada. Además se expone la arquitectura y los patrones de diseño empleados.

Capítulo 4. Implementación y prueba: Contiene los diagramas de componentes y de despliegue, además de la realización y descripción de los casos de prueba.

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

1.1 Introducción

En la actualidad por la gran cantidad de información digital que existe se hace necesario crear y perfeccionar técnicas para su localización, de modo que los usuarios de Internet o de un sitio específico puedan localizar la información de su interés de forma rápida. Por ello investigar y conocer los métodos para recuperar la información es necesario para facilitar a los usuarios acceso al contenido que esté buscando.

En el presente capítulo se realiza un estudio sobre las aplicaciones informáticas que gestionan la RI, qué herramientas y bibliotecas se emplean para la RI, además de los modelos y técnicas que estas utilizan para llevar a cabo la RI. Como se trabajará para un portal educativo se deben estudiar sus ventajas, las tendencias vigentes, y las tecnologías que se utilizan en la actualidad para la creación de portales. Con el propósito de lograr un correcto desarrollo de la solución que se va a implementar es necesario definir las herramientas, metodologías y tecnologías que serán utilizadas.

1.2 La recuperación de la información en sistemas digitales

La recuperación de la información es un concepto muy polémico sobre el que existen varias definiciones. Una de ellas plantea que la RI es un *“proceso donde se accede a una información previamente almacenada, mediante herramientas informáticas que permiten establecer ecuaciones de búsqueda específicas. Dicha información ha debido de ser estructurada previamente a su almacenamiento.”* (6). Mientras en esta definición se le trata como un proceso de acceso a la información hay otra que define la RI como *“el conjunto de tareas mediante las cuales el usuario localiza y accede a los recursos de información que son pertinentes para la resolución del problema planteado.”* (7) A pesar de la diferencia en los conceptos se puede decir de forma general que la RI *“tiene por objeto: identificar, localizar, seleccionar y acceder a los recursos de información útiles al usuario.”* (8) Además se define como un Sistema de Recuperación de la Información (SRI) al conjunto de herramientas de búsqueda idóneas para el tipo de información a buscar. (7)

La RI ocupa un papel importante en varias disciplinas como: psicología cognitiva², la arquitectura de la información³, inteligencia artificial⁴, lingüística⁵, semiótica⁶, informática⁷, biblioteconomía⁸, archivística⁹ y

² Como se denota a los sistemas en psicología que se encargan del estudio de la cognición, es decir, los procesos mentales implicados en el conocimiento.

³ Disciplina y arte encargada del estudio, análisis, organización, disposición y estructuración de la información en espacios de información.

⁴ Es la ciencia e ingenio de hacer máquinas inteligentes, especialmente programas de cómputo inteligentes.

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

documentación¹⁰. El presente trabajo hará énfasis solamente en el papel de la RI en el campo de la informática.

La idea del uso de computadoras para la búsqueda de información relevante se popularizó a raíz de un artículo: “*As We May Think*” de Vannevar Bush¹¹, en el año 1945. Los primeros sistemas automatizados de recuperación de la información fueron presentados durante la década de 1950 a 1960. A principios de 1970 comenzaron a utilizarse sistemas de recuperación a larga escala, como el Sistema de Diálogo Lockheed. Más de veinte años después, en 1992, el Departamento de Defensa de los Estados Unidos en conjunto con el Instituto Nacional de Standards y Tecnología (NIST), patrocinaron la Conferencia de Recuperación de Texto (TREC) como parte del programa TIPSTER; esta conferencia proporcionó ayuda al facilitar la infraestructura necesaria para evaluar las metodologías de recuperación de texto. (9)

Entre los tipos de información que se pueden recuperar en la actualidad se encuentran: archivos de audio, imágenes, archivos de video y archivos de texto. (7) Para llevar a cabo este proceso de RI se realizan varias actividades como: definir las necesidades de búsqueda, seleccionar las herramientas de búsqueda, definir una estrategia de búsqueda (técnicas y modelos de RI) y evaluar los resultados para ver si se obtuvo información relevante o de lo contrario redefinir la estrategia de búsqueda. (7)

El proceso de RI se puede ver representado en el siguiente gráfico:

⁵ Estudio científico tanto de la estructura de las lenguas naturales y de aspectos relacionados con ellas como su evolución histórica.

⁶ Disciplina que aborda la interpretación y producción del sentido. Esto significa que estudia fenómenos significantes, objetos de sentidos, sistemas de significación, lenguajes, discursos y los procesos a ellos asociados: la producción e interpretación.

⁷ Ciencia que estudia métodos, procesos, técnicas, con el fin de almacenar, procesar y transmitir información y datos en formato digital.

⁸ Conjunto de actividades técnicas y científicas que tiene como finalidad el conocimiento de la información, de los materiales en que se presenta y del lugar en que se hace uso de ella.

⁹ Estudio teórico y práctico de los principios, procedimientos y problemas concernientes al almacenamiento de documentos.

¹⁰ Ciencia del procesamiento de la información, que proporciona información sobre algo con un fin determinado.

¹¹ (1890-1974) fue un ingeniero y científico estadounidense.

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

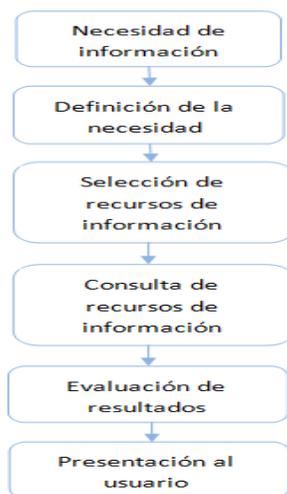


Figura 1: Proceso de RI. (7)

1.2.1 Técnicas y modelos de recuperación de la información

Para implementar el proceso de RI se hace necesario tener conocimiento acerca de las técnicas y modelos de recuperación de la información que existen para identificar los más adecuados en dependencia de las necesidades propias.

En nuestros días con el incremento de la información digital se hace necesario contar con herramientas informáticas adecuadas para su recuperación. Actualmente los ordenadores hacen posible representar un documento utilizando su contenido completo, denominado representación a texto completo, la cual es la forma más completa que existe, pero implica un coste computacional muy alto para colecciones grandes de documentos. Por lo anterior se busca encontrar el equilibrio entre el número de términos utilizados para representar su contenido y el coste computacional asociado para obtener la salida en un tiempo razonable. (10)

Técnicas de recuperación de la información

Las técnicas de recuperación de la información son las prácticas a seguir para recuperar la información de los documentos según criterios de búsqueda, que utilizan algoritmos y modelos de RI, para obtener resultados relevantes con respecto a lo que se desea encontrar. Entre las técnicas más usadas se encuentran:

Sistemas de recuperación de lógica difusa: Esta técnica permite establecer consultas con frases normales, de forma que la máquina al realizar la búsqueda elimina signos de puntuación, artículos, conjunciones, plurales, tiempos verbales, dejando solo aquellas palabras que el sistema considera relevantes. (6)

Técnicas de ponderación de términos: La ponderación pretende darle un valor adecuado a la búsqueda

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

dependiendo de los intereses del usuario. Los documentos recuperados se encuentran en función del valor obtenido en la ponderación. El valor depende de los términos convenientes que el documento contenga y la frecuencia con que se repitan. (6)

Técnica de *clustering*: Es un modelo probabilístico que permite las frecuencias de los términos de búsqueda en los documentos recuperados. Se atribuyen unos valores o pesos que actúan como agentes para agrupar los documentos por orden de importancia, mediante algoritmos *ranking* (proveen ciertos parámetros para ordenar objetos). (6)

Técnicas de retroalimentación por relevancia: Pretenden obtener el mayor número de documentos relevantes tras establecer varias estrategias de búsqueda. La idea es que, tras determinar unos criterios de búsqueda y observar los documentos recuperados, se vuelva a repetir nuevamente la consulta pero esta vez con los elementos interesantes, seleccionados de los documentos primeramente recuperados. (6)

Técnicas de *stemming*: Pretenden eliminar las posibles confusiones semánticas que puedan ocurrir en la búsqueda de un concepto, para ello truncan la palabra y buscan solo por la raíz. (6)

La comprensión de las diferentes técnicas permite conocer cómo la herramienta o biblioteca que se seleccione para integrar con el proyecto realiza la búsqueda y recuperación de la información, y de esta forma entender mejor los procesos internos que son realizados en esta. Por este motivo es necesario poseer un conocimiento básico de cada técnica de recuperación.

Modelos de recuperación de la información

Los modelos de recuperación son las reglas por las que se determina la relevancia de los documentos encontrados según las técnicas de RI. Estos se diferencian en el método elegido para representar los documentos y las consultas, y para realizar las búsquedas de información. (10)

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

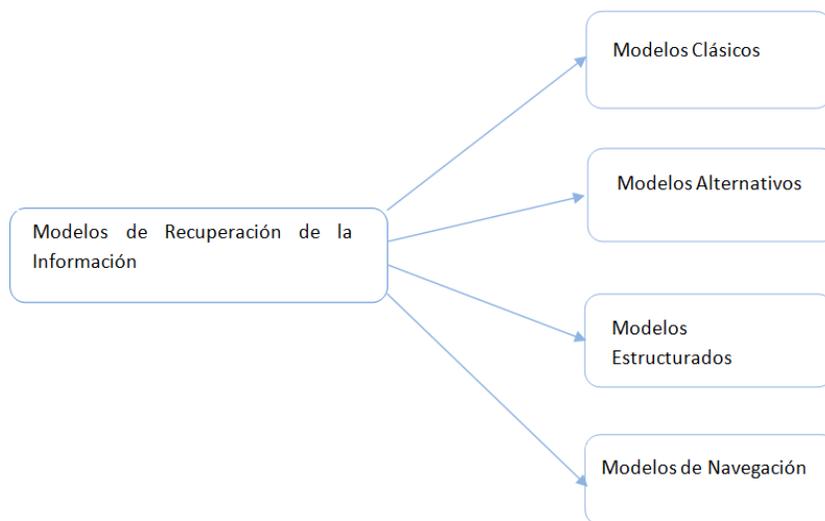


Figura 2: Modelos de Recuperación de la Información.

Modelos Clásicos:

- Consideran que cada documento está descrito por un conjunto de términos indexados, en la cual la relevancia de un término de cada documento está dado por un peso.

Modelos Alternativos:

- Se extiende de los modelos clásicos.
- Son generalmente bastante costosos de implementar y no siempre dan grandes resultados.

Modelos Estructurados:

- Tratan de combinar la información del contenido del texto con la estructura del mismo.
- Se pierde la noción de relevancia, y se está en presencia de una recuperación de datos.

Modelos de Navegación:

- No se basan en consultas del usuario, sino que el usuario navega a través de una jerarquía hasta encontrar los documentos relacionados a lo que anda buscando.
- Son una guía jerárquica de directorios que va de los temas más generales a los más particulares.
- Listan lugares (URLs) y los clasifican en categorías, además de añadir comentarios identificativos sobre ellos. (11)

El conocimiento de las normas para determinar la relevancia de los documentos recuperados es importante para determinar cómo serán devueltos los resultados de la recuperación de información. Según las características expuestas de los modelos, los que más se adaptan a las necesidades del proyecto y presentan

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

menos desventajas en cuanto a su uso son los modelos clásicos. Dentro de estos existen varios tipos como el booleano, el espacio vectorial y el probabilístico.

1.2.2 Bibliotecas y herramientas de búsqueda

Bibliotecas

Las bibliotecas informáticas son un conjunto de funciones relacionadas entre sí, estas contienen código y datos, que proporcionan servicios a programas y pasan a formar parte de estos. Existen diferentes bibliotecas para la recuperación de la información, algunas de ellas son la base para herramientas de búsqueda como los motores de búsqueda. Entre las bibliotecas para la RI se encuentran:

Lemur: Como SRI permite todas las etapas desde la indización a la búsqueda de documentos. Aporta una poderosa API implementada en C++ y está diseñada para trabajar en todos los sistemas operativos, permite la indización incremental e indiza atributos de los documentos. Basada en ella se creó el motor de búsqueda Indri. (12)

Xapian: Biblioteca de funciones *OpenSource* de RI para crear motores de búsqueda. Está escrita en C++, pero también se encuentra disponible en otros lenguajes como Perl, Python, PHP, Java, C# y Ruby. Contiene una API potente y adaptable que le facilita al programador los procesos de indización y búsqueda. (12)

Lucene: Biblioteca escalable de alto rendimiento para la recuperación de información, que permite agregar capacidades de indexación y búsqueda a las aplicaciones. Se trata de una API de indexación y búsqueda con simplicidad en su utilización que requiere básicamente que el usuario aprenda el uso de sus clases. Escrita originalmente en Java ha sido traducida a Perl, Python, C++, C#, .Net y Ruby. Basada en ella se creó el motor de búsqueda Apache Solr. (10)

Características	Lemur	Xapian	Lucene
Multiplataforma.	Sí	Sí	Sí
Técnica de <i>stemming</i> .	Sí	Sí	Sí
Lenguajes de programación.	C++ (soporte para otros lenguajes de programación)	C++ (soporte para otros lenguajes de programación)	Java (soporte para otros lenguajes de programación)
Código abierto.	Sí	Sí	Sí
Buscar mientras se actualiza el índice.	No	No	Sí
Indización incremental.	Sí	No	Sí
Modelo de RI.	Probabilístico	Probabilístico	Espacio vectorial

Tabla 1: Comparación entre bibliotecas. (12)

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

Por las características antes expuestas se escoge como biblioteca Lucene, que es de alto rendimiento a la vez que es simple en su utilización. Permite además buscar mientras se actualiza el índice, la indización incremental y agregar capacidades de indexación y búsqueda a las aplicaciones. Lucene utiliza la técnica de *stemming* para eliminar confusiones semánticas al buscar por la raíz de la palabra. Igualmente hace uso del modelo de recuperación clásico: espacio vectorial, el cual en lugar de predecir si un documento es pertinente o no, da un *ranking* de los documentos de acuerdo con su grado de similitud a la consulta.

Herramientas

Las herramientas encargadas de la RI en Internet pueden ser fundamentalmente clasificadas en dos grupos: directorios *web* o índices temáticos y los motores de búsqueda o buscadores *web*. Además existen otras, no solo para Internet, como son: bases de datos, lenguajes de indización y control terminológico (índices, palabras clave, *tesauros*), lenguajes de interrogación y ecuaciones de búsqueda (ecuaciones simples o compuestas) y operadores (lógicos, posicionales, de existencia, de comparación y de truncamiento). En este trabajo se profundizará solamente en los dos grupos más usados de herramientas de RI para Internet, o sea, directorios y buscadores *web*.

Se puede definir a un directorio *web* como un sistema de búsqueda por temas o categorías con jerarquía (aunque también suelen incluir sistemas de búsqueda por palabras clave). Se trata de bases de datos de direcciones *web* elaboradas manualmente, es decir, hay personas que se encargan de asignar cada página *web* a una categoría o tema determinado. (13)

Por otro lado los motores de búsqueda son sistemas de búsqueda por palabras claves, utilizan programas automáticos, generalmente llamados *crawlers* o *spiders* (arañas), que rastrean la *web* recolectando páginas o parte de ellas, y las almacenan en bases de datos documentales, proporcionando una herramienta para hacer búsquedas sobre estas bases de datos. También pueden ser usados para un sitio en específico, en el cual el buscador localizaría los contenidos existentes en el sitio. (13)

Motores de búsqueda	Directorios <i>web</i>
La información se actualiza automáticamente por la red.	La información se actualiza mediante la mano humana que se da de alta en el directorio cuando crea un lugar <i>web</i> .
Recogen toda la información almacenada en la página.	No almacenan todos los contenidos <i>web</i> , solamente los campos más relevantes como son el título y las palabras clave.
Almacenan la información mediante una base de datos propia.	Almacenan información mediante directorios, clasificados en categorías.
La búsqueda se realiza en la base de datos mediante la ecuación de búsqueda.	La búsqueda se realiza jerárquicamente según las categorías establecidas.
La presentación de los resultados se establece por orden de relevancia según unos criterios	La presentación de los resultados se lleva a cabo mediante un listado de todos los documentos correspondientes en la

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

establecidos en la ecuación de búsqueda.	categoría, sin ningún criterio de presentación.
Apropiados para localizar información específica.	Apropiados para localizar información general sobre un tema.

Tabla 2: Comparación entre motores de búsqueda y directorios. (6)

Por estas características la mejor opción como herramienta de búsqueda para el sitio sería un motor de búsqueda. Entre los más populares se encuentran:

Apache Solr: es una plataforma de búsquedas basada en Apache Lucene, que funciona como un servidor de búsquedas. Sus principales características incluyen búsquedas de texto completo, resaltado de resultados, *clustering* dinámico, y manejo de documentos ricos (como *Word* y PDF). Es escalable, permitiendo realizar búsquedas distribuidas y replicación de índices, corre sobre un contenedor de *servlets* Java como Apache Tomcat y actualmente se está usando en muchos de los sitios más grandes de Internet (14). Además permite la búsqueda facetada, o filtrada según ciertos parámetros de los contenidos encontrados.

Minion: está diseñado para ser altamente configurable y es recomendable su uso en investigaciones, así como en entornos empresariales. En mayo 2008, se publicó su código fuente, y la versión desarrollada hasta el momento. Es utilizado por el sistema de recomendación Aura, desarrollado en los laboratorios de Sun, así como en aplicaciones *web* de la compañía. (15)

Terrier: se comenzó a desarrollar en el 2001 y ha estado disponible al público desde noviembre de 2004. Su nombre es un acrónimo de *TERabyte RetrIEvER*. Permite realizar experimentación con fuentes documentales de gran escala. Desde la publicación de su API, existen más de 40 instituciones académicas y empresariales que manifiestan interés en su utilización. (15)

Indri: es un motor de búsqueda basado en el proyecto Lemur; consiste en un esfuerzo cooperativo entre dos universidades. El sistema se concluyó en el 2004. Desde la perspectiva empresarial, es eficiente y fácil de integrar en un SRI. (15)

En la Tabla 3 se hará una comparación entre los motores anteriormente mencionados de acuerdo con ciertos indicadores. Estos indicadores se especifican seguidamente para una mejor comprensión.

Estructura del índice: indica la forma en que el indexador almacena el índice.

Modelo RI: modelo que implementa el motor de búsqueda correspondiente.

Creación de resumen: indica si tiene la capacidad de hacer un resumen del contenido de un documento recuperado, generalmente, se resaltan los términos que coinciden con las consultas.

Tipo de operadores de búsqueda: tipo de búsquedas que es capaz de hacer, y si acepta operadores de consulta.

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

Lenguaje de programación: Lenguaje de programación usado para implementar el indexador y las herramientas de búsqueda.

Lematización: expresa si el indexador es capaz de hacer operaciones de lematización (términos con el mismo origen) sobre las palabras.

Índice sensible a mayúsculas y minúsculas: El índice tiene en cuenta las mayúsculas y minúsculas del texto en cada documento.

Licencia: Determina las condiciones para usar y modificar el motor de búsqueda.

Características	Apache Solr/ Apache Lucene	Minion	Terrier	Indri/Lemur
Estructura del índice.	Índice invertido.	Índice invertido.	Índice invertido.	Índice invertido.
Modelo RI.	Combinación del modelo espacio vectorial y el booleano.	Combinación del modelo booleano, consultas relacionales y de proximidad.	Divergencia de aleatoriedad, basado en el modelo probabilístico.	Combinación del modelo de redes de inferencia y modelado de lenguaje natural.
Creación de resumen.	Sí	Sí	No	Sí
Tipo de operadores de búsqueda.	Booleanos, frases, términos <i>wildcard</i> ¹² , y campos.	Booleanos, proximidad, pasajes, y campos.	Booleanos, operadores +/-, frases, proximidad, y campos.	Booleanos, frases, sinónimos, términos <i>wildcard</i> , pasajes, y campos.
Lenguaje de programación.	Java, C++, C#, Python, Perl, Ruby, y PHP	Java	Java	Java, C++, C#, y PHP
Lematización (Técnica de <i>stemming</i>).	Sí	Sí	Sí	Sí
Índice sensible a mayúsculas y minúsculas.	Sí	Sí	Solo para fuentes documentales TREC.	Sí
Licencia	Apache Software Foundation (ASF).	GNU Public License (GPL).	Mozilla Public License (MPL).	Berkeley Software Development (BSD).

Tabla 3: Comparación entre motores de búsqueda. (15)

Por sus características, la popularidad en su uso, la búsqueda facetada necesaria para cumplir los requisitos solicitados por el cliente, además de que se puede integrar a la arquitectura del proyecto (Drupal) y el uso de la biblioteca Lucene; se selecciona el motor de búsqueda Apache Solr como herramienta de RI. Solr utiliza la

¹² Carácter comodín, en informática, un carácter tal como el asterisco (*) que representa cualquier otro carácter o cadena de caracteres.

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

técnica de *stemming* ya que está basado en Lucene que hace uso de ella. De la misma forma utiliza una combinación de los modelos clásicos: espacio vectorial (usado por Lucene) y booleano (modelo basado en la teoría de conjuntos y el álgebra booleana).

1.2.3 Lenguajes y tipos de búsqueda

Lenguajes de recuperación de la información

Una manera para acceder a la información digital en las distintas áreas del conocimiento es mediante la RI sobre documentos, para ello existe una variedad de lenguajes de consulta y recuperación de información.

Entre los más significativos se encuentran:

XQUERY: es un lenguaje de consultas estándar que utiliza la notación XML para definir consultas y manejar los resultados, se centra en encontrar y extraer elementos y atributos de documentos XML. Está definido en términos de un modelo formal abstracto, no en términos de texto XML. Cada entrada a una consulta es una instancia de un modelo de datos, y la salida de una consulta también.

XQL: es un lenguaje de consultas que se basa en operadores de búsqueda de un modelo de datos para documentos XML que puede realizar consultas en infinidad de tipos de documentos como son: documentos estructurados, colecciones de documentos, bases de datos, estructuras DOM y catálogos. Es una notación para obtener información de un documento, esta puede ser un conjunto de nodos, información sobre las relaciones entre nodos, o valores derivados.

SQL: *Structured Query Language* o lenguaje de consulta estructurado. Este es el lenguaje que se utiliza para conectarse a una base de datos. Son sentencias, que realizan un query (consulta) a la base de datos a fin de que esta le responda con una cantidad de datos limitada según lo buscado. (2)

Puesto que la herramienta seleccionada es un motor de búsqueda, los cuales guardan la información dentro de una base de datos, es necesaria una forma de realizar consultas sobre esta para obtener resultados relevantes. Por esta razón el lenguaje de consulta que más responde a las necesidades del proyecto es SQL, ya que está diseñado para realizar consultas a sobre una base de datos.

Tipos de búsqueda

Existe una gran diversidad de búsquedas utilizadas por los buscadores *web*, cada una con sus características particulares. Seleccionar y estudiar el tipo de búsqueda que realizará un buscador es un aspecto esencial para el resultado del mismo.

Full Text Search: “La búsqueda de texto completo proporciona la capacidad para identificar los documentos en lenguaje natural que satisfacen una consulta y, opcionalmente, para ordenarlos por relevancia a la consulta”.

(16)Esta provee indexación de texto, permitiéndole a los documentos un preprocesado y un índice revisado

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

posteriormente a cualquier búsqueda rápida. (17) Es empleado por algunos motores de búsqueda, tales como AltaVista¹³.

Aunque es el tipo de búsqueda más utilizado en la actualidad presenta como desventaja principal que el usuario no siempre conoce lo que está buscando por lo que es necesario que el usuario introduzca términos que sean lo más semejante posible al resultado deseado. Existe otro tipo de búsqueda que da respuesta a este problema: *Fuzzy Search*.

Fuzzy Search: “En informática, una cadena coincidente aproximada es la técnica de búsqueda de coincidencias aproximadas a un patrón en una cadena”. (18) Las búsquedas *fuzzy* o difusa tienen objetivo que, cuando la persona no conoce cómo se escribe con exactitud lo que quiere buscar, el sistema retorna términos que en su deletreo sean parecidos a la entrada de la búsqueda. Es usada por algunas herramientas como Lucene¹⁴.

Por escogerse la biblioteca Lucene se usará como tipo de búsqueda las búsquedas difusas (*Fuzzy Search*) que al mismo tiempo resuelven la principal desventaja de la búsqueda de texto completo; permitiéndole al usuario, que no conoce con precisión cómo se escribe lo que desea buscar, encontrar más resultados que con la *Full Text Search*.

1.3 Portales educativos

Los portales educativos por sus características como distribuidores de información y almacenes de recursos, constituyen para la escuela actual un soporte ideal y eficiente de información relacionado con el mundo educativo. (3)

El origen de los portales educativos es una especificación de los portales de carácter generalista. El porqué de su aparición podría entenderse a partir de la desalentadora e incesante búsqueda de información específica en Internet y la constante necesidad del usuario de encontrar mayor precisión y diversidad sobre un tema concreto. (3)

Entre los portales educativos más significativos a nivel mundial se pueden encontrar los siguientes:

- SchoolNet (Canadá)¹⁵
- Education.fr (Francia)¹⁶
- Educ.ar (Argentina)¹⁷

¹³ Disponible en: <http://es.altavista.com/>

¹⁴ Disponible en: <http://lucene.apache.org/java/docs/index.html>

¹⁵ Disponible en: <http://www.schoolnet.ca>

¹⁶ Disponible en: <http://www.education.fr>

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

Ventajas de los portales educativos

- Proporcionar información de todo tipo a profesores, estudiantes y padres, así como instrumentos para realizar búsquedas de información.
- Los estudiantes pueden obtener asistencia personalizada, ya que a diferencia de la educación tradicional, en un portal pueden o tienen herramientas para la comunicación asincrónica o sincrónica por los profesores.
- Proporcionar recursos didácticos de todo tipo, gratuitos y utilizables directamente desde el lugar donde se encuentre publicado el portal (materiales didácticos en línea).
- Proporcionar instrumentos para la comunicación: correo electrónico, *chat*, espacios para alojar páginas *web*. (2)

1.3.1 Tendencias actuales. Comunidades virtuales

Una nueva tendencia en los portales educativos son las comunidades virtuales. Estas son sitios creados por una o más personas que establecen relaciones a partir de temas comunes. Los miembros de estas comunidades crean páginas, publican eventos, administran sus foros de discusión, salas de *chat*, álbumes de fotos, y archivos para compartir. (19)

El término Comunidad Virtual (CV) fue acuñado por Howard Rheingold¹⁸ en 1996 definiéndola como *"agregaciones sociales que emergen de la red cuando un número suficiente de personas entablan discusiones públicas durante un tiempo lo suficientemente largo, con un suficiente sentimiento humano, para formar redes de relaciones personales en el ciberespacio"*. (8) Otro concepto más actual es el dado en el 2003 por Kardaras¹⁹, Karakostas²⁰ y Papathanassiou²¹ *"grupo de individuos que se comunica a través de medios electrónicos, como Internet, que comparten intereses, sin necesidad de localizarse en un mismo lugar, contactar físicamente o pertenecer a una determinada etnia."* (8)

En dependencia del tipo de construcción de conocimiento que se desarrolle en una CV, existen al menos cuatro tipos de comunidades propuestas por Jonassen²², Peck²³ y Wilson²⁴ en 1999:

¹⁷ Disponible en: <http://www.educ.ar>

¹⁸ Crítico y ensayista estadounidense. Está especializado en las implicaciones culturales, sociales y políticas de las nuevas tecnologías de la información, como Internet y la telefonía móvil.

¹⁹ Dr. Dimitris Kardaras. Profesor Asistente en Sistemas de Gestión de la Información. Universidad de Atenas de Economía y Negocios.

²⁰ Bill Karakostas. Profesor Titular. Posee un doctorado en Ingeniería de *Software* de la Universidad de Manchester.

²¹ Dr. Eleutherios Papathanassiou. Universidad de Atenas de Economía y Negocios, Informática de Negocios, Departamento de Administración de Empresas. Ph.D. en Ciencias de la Computación.

²² Profesor de Sistemas de Instrucción de la Universidad de Pennsylvania.

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

CV de discurso: Se desarrolla a partir de la dimensión de Ser Social-Dialógico del hombre y la necesidad de relacionarse con otros, en torno a temas de interés común. Esto es facilitado por la integración de los distintos medios de comunicación masiva que proporcionan numerosas herramientas para el desarrollo de este tipo de comunidades.

CV de práctica: Surgen cuando se necesita aprender sobre algún tema nuevo, por lo que su característica es estar enfocadas al aprendizaje de nuevos conocimientos y habilidades sobre la base de lo que pueden transferir los más experimentados y por el contacto con el contexto natural en el que estos aprendizajes se utilizan.

CV de construcción de conocimiento: Tienen como objetivo que los estudiantes desarrollen de manera activa y estratégica el aprendizaje, reconociendo lo que el estudiante sabe e induciéndolo a la construcción de su propio conocimiento.

CV de aprendizaje: Surgen cuando sus integrantes comparten intereses, conocimientos, valores y metas. A partir de esto se desarrollan diferentes tareas para lograr el aprendizaje, construir conocimientos, reflexionar sobre los conocimientos construidos e incluso reflexionar cómo ha sido ese proceso. (8)

1.3.2 Tecnologías para desarrollar portales educativos. CMS

El impresionante desarrollo y constante perfeccionamiento de los portales *web*, han traído consigo el importante reto de construirlos de acuerdo con ciertos estándares de calidad, lo que les impone a los desarrolladores *web* una tarea inaplazable de administración de su portal, eficiente, rápida y de calidad. Para cumplir con estos propósitos, se cuenta con herramientas de *software* de acceso público llamadas Sistema de Gestión de Contenidos (*Content Management Systems*). (2)

Los Sistemas de Gestión de Contenidos (CMS) incorporan potentes herramientas para la administración y gestión de contenidos. Suelen ofrecer procesos automatizados de recolección de información (noticias, vínculos, comentarios) y dar soporte a la creación de galerías de elementos multimedia. (2)

Entre los CMS más usados actualmente destaca Drupal que es un CMS modular y muy configurable. Es desarrollado y mantenido por una activa comunidad de usuarios. La utilización de este CMS forma parte de los requisitos del proyecto Portal Educativo por lo que es imperativo su uso para el desarrollo del módulo de búsqueda; ya que el portal desarrollado por el proyecto utiliza Drupal, por consiguiente, también sus módulos.

Algunas características de Drupal son:

²³Dr. Kyle L. Peck. Universidad de Pennsylvania. Ph.D. en Psicología de la Educación / Formación.

²⁴Doctor Bruce G. Wilson. Universidad de Pennsylvania.

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

- Programa de código abierto, con licencia GNU/GPL.
- Escrito en PHP.
- Destaca por la calidad de su código y de las páginas generadas.
- Su flexibilidad y adaptabilidad lo hacen adecuado para realizar diferentes tipos de sitios *web*.
- Genera un código HTML extremadamente limpio y bien estructurado.
- Existencia de gran cantidad de módulos contribuidos.
- El sistema base de gestión de contenido es limpio, flexible y potente. (20)

1.4 Metodología de desarrollo

Las Metodologías de Desarrollo de *Software* surgen por la necesidad de utilizar una serie de procedimientos, técnicas, herramientas y soporte documental para desarrollar un producto de *software*. Estas guían a los desarrolladores al crear un nuevo *software* y se dividen en dos grandes grupos: pesadas y ágiles. (21) Entre las metodologías pesadas se encuentra RUP que es la propuesta por el proyecto al que pertenece el módulo a desarrollar.

1.4.1 RUP

Proceso Unificado de Desarrollo (RUP: *Rational Unified Process*): es una Metodología de Desarrollo de *Software* pesada que está basada en componentes e interfaces bien definidas, y junto con el Lenguaje Unificado de Modelado (UML), constituye la metodología estándar más utilizada para el análisis, implementación y documentación de sistemas orientados a objetos.

Principales ventajas:

- Coste del riesgo a un solo incremento.
- Reduce el riesgo de no sacar el producto en el calendario previsto.
- Acelera el ritmo de desarrollo.
- Se adapta mejor a las necesidades del cliente. (22)

Se ha escogido esta metodología pesada porque responde a las necesidades planteadas y es útil cuando el cliente se encuentra lejos del equipo de trabajo y el proyecto a desarrollar es grande. Además debido a que es la metodología escogida por los lineamientos de la directiva del centro de desarrollo en el cual se enmarca el proyecto donde se desarrolla el Portal Educativo.

1.5 Tecnologías y Herramientas de Desarrollo

1.5.1 Tecnologías

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

Las tecnologías actuales evolucionan cada día más, por lo que es necesario mantener una constante actualización sobre el desarrollo de las mismas. Para una correcta implementación del módulo propuesto se hace necesario contar con tecnologías que permitan llegar a una satisfactoria solución del problema a resolver.

Lenguajes de programación

Como para la implementación del módulo se hará uso del CMS Drupal, los lenguajes de programación a utilizar deben ser los que emplee Drupal para la implementación de sus módulos, estos son: HTML, JavaScript, PHP y CSS.

HTML (*Hypertext Markup Language*): Lenguaje de programación que utiliza marcas para describir la forma en la que deberían aparecer el texto y los gráficos en un navegador *web* que, a su vez, están preparados para leer esas marcas y mostrar la información en un formato estándar. La versión HTML5 incluye novedades significativas en diversos ámbitos; ya no solo trata de incorporar nuevas etiquetas o eliminar otras, sino que supone mejoras en áreas que hasta ahora quedaban fuera del lenguaje y para las que se necesitaba utilizar otras tecnologías. (23)

JavaScript: Es un lenguaje interpretado por lo que no es necesario compilar los programas para ejecutarlos, se pueden probar en cualquier navegador sin necesidad de procesos intermedios y es orientado a objetos. Entre sus ventajas se encuentra:

- No requiere tiempo de compilación.
- Los scripts pueden desarrollarse en un período de tiempo relativamente corto.
- Posee características de interfaz, que son gestionadas por el navegador y por el código HTML. (24)

PHP: Este lenguaje es desarrollado por el Grupo PHP. (25) Entre sus características se encuentran:

- Es de código abierto.
- Especialmente adecuado para desarrollo *web*.
- Puede ser incrustado en HTML.
- Posee soporte para una gran cantidad de bases de datos.
- Es un lenguaje multiplataforma.
- Es usado por el CMS Drupal para la implementación de sus módulos. (25)

CSS: Lenguaje de hojas de estilos creado para controlar la presentación de los documentos electrónicos definidos con HTML y XHTML. CSS es la mejor forma de separar los contenidos y su presentación, y es imprescindible para la creación de páginas *web* complejas. Este lenguaje se utiliza para definir el aspecto de todos los contenidos, es decir, el color, tamaño y tipo de letra de los párrafos de texto, la separación entre titulares y párrafos y la tabulación con la que se muestran los elementos de una lista. (26)

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

Lenguaje de modelado. UML

UML es el lenguaje de modelado de sistemas de *software* más conocido y utilizado en la actualidad; está respaldado por el OMG (*Object Management Group*). (27)

Sus principales ventajas son:

- Se puede usar para modelar distintos tipos de sistemas: sistemas de *software*, sistemas de *hardware* y organizaciones del mundo real.
- Es una consolidación de muchas de las notaciones y conceptos más usados. (28)

1.5.2 Herramientas de Desarrollo

Es necesario hacer una correcta selección de las herramientas a utilizar durante el proceso de desarrollo del *software* para lograr una buena calidad, la satisfacción de los requisitos del cliente y un mejor aprovechamiento del tiempo empleado.

Entorno Integrado de Desarrollo (IDE). Netbeans

El IDE NetBeans es una herramienta para programadores pensada para escribir, compilar, depurar y ejecutar programas.

Características de IDE NetBeans versión 7:

- Escrito en Java (puede servir para cualquier otro lenguaje de programación como PHP).
- Existen módulos para extender el IDE NetBeans.
- Permite el auto completamiento de código.
- Es un producto libre y gratuito sin restricciones de uso (código abierto). (29)

Se decide utilizar NetBeans porque es una herramienta de código abierto, posee soporte para el lenguaje PHP y el equipo de desarrollo cuenta con experiencia de trabajo con este IDE.

Servidor Web. Apache Tomcat

Un servidor *web* es un programa que sirve para atender y responder a las diferentes peticiones de los navegadores, proporcionando los recursos que se soliciten usando el protocolo HTTP o HTTPS. (24) Apache Tomcat es un servidor *web* HTTP multiplataforma, además de servidor de aplicaciones o contenedor de *servlets*. Entre sus ventajas se encuentran que es: modular, de código abierto, extensible y popular (facilidades para conseguir ayuda o soporte). (30) Se caracteriza por ser un servidor ligero, con una alta configurabilidad y una amplia explotación. La arquitectura que utiliza es cliente-servidor, el cliente hace la petición al servidor y este atiende dicha petición. (24).

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

Se selecciona ya que es fácil de configurar, fiable, puede funcionar como servidor *web* en sí mismo y funciona en todos los sistemas operativos que dispongan de la máquina virtual de Java. Además es el servidor *web* orientado a utilizar por el proyecto.

Sistema Gestor de Base de Datos. PostgreSQL

PostgreSQL es un sistema de gestión de base de datos relacional orientada a objetos, publicado bajo la licencia BSD. (16) Es multiplataforma, permite una fácil gestión de la base de datos que contenga el sistema. Sirve de soporte a los lenguajes más utilizados como PHP, Java, C++, Python, C y C#. El número de base de datos que puede contener es ilimitado. Es considerado como el gestor libre más avanzado del mundo. (31)

Se selecciona porque además de las características anteriores es soportado por Drupal y utilizado por el proyecto Portal Educativo.

Herramienta de Modelado. Visual Paradigm

Visual Paradigm es una herramienta CASE (Ingeniería de *Software* Asistida por Computación) que propicia un conjunto de ayudas para el desarrollo de programas informáticos, desde la planificación, pasando por el análisis y el diseño, hasta la generación del código fuente de los programas y la documentación.

Esta herramienta permite aumentar la calidad del *software*, a través de la mejora de la productividad en el desarrollo y mantenimiento del *software*. Aumenta el conocimiento informático de una empresa ayudando así a la búsqueda de soluciones para los requisitos. También permite la reutilización del *software*, portabilidad y estandarización de la documentación, además del uso de las distintas metodologías propias de la Ingeniería del *Software*. (32)

El equipo de desarrollo cuenta con experiencia en el trabajo con esta herramienta y es la que utiliza el proyecto, por lo que se seleccionó como herramienta de modelado.

1.6 Conclusiones parciales

En este capítulo se realizó un estudio de las herramientas y bibliotecas actuales utilizadas para la RI, esto permitió escoger la herramienta que será integrada a la arquitectura del proyecto para facilitar la recuperación de información en el sitio. Para utilizar la herramienta de búsqueda es necesario desarrollar un módulo que la integre con la arquitectura del portal. Para la implementación de la solución se deben usar tecnologías libres por lo que se realizó un estudio detallado de los lenguajes, metodologías y herramientas desarrollo y se definieron aquellas tecnologías que respondían a las necesidades y estaban en concordancia con la arquitectura del proyecto en el cual se enmarca el módulo a desarrollar.

CAPÍTULO 2: CARACTERÍSTICAS DEL SISTEMA

CAPÍTULO 2: CARACTERÍSTICAS DEL SISTEMA

2.1 Introducción

Para un adecuado desarrollo del sistema se deben conocer los requisitos, funcionales y no funcionales, que debe cumplir el mismo. Se deben definir cómo serán la estructura y componentes de la solución propuesta para el sistema; quiénes serán los actores y cómo se relacionarán con el sistema para cumplir con los requisitos de este. Además se debe modelar el negocio para tener una visión de los procesos que se realizan en el sistema.

2.2 Propuesta de sistema

Para dar solución al problema planteado, se decidió implementar un módulo para el CMS Drupal que permita al usuario final recuperar la información existente en el portal. El módulo Búsqueda permite recuperar los distintos contenidos que existen en el sitio, como: avisos, eventos, comunidades virtuales, recursos educativos y boletines. También hace uso de la práctica de búsqueda facetada, para filtrar los resultados. Este módulo permite la realización de búsqueda simple, donde se pueden filtrar los resultados; y búsqueda avanzada, donde se pueden buscar varias palabras claves intercaladas con los operadores AND y OR, restringir la búsqueda por autor, tipo de contenido y rango de fechas.

Se analizaron la mayoría de los módulos realizados por la Comunidad de Drupal en esta esfera y se seleccionaron como módulos contribuidos: el Apache Solr, que integra el buscador Apache Solr con la arquitectura de Drupal; el Facetapi, que permite la creación de facetas para la búsqueda; el CTools que es requerido por el Facetapi. Además para ampliar las funcionalidades de la búsqueda se hizo uso de los módulos: ApacheSolr_attachments, para realizar la búsqueda dentro de documentos adjuntados, y el ApacheSolr_autocomplete, que permite el auto completamiento de la búsqueda. Todos estos módulos fueron personalizados según los requerimientos del sistema.

Para el despliegue del sistema es necesaria la instalación del Apache Solr en un servidor al que tenga acceso el servidor donde esté montado el portal; y que a este, a su vez, tengan acceso los ordenadores cliente donde los usuarios podrán realizar las peticiones de búsqueda.

2.3 Modelo de dominio

Al desarrollar cualquier sistema se hace necesario dividirlo en secciones, estas se pueden representar mediante modelos que permitan abstraer sus características esenciales. Como paso previo al diseño del sistema es preciso realizar un Modelo de Negocio o un Modelo de Dominio para capturar y enunciar la visión del proceso.

CAPÍTULO 2: CARACTERÍSTICAS DEL SISTEMA

Para la construcción de este sistema no se realizó un modelo del negocio ya que no son claramente visibles los procesos del negocio, por tanto, se realizó un modelo de dominio el cual se muestra a continuación.

- Usuario: Persona encargada de realizar la búsqueda en el sistema, sea una búsqueda simple o avanzada, y usando o no filtros.
- Tipo de Contenido: Representa los diferentes tipos de contenido que existen en el portal y que pueden ser buscados y recuperados.

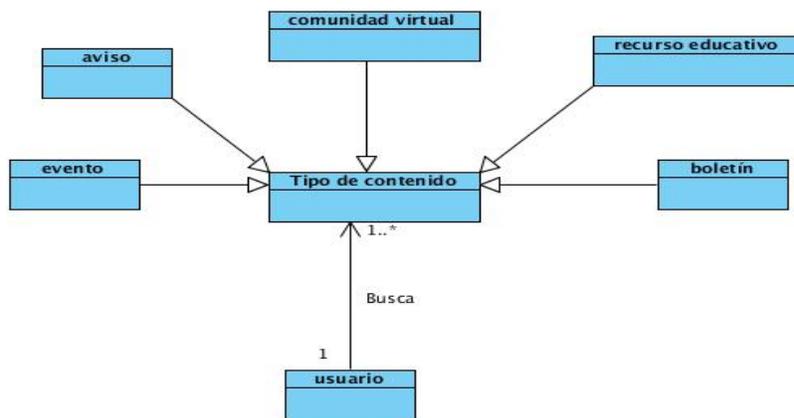


Figura 3: Diagrama del Modelo de Dominio.

2.4 Especificación de los requisitos de *software*

Los requerimientos para un sistema de *software* determinan lo que hará el sistema y definen las restricciones de su operación e implementación.

2.4.1 Requerimientos Funcionales

RF-1: Realizar búsqueda simple.

RF-2: Filtrar por comunidades virtuales.

RF-3: Filtrar por boletín.

RF-4: Filtrar por avisos.

RF-5: Filtrar por eventos.

RF-6: Filtrar por recursos educativos.

RF-7: Filtrar búsqueda por clasificación de eventos.

- Autor.
- Fecha de envío.
- Fecha de modificación.

RF-8: Filtrar búsqueda por clasificación de avisos.

CAPÍTULO 2: CARACTERÍSTICAS DEL SISTEMA

- Autor.
- Fecha de envío.
- Fecha de modificación.

RF-9: Filtrar búsqueda por clasificación de boletín.

- Año.
- Mes.
- Número.
- Autor.
- Fecha de envío.
- Fecha de modificación.

RF-10: Filtrar búsqueda por clasificación de comunidades virtuales.

- Autor.
- Fecha de envío.
- Fecha de modificación.
- Categoría.

RF-11: Filtrar búsqueda por clasificación de recurso educativo.

- Autor.
- Fecha de envío.
- Fecha de modificación.

RF-12: Realizar búsqueda avanzada.

2.4.2 Requerimientos No Funcionales

RNF-1: Usabilidad. Comprensión del usuario final.

El sistema debe garantizar una ayuda legible y comprensible, capaz de enseñar y guiar el uso del *software* al usuario final.

RNF-2: Accesibilidad.

Los elementos que conforman una interfaz de usuario deben tener un nombre sugerente, que logre que el usuario perciba la acción, además de mostrar una breve descripción de su función o razón de ser, en la propia interfaz.

En el caso de los formularios deben especificar las reglas de sus campos, como pueden ser: los obligatorios, aquellos que solo aceptan tipos de datos específicos o longitudes de cadenas limitadas.

CAPÍTULO 2: CARACTERÍSTICAS DEL SISTEMA

RNF-3: Confiabilidad.

Se deben garantizar el correcto acceso a cada funcionalidad, dependiendo de las credenciales de cada usuario.

RNF-4: Requerimientos de *Hardware*.

Los requisitos de *hardware* son:

- Microprocesador: 2 Core-Quad.
- Memoria de acceso aleatorio (RAM): 4GB.
- Capacidad de almacenamiento: 1TB.

RNF-5: Requerimientos de *Software*.

En los servidores se utilizan los siguientes *software*:

- Sistema operativo: GNU/Linux.
- Servidor de aplicaciones *web*: Apache 2.0.
- Sistema de gestión de contenido (CMS): Drupal 7.
- Gestor de base de datos: Postgres 9.0.
- Aplicaciones auxiliares: Apache Solr 3.4.

En el cliente, aunque solo se precisa de un navegador *web*, es necesario tener en consideración cuáles son adecuados y recomendables para hacer uso del sistema:

- Firefox (versión ≥ 4.0) (Recomendables).
- Internet Explorer (versión ≥ 9.0) (Recomendables).
- Google Chrome.
- Opera.

2.5 Definición de los casos de uso

Un caso de uso se define como “*una unidad coherente de funcionalidad, expresada como transacción entre los actores y el sistema.*” (33) Existen diferentes símbolos para la representación de los diagramas de casos de uso, para el diagrama de este sistema se usaron los símbolos siguientes:

CAPÍTULO 2: CARACTERÍSTICAS DEL SISTEMA

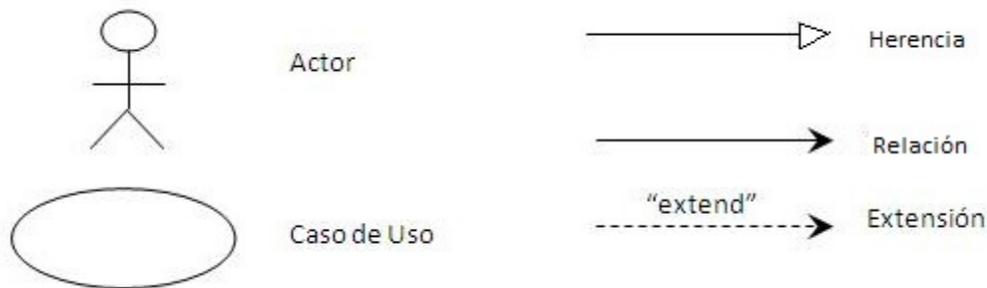


Figura 4: Simbología del Diagrama de Casos de Uso.

Patrones de Casos de Uso

Extensión: Consiste en dos casos de uso y una relación de extensión entre ellos. Este patrón se aplica cuando un flujo puede extender el flujo de otro caso de uso así como ser realizado en sí mismo. Ejemplo: el CU Filtrar por clasificación de eventos extiende del CU Filtrar por eventos.

Múltiples actores roles comunes: Ocurre cuando a un caso de uso ingresan más de dos actores y estos tienen un rol común. Ejemplo: el actor usuario registrado comparte un rol común con el actor usuario al acceder al CU Realizar búsqueda simple.

2.5.1 Definición de actores

Actor	Justificación
Usuario registrado	Usuario autenticado que interactúa con el <i>software</i> .
Usuario	Persona que todavía no se ha autenticado y que interactúa con el <i>software</i> .

Tabla 4: Definición de los actores del sistema.

Listado de casos de uso

Estos son el listado de cuatro de los casos de uso del sistema más significativos. Para visualizar el listado de los otros casos de uso remitirse al Anexo 1.

CU-1	Realizar búsqueda simple.
Actor	Usuario o usuario registrado.
Descripción	El caso de uso le permite al usuario realizar una búsqueda simple en el sistema.
Referencia	RF-1

Tabla 5: Definición del CU Realizar búsqueda simple.

CAPÍTULO 2: CARACTERÍSTICAS DEL SISTEMA

CU-2	Filtrar por comunidades virtuales.
Actor	Usuario.
Descripción	El caso de uso le permite al usuario filtrar la búsqueda por comunidades virtuales.
Referencia	RF-2, CU-1

Tabla 6: Definición del CU Filtrar por comunidades virtuales.

CU-10	Filtrar por clasificación de comunidades virtuales.
Actor	Usuario.
Descripción	El caso de uso le permite al usuario, después filtrar por comunidades virtuales, filtrar estas por autor, fecha de envío, fecha de modificación o categoría.
Referencia	RF-10, CU-2

Tabla 7: Definición del CU Filtrar por clasificación de comunidades virtuales.

CU-12	Realizar búsqueda avanzada.
Actor	Usuario o usuario registrado.
Descripción	El caso de uso le permite al usuario realizar una búsqueda avanzada en el sistema, restringiendo la búsqueda, por autor, tipo de contenido y rango de fecha; y utilizando los operadores lógicos AND y OR.
Referencia	RF-12

Tabla 8: Definición del CU Realizar búsqueda avanzada.

2.5.2 Diagrama de casos de uso

CAPÍTULO 2: CARACTERÍSTICAS DEL SISTEMA

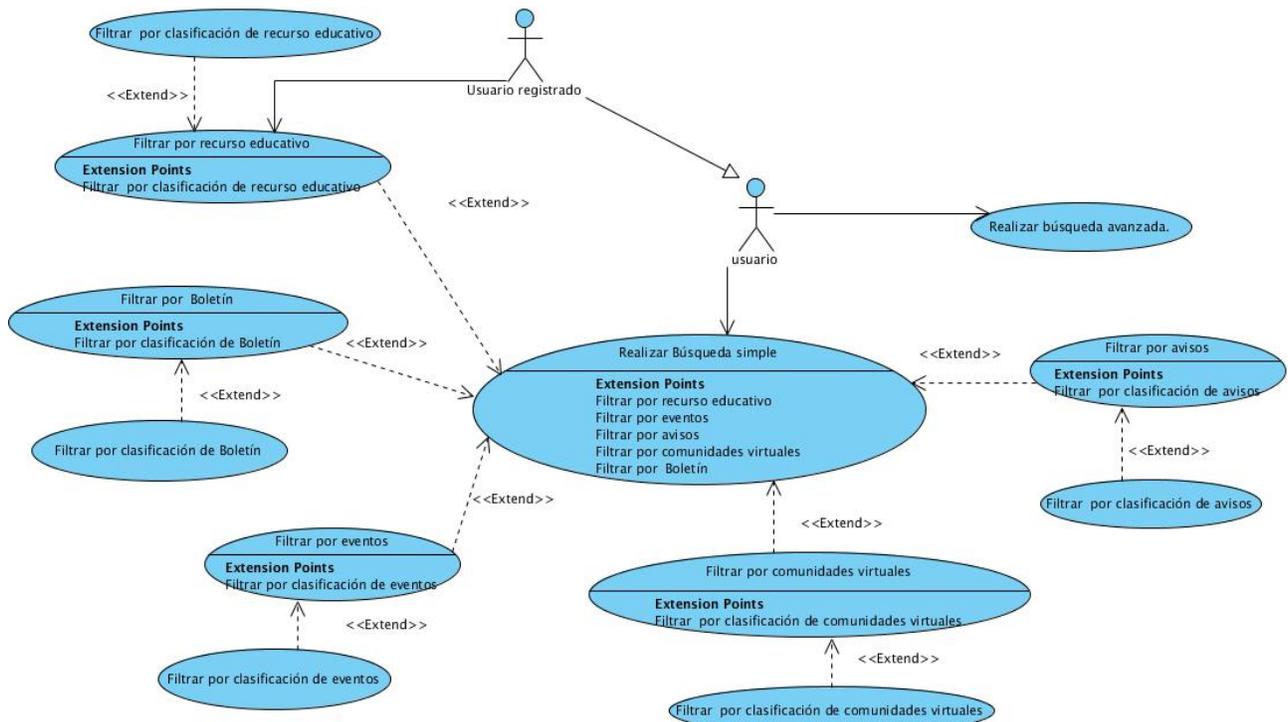


Figura 5: Diagrama de Casos de Uso del Sistema.

2.5.3 Descripción de casos de uso

Para especificar cada una de las funcionalidades que deben ser implementadas se describen los casos de uso del sistema. A continuación se muestra la descripción de cuatro de los casos de uso más significativos. Para visualizar las especificaciones de los otros casos de uso remitirse al Anexo 2.

Caso de uso	
CU-1	Realizar búsqueda simple.
Propósito	Realizar la búsqueda general en el sistema de los diferentes tipos de contenido.
Actores: Usuario o usuario registrado.	
Resumen: El usuario puede hacer una búsqueda general en el sistema, introduciendo una o varias palabras claves en un mismo campo. Los contenidos que se pueden localizar son los avisos, boletines y sus números, comunidades virtuales, los eventos y en el caso de los recursos educativos, solo si es usuario registrado y tiene acceso a él. Esta búsqueda se puede filtrar antes de realizarse y después.	
Referencias	RF-1
Acción del actor	Respuesta del sistema
1 El caso de uso se inicia cuando el actor introduce una o más palabras en el formulario de búsqueda	2 Se buscan las palabras en los contenidos a los que tiene acceso el usuario (los avisos, boletines y sus números,

CAPÍTULO 2: CARACTERÍSTICAS DEL SISTEMA

general.	comunidades virtuales y los eventos, en el caso del usuario; el usuario registrado puede acceder además a los recursos educativos a los que tenga ese privilegio).
4 Puede filtrar los resultados obtenidos por el tipo de contenido que desee (avisos, boletines y sus números, comunidades virtuales, recursos educativos y eventos).	3 Se devuelven los resultados encontrados. En el formato de título y fragmento de la descripción con las palabras clave destacadas, además del autor y la fecha de modificación.
	5 El CU termina.
Flujos Alternos	
Flujo alternativo al paso 2 “No tiene privilegios el usuario”	
Acción del actor	Respuesta del sistema
	Solo se busca en los contenidos a los que tenga acceso el usuario.
Flujo alternativo al paso 3 “No se encontró resultados”	
Acción del actor	Respuesta del sistema
	Se muestra un mensaje diciendo que no se encontraron resultados relacionados con la búsqueda solicitada y recomienda cómo realizar una búsqueda más eficiente.
Puntos de extensión	
Línea 4 el usuario puede filtrar los resultados obtenidos. Ver CU-2 al CU-6.	

Tabla 9: Descripción del CU Realizar búsqueda simple.

Caso de uso	
CU-2	Filtrar por comunidades virtuales.
Propósito	Filtrar por comunidades virtuales la búsqueda que se va a realizar.
Actores: Usuario.	
Resumen: El caso de uso se inicia cuando el usuario selecciona la opción que le permite filtrar la búsqueda que va a realizar por comunidades virtuales. El sistema restringe la búsqueda a las comunidades virtuales solamente.	
Referencias	RF-2, CU-1.
Acción del actor	Respuesta del sistema
1 El caso de uso se inicia cuando el actor selecciona la opción de filtrar por comunidades virtuales.	2 Se filtra la búsqueda para cuando se vaya a realizar solo busque en las comunidades virtuales.

CAPÍTULO 2: CARACTERÍSTICAS DEL SISTEMA

3 Puede ahora realizar una búsqueda filtrada solo en las comunidades virtuales. (Para mantener los filtros seleccionados debe dar clic en la casilla para mantener los filtros anteriores).	4 El CU termina.
Puntos de extensión	
Línea 3 el usuario puede restringir más la búsqueda. Ver CU-10.	

Tabla 10: Descripción del CU Filtrar por comunidades virtuales.

Caso de uso	
CU-10	Filtrar por clasificación de comunidades virtuales.
Propósito	Filtrar la búsqueda realizada en las comunidades virtuales, por ciertos campos de las comunidades virtuales.
Actores: Usuario.	
Resumen: El caso de uso se inicia cuando el usuario selecciona la opción que le permite restringir la búsqueda realizada en las comunidades virtuales por autor, fecha de envío, fecha de modificación o categoría. Además de poder escoger otro tipo de contenido que quiera filtrar. El sistema restringe la búsqueda según la opción escogida o busca en el otro tipo de contenido seleccionado.	
Referencias	RF-10, CU-2.
Acción del actor	Respuesta del sistema
1 El caso de uso se inicia cuando el actor después de haber realizado una búsqueda inicial filtrada por las comunidades virtuales desea restringir más la búsqueda y selecciona una de las opciones presentadas: <ul style="list-style-type: none"> ✓ Autor. ✓ Fecha de envío. ✓ Fecha de modificación. ✓ Categoría. O puede buscar por otro tipo de contenido.	2 Se restringe la búsqueda en las comunidades virtuales según la opción escogida. O busca en el tipo de contenido seleccionado.
3 Puede buscar en la nueva restricción escogida. (Para mantener los filtros seleccionados debe dar clic en la casilla para mantener los filtros anteriores.).	4 El CU termina.

Tabla 11: Descripción del CU Filtrar por clasificación de comunidades virtuales.

CAPÍTULO 2: CARACTERÍSTICAS DEL SISTEMA

Caso de uso	
CU-12	Realizar búsqueda avanzada.
Propósito	Realizar la búsqueda avanzada, en el sistema, de los diferentes tipos de contenido.
Actores: Usuario o usuario registrado.	
Resumen: El usuario puede hacer una búsqueda avanzada en el sistema, para ello puede especificar palabras claves en varios campos, como máximo 3, que a su vez pueden combinarse con los operadores lógicos AND y OR. Se puede hacer una selección múltiple de por cuales contenidos se desea buscar, además de un rango de fecha de envío del contenido y el autor del mismo.	
Referencias	RF-12
Acción del actor	Respuesta del sistema
1 El caso de uso se inicia cuando el actor introduce una o más palabras en los formularios de búsqueda avanzada, puede combinar las palabras con los operadores lógicos AND y OR. Igualmente puede hacer una selección múltiple de por cuales contenidos desea buscar, además de un rango de fecha de envío del contenido y el autor del mismo.	2 Se buscan las palabras en los contenidos a los que tiene acceso el usuario (los avisos, boletines y sus números, comunidades virtuales y los eventos, en el caso del usuario; el usuario registrado puede acceder además a los recursos educativos a los que tenga ese privilegio), siguiendo las especificaciones entradas para filtrar la búsqueda.
	3 Se devuelven los resultados encontrados; en el formato de título y fragmento de la descripción con las palabras clave destacadas, además del autor y la fecha de modificación.
	4 El CU termina.
Flujos Alternos	
Flujo alternativo al paso 2 “No tiene privilegios el usuario”	
Acción del actor	Respuesta del sistema
	Se busca en los contenidos a los que tenga acceso el usuario.
Flujo alternativo al paso 3 “No se encontró resultados”	
Acción del actor	Respuesta del sistema
	Se muestra un mensaje diciendo que no se encontraron resultados relacionados con la búsqueda solicitada y recomienda cómo realizar una búsqueda más eficiente.

Tabla 12: Descripción del CU Realizar búsqueda avanzada.

CAPÍTULO 2: CARACTERÍSTICAS DEL SISTEMA

2.6 Conclusiones parciales

Durante este capítulo se expuso la descripción de la propuesta del sistema, se definieron los 12 requisitos funcionales y los 5 no funcionales. Además se mostró el modelo de dominio, el diagrama de casos de uso del sistema, la descripción de los actores y una descripción de cuatro de los casos de uso más significativos. Todo lo anterior permite tener una clara visión del sistema, los procesos que suceden dentro de este y quiénes participan en cada acción realizada. Este conocimiento sobre las características del sistema posibilita desarrollar el módulo convenientemente.

CAPÍTULO 3: ANÁLISIS Y DISEÑO DEL SISTEMA

CAPÍTULO 3: ANÁLISIS Y DISEÑO DEL SISTEMA

3.1 Introducción

Para tener una clara idea de cómo debe funcionar el módulo de búsqueda se debe realizar un correcto análisis y diseño del sistema, porque de ello dependerá que la implementación de dicho módulo esté de acuerdo con lo especificado en los requisitos. Para realizar el análisis y diseño se deben aplicar ciertas técnicas y principios (guiados por la metodología de *software* utilizada) con el propósito de definir el sistema, con suficientes detalles como para permitir su interpretación y realización física (su implementación).

En el presente capítulo se describen los procesos que se llevan a cabo durante el diseño del sistema basándose en los casos de uso, donde se representarán los diagramas de clases del diseño, los diagramas de colaboración y el diagrama entidad-relación de la base de datos. También se muestran la arquitectura del sistema y los patrones de diseño y arquitectónico.

3.2 Arquitectura

Drupal es un CMS cuya lógica está programada en PHP, siguiendo un modelo de programación estructurada, y que hace uso de un sistema de bases de datos relacional (por ejemplo, MySQL o PostgreSQL).

El código que constituye el núcleo de Drupal está formado por un conjunto de bibliotecas que permiten gestionar los procesos de arranque del sistema. Estas bibliotecas ofrecen además un conjunto de servicios que posibilitan la integración de las funcionalidades adicionales de los módulos. Drupal es, pues, un sistema con una arquitectura modular que permite ampliar sus funcionalidades a través de unos métodos uniformes de desarrollo e integración de nuevos módulos. (34) Al mismo tiempo se puede decir que Drupal utiliza el patrón arquitectónico Modelo-Vista-Controlador (MVC). Este patrón separa todo el proceso en tres capas, típicamente son el Modelo (los datos de una aplicación), la Vista (interfaz de usuario) y el Controlador (atiende los cambios en la vista y se lo comunica al modelo).

En la Figura 6 se muestran las capas en las cuales se estructura Drupal, donde la vista son la “Lógica de los Temas” y el “JavaScript/Ajax”; el modelo es la base de datos, y el controlador es el “Núcleo de Drupal” y los “Módulos”, donde a través de las funciones *hooks*, se facilita la comunicación entre el “Núcleo de Drupal” y los “Módulos”.

CAPÍTULO 3: ANÁLISIS Y DISEÑO DEL SISTEMA

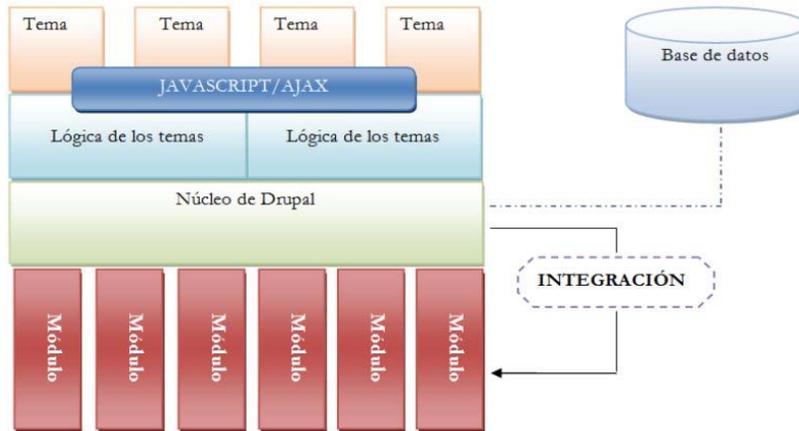


Figura 6: Arquitectura de Drupal. (34)

3.3 Modelo de Análisis

En el análisis se refinan y estructuran los requisitos obtenidos con anterioridad, de manera que sean más fáciles su comprensión, preparación, modificación y mantenimiento.

En el Modelo de Análisis están contenidas las clases de análisis, las cuales se centran en los requisitos funcionales. RUP propone clasificar dichas clases en:

Clase Entidad: Modela información que posee larga vida.



Figura 7: Ejemplo de Clase Entidad.

Clase Interfaz: Modela la interacción entre el sistema y sus actores.



Figura 8: Ejemplo de Clase Interfaz.

Clase Controladora: Coordina la realización de uno o unos pocos casos de uso coordinando las actividades de los objetos que implementan la funcionalidad del caso de uso. (35)



Figura 9: Ejemplo de Clase Controladora.

CAPÍTULO 3: ANÁLISIS Y DISEÑO DEL SISTEMA

A continuación se muestran los diagramas de clases de análisis de dos de los casos de uso más significativos, para visualizar los diagramas de clases de análisis de otros casos de uso remitirse al Anexo 3.

3.3.1 Diagramas de Clases del Análisis

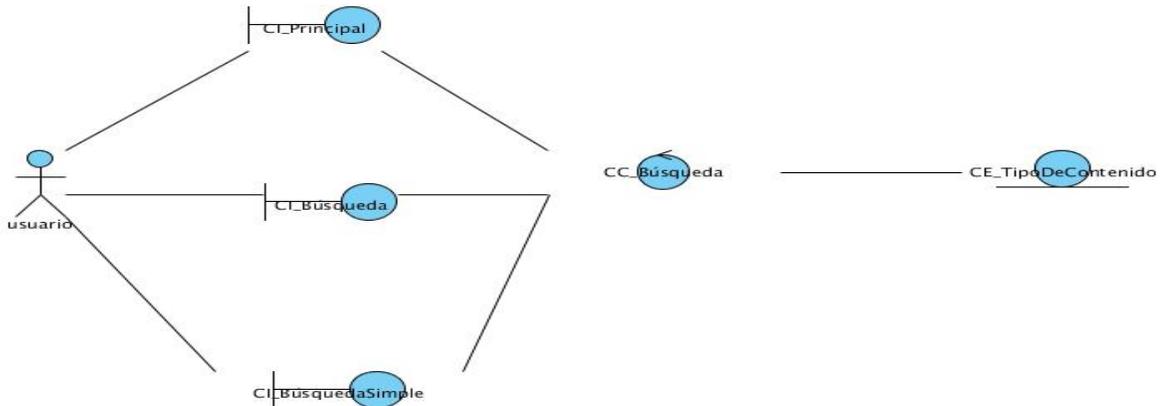


Figura 10: Diagrama de Clases de Análisis del Caso de Uso Realizar búsqueda simple.

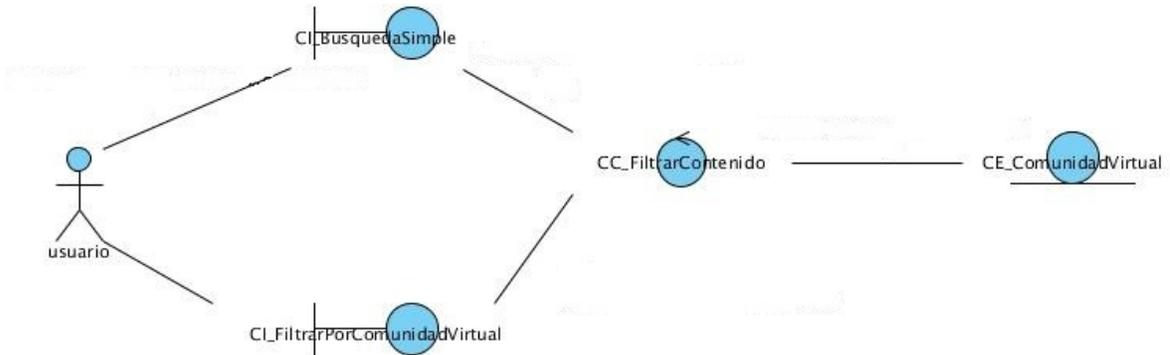


Figura 11: Diagrama de Clases de Análisis del Caso de Uso Filtrar por comunidades virtuales.

3.4 Modelo de Diseño

El diseño es un refinamiento del análisis que tiene en cuenta los requisitos no funcionales, es decir, cómo cumple el sistema sus objetivos. Proporciona una completa idea de lo que es el *software*. (35)

3.4.1 Diagramas de Interacción. Diagramas de Colaboración

Los diagramas de interacción se utilizan para modelar los aspectos dinámicos de un sistema. Esto implica modelar instancias concretas o prototípicas de clases, interfaces, componentes y nodos, junto con mensajes enviados entre ellos, todo en el contexto de un escenario que ilustra un comportamiento. Un diagrama de colaboración es un diagrama de interacción que destaca la ordenación temporal de los

CAPÍTULO 3: ANÁLISIS Y DISEÑO DEL SISTEMA

mensajes. (36) A continuación se muestran los diagramas de colaboración de dos de los casos de uso más significativos, para visualizar los diagramas de colaboración de otros casos de uso remitirse al Anexo 4.

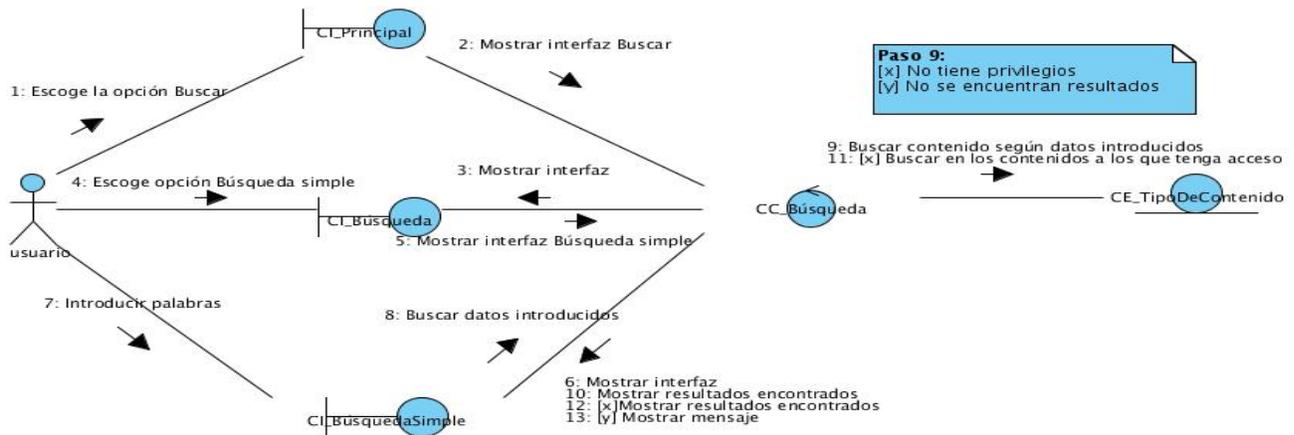


Figura 13: Diagrama de Colaboración del Caso de Uso Realizar búsqueda simple.

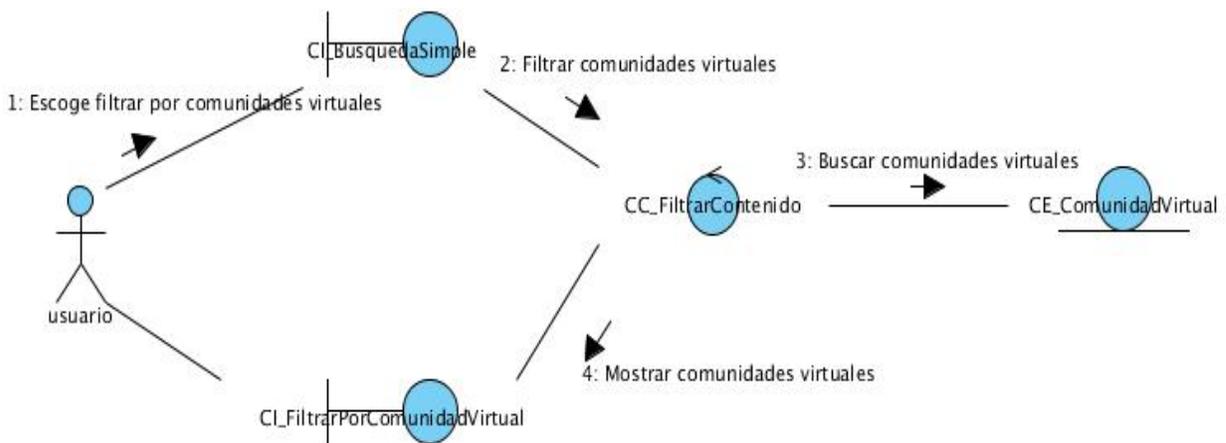


Figura 14: Diagrama de Colaboración del Caso de Uso Filtrar por comunidades virtuales.

3.4.2 Diagramas de Clases de Diseño

Un diagrama de clases se utiliza para visualizar las relaciones entre las clases que involucran el sistema. Una clase es una descripción de un conjunto de objetos que comparten los mismos atributos, operaciones, métodos, relaciones y semántica. Un diagrama de clases está compuesto por los siguientes elementos: (36)

- Clase: atributos y métodos.
- Relaciones: herencia, composición, agregación, asociación y uso.

CAPÍTULO 3: ANÁLISIS Y DISEÑO DEL SISTEMA

A continuación se muestran los diagramas de clases de diseño *web* de dos de los casos de uso más significativos, para visualizar los diagramas de clases de diseño *web* de otros casos de uso remitirse al Anexo 5.

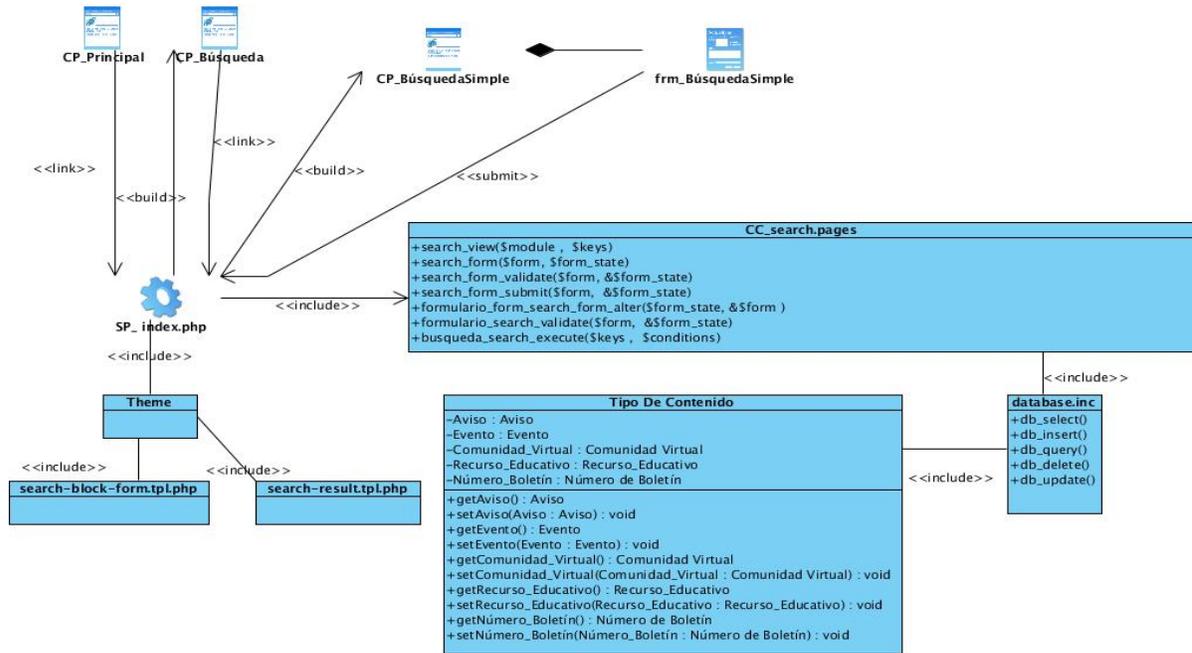


Figura 15: Diagrama de Clases del Caso de Uso Realizar búsqueda simple.

CAPÍTULO 3: ANÁLISIS Y DISEÑO DEL SISTEMA

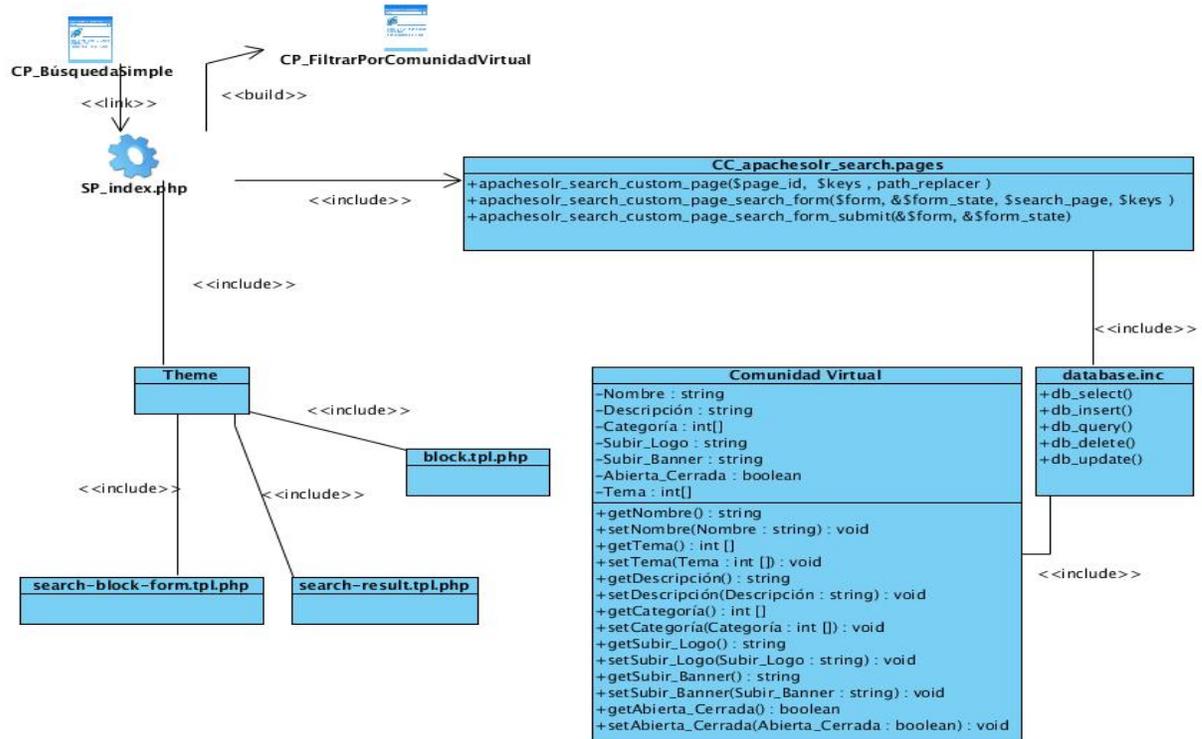


Figura 16: Diagrama de Clases del Caso Uso Filtrar por comunidades virtuales.

3.5 Diseño de la base de datos

3.5.1 Modelos de Datos

Describe la representación lógica y física de los datos persistentes usados por el sistema. A continuación se presenta el modelo de datos del sistema:

CAPÍTULO 3: ANÁLISIS Y DISEÑO DEL SISTEMA

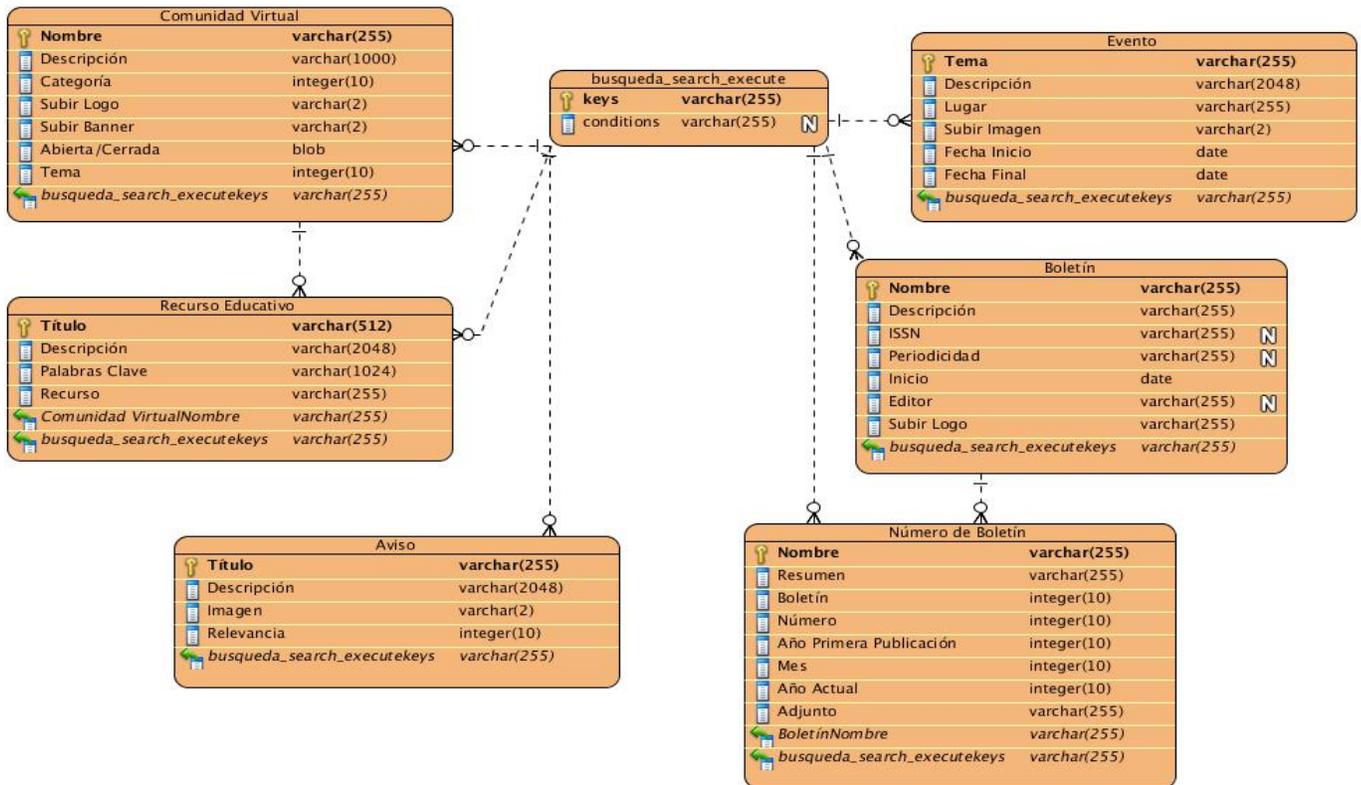


Figura 17: Modelo de Datos.

Para una mejor comprensión del modelo de datos se describen a continuación las tablas que lo componen:

- Comunidad Virtual: almacena los datos de las comunidades virtuales existentes en la base de datos y se relaciona con la tabla Recurso Educativo.
- Recurso Educativo: contiene los datos de los recursos educativos existentes en la base de datos.
- Aviso: contiene los datos de los avisos existentes en la base de datos.
- Evento: contiene los datos de los eventos existentes en la base de datos.
- Boletín: contiene los datos de los boletines existentes en la base de datos, se relaciona con la tabla Número de Boletín.
- Número de Boletín: contiene los datos de los números de boletines existentes en la base de datos.
- Busqueda_search_execute: contiene las palabras clave por las que se realiza la búsqueda en el sistema. Se relaciona con las demás tablas existentes en el modelo de datos.

CAPÍTULO 3: ANÁLISIS Y DISEÑO DEL SISTEMA

3.6 Patrones de Diseño

Los patrones de diseño son descripciones de clases cuyas instancias colaboran entre sí. Cada patrón es adecuado para ser adaptado a un cierto tipo de problema. (36)

3.6.1 Patrones GOF

En Drupal se puede percibir el uso de patrones de diseño GOF (*Gang of Four*) como son:

Patrón creacional:

- *Singleton* (instancia única): garantiza la existencia de una única instancia para una clase y la creación de un mecanismo de acceso global a dicha instancia. (37) Pensando en los módulos y temas como objetos, en general estos no encapsulan datos, los que separa a un módulo de otro es el conjunto de funciones que posee por lo que se debe pensar como una clase con un ejemplo aislado. (38)

Patrones estructurales:

- *Decorator* (Envoltorio): Añade funcionalidad a una clase dinámicamente. (37) Drupal hace un amplio uso de este patrón por ejemplo el uso de `hook_nodeapi()`, que permite a los módulos extender arbitrariamente el comportamiento de todos los nodos. (38)
- *Bridge* (Puente): Desacopla una abstracción de su implementación. (37) Los módulos necesitan ser escritos en una forma que es independiente del sistema que se está utilizando en la base de datos, y proporciona la capa de abstracción para ello. La nueva capa de base de datos se puede escribir para que conforme la API definida por el puente, añadiendo soporte para más sistemas de bases de datos sin la necesidad de modificar el código del módulo. (38)

Patrones de comportamiento:

- *Chain of Responsibility* (Cadena de responsabilidad): Permite establecer la línea que deben llevar los mensajes para que los objetos realicen la tarea indicada. (37) Se puede distinguir en el sistema de menús de Drupal, en cada solicitud de la página, el menú del sistema determina si hay un módulo para gestionar la solicitud, si el usuario tiene acceso a los recursos solicitados, y que la función se llame para hacer el trabajo. (38)
- *Command* (Orden): Encapsula una operación en un objeto, permitiendo ejecutar dicha operación sin necesidad de conocer el contenido de la misma. (37) El propio sistema de gancho (*hook*), que utiliza Drupal, usa este modelo, a fin de que los módulos no tengan que definir cada *hook*, sino solo los que vayan a aplicar. (38)

CAPÍTULO 3: ANÁLISIS Y DISEÑO DEL SISTEMA

- *Observer* (Observador): Define una dependencia de uno-a-muchos entre objetos, de forma que cuando un objeto cambie de estado se notifique y actualicen automáticamente todos los objetos que dependen de él. (37) Cuando una modificación es hecha a un vocabulario en el sistema de taxonomía de Drupal, el *hook taxonomy* es llamado en todos los módulos que lo implementan. (38)

3.7 Conclusiones parciales

En este capítulo se elaboraron los diagramas de clases del diseño, los de análisis, los de colaboración y el modelo de datos, lo que ha permitido obtener una visión del sistema, cómo este estará conformado y cómo serán las relaciones entre cada funcionalidad. Se mostró además la arquitectura presente en la aplicación, el patrón arquitectónico MVC, y los de diseño GOF, a usar durante el desarrollo del sistema. Estos patrones permiten implementar el sistema siguiendo ciertos parámetros de diseño e implementación para facilitar la realización del módulo de la forma más eficiente y rápida y haciendo uso de buenas prácticas en la codificación de este.

CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA

4.1 Introducción

Es necesario asegurar que el sistema que fue diseñado, desde sus inicios, enfocado al objetivo deseado y cumpliendo los requisitos trazados, cumpla con este diseño. Además se debe comprobar después de la realización del sistema si este cumplió consecuentemente con lo planteado con anterioridad, por lo que se deben realizar pruebas que demuestren la efectividad de lo implementado.

El objetivo del presente capítulo es abordar la implementación sobre la base de las funcionalidades, donde se describe cómo se implementan los elementos del modelo de diseño. Se modelan los diagramas de componentes y de despliegue, dando una visión de cómo quedará desarrollado el sistema. Además se especifica el tipo de prueba realizada a todas las funcionalidades necesarias.

4.2. Implementación

4.2.1 Diagrama de Despliegue

El diagrama de despliegue es utilizado para capturar los elementos de *hardware* y las conexiones entre ellos. Se utiliza para visualizar la distribución de los componentes de *software* en los nodos físicos. (39)

A continuación se muestra el diagrama de despliegue del sistema:

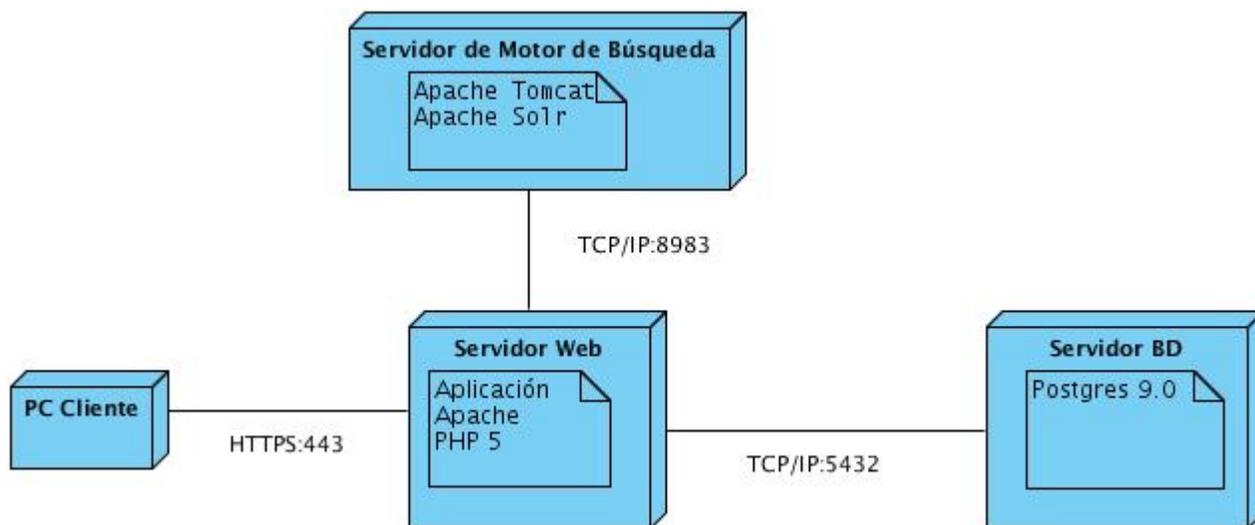


Figura 18: Diagrama de Despliegue.

- PC Cliente: se conecta al servidor web para visualizar la aplicación web.

CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA

- Servidor Web: contiene la aplicación a la cual se accederá desde las máquinas clientes y el servidor Apache.
- Servidor de Motor de Búsqueda: contiene el servidor Apache Tomcat y la herramienta Apache Solr.
- Servidor BD: contiene el gestor de base de datos PostgreSQL.

4.2.2 Diagrama de Componentes

Los diagramas de componentes muestran la organización de los componentes del sistema; modelan los aspectos físicos de este, es decir, un conjunto de elementos físicos y sus relaciones. Un componente se corresponde con una o varias clases, interfaces o colaboraciones; en este caso se corresponde con los módulos del sistema implementado.

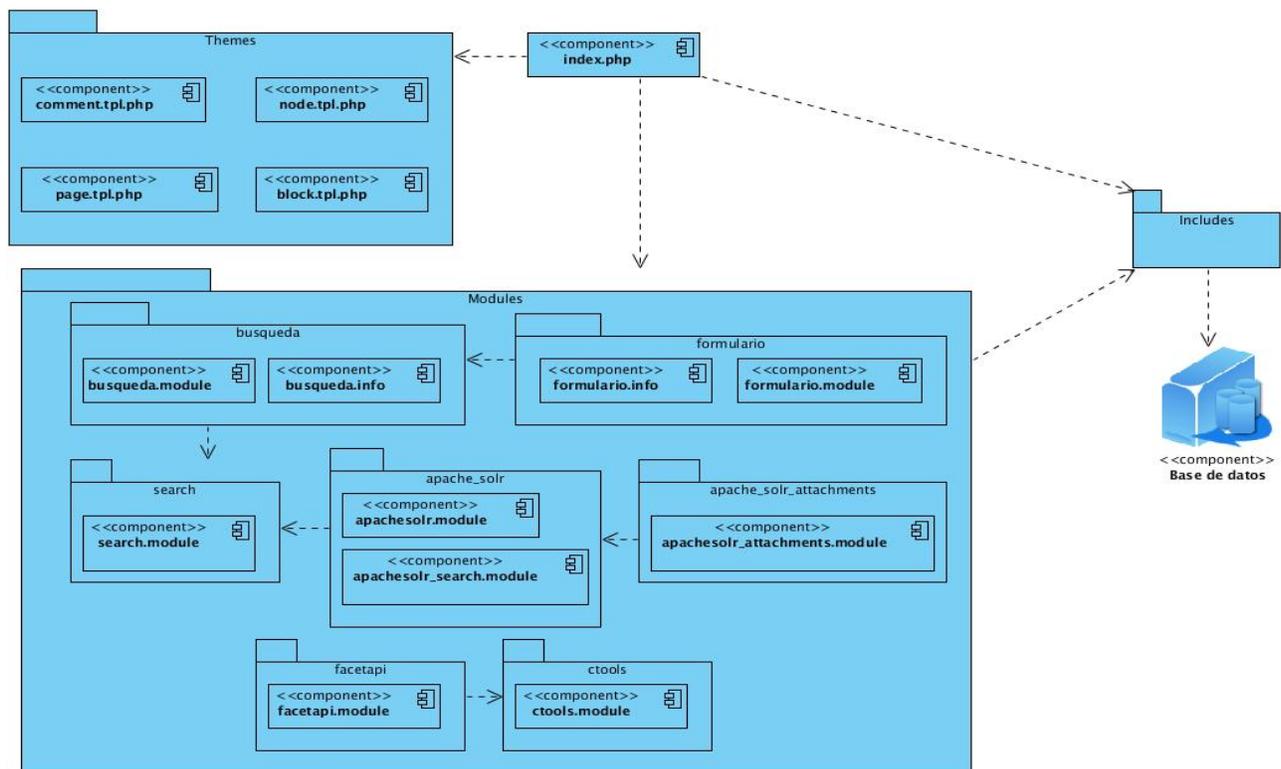


Figura 19: Diagrama de Componentes.

Themes: Contiene los elementos principales para determinar la vista del sistema.

Modules: Contiene los elementos para determinar las funcionalidades del sistema. Cada módulo aporta funcionalidades específicas, relacionadas unas con otras.

4.3 Pruebas de software

CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA

Las pruebas de *software* se definen como “el proceso de ejecución de un programa con la intención de descubrir un error.” (40) Se puede decir que no son más que la verificación constante del comportamiento del *software* a partir de un conjunto finito de casos de prueba comprobando la satisfacción de los requerimientos solicitados por el cliente. Algunas de las actividades que forman parte de estas pruebas son:

- Encontrar y documentar defectos en la calidad de *software*.
- Advertir sobre la calidad percibida del *software*.
- Validar y probar las especificaciones hechas en el diseño y especificación de requerimientos a través de una demostración concreta. (41)

4.3.1 Pruebas Unitarias

Las pruebas unitarias se pueden definir como pruebas a menor escala y son las encargadas de probar el correcto funcionamiento de un módulo de código. (42) En Drupal existen varios módulos que permiten realizar pruebas unitarias, uno de ellos se llama QUnit, en su versión 7.x-1.0-alpha1, que hace uso de QUnit, el cual es un *framework* de pruebas unitarias de JavaScript. Este módulo es una colaboración para hacer un marco de pruebas de trabajo para Drupal alrededor del marco de pruebas de QUnit. (43) Haciendo uso de este módulo se le realizaron pruebas a la aplicación cuyos resultados se pueden observar en la siguiente imagen:

CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA

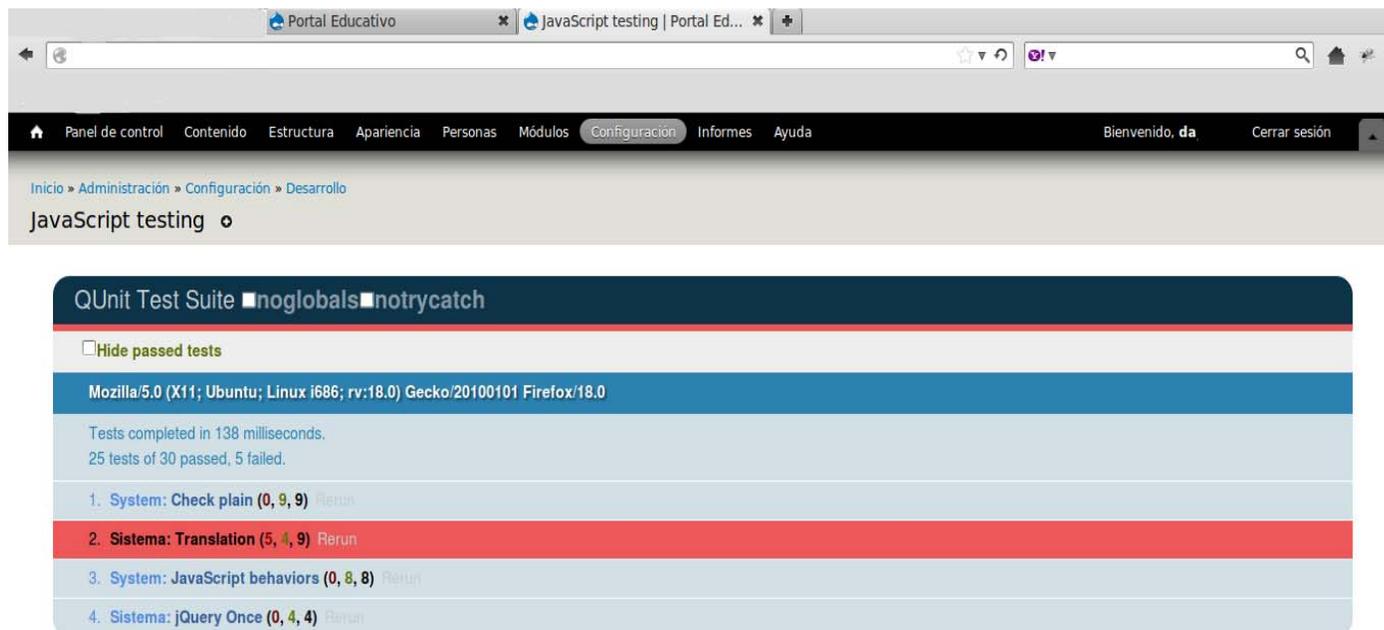


Figura 20: Pruebas usando el módulo QUnit.

Se concluye que de las 30 pruebas que el módulo QUnit realizó al sitio, 25 resultaron exitosas y 5 fallidas. No obstante, las pruebas fallidas están relacionadas con la traducción de textos de los diferentes módulos por lo que no constituyen fallas de funcionalidad. Por lo cual se concluye que las pruebas unitarias realizadas a la aplicación resultaron exitosas.

4.3.2 Modelo de prueba. Pruebas de Funcionalidad

Las pruebas de funcionalidad aseguran el trabajo apropiado de los requisitos funcionales, incluyendo la navegación, entrada de datos, procesamiento y obtención de resultados. Estas se ejecutan en cada funcionalidad de caso de uso. Tienen como objetivo conocer si mediante la introducción de datos válidos, se obtienen los resultados esperados, además de que sean mostrados los mensajes apropiados de error cuando se usan datos inválidos. (36)

Descripción de los casos de prueba

Los casos de prueba (CP) especifican una forma de probar el sistema, incluyendo las entradas con las que se ha de probar, los resultados esperados y las condiciones bajo las que ha de probarse. (36) Para comprobar el correcto funcionamiento del módulo Búsqueda se definieron un conjunto de casos de prueba que reúnen la mayor cantidad de escenarios posibles. A continuación se muestran los casos de prueba

CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA

Realizar búsqueda simple y Realizar búsqueda avanzada.

Caso de prueba Realizar búsqueda simple

Descripción general: Permite al usuario realizar una búsqueda general en los contenidos del sitio.

Escenario.	Descripción.	Palabras clave	Respuesta del sistema.	Flujo central.
EC 1.1 Se realiza la búsqueda simple satisfactoriamente.	El usuario introduce una o más palabras clave en el cuadro de texto de la búsqueda simple.	V (cualquier conjunto de caracteres)	Realiza una búsqueda general en los contenidos del portal.	1 Clic en menú Buscar. 2 Clic pestaña búsqueda simple. 3 Introducir una o varias palabras clave en el cuadro de texto de la búsqueda simple. 4 Clic en el botón de buscar.
EC 1.2 Campo vacío.	El usuario deja de insertar la palabra clave.	I (nulo)	El sistema emite un mensaje: "Por favor, escriba algunas palabras clave."	1 Clic en menú Buscar. 2 Clic pestaña búsqueda simple. 3 Introducir una o varias palabras clave en el cuadro de texto de la búsqueda simple. 4 Clic en el botón de buscar.
EC 1.3 No se encuentran resultados.	El usuario inserta una palabra clave que no se encuentra.	V (cualquier conjunto de caracteres)	El sistema emite un mensaje: "Su búsqueda no produjo resultados. Verifique que las palabras estén bien escritas. No encierre la frase entre comillas si desea buscar por cada palabra individualmente. Fabulosa búsqueda generalmente devolverá más resultados que "fabulosa búsqueda". Considere relajar su	1 Clic en menú Buscar. 2 Clic pestaña búsqueda simple. 3 Introducir una o varias palabras clave en el cuadro de texto de la búsqueda simple. 4 Clic en el botón de buscar.

CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA

			búsqueda con OR. Fabulosa OR búsqueda generalmente devolverá más resultados que fabulosa búsqueda.”	
--	--	--	---	--

Tabla 13: Caso de prueba Realizar búsqueda simple.

No	Nombre del campo.	Clasificación.	Valor Nulo.	Descripción.
1	Palabras clave.	Campo de texto.	No.	En este campo se insertarán las palabras a buscar.

Tabla 14: Descripción de las variables del CP Realizar búsqueda simple.

Caso de prueba Realizar búsqueda avanzada

Descripción general: Permite al usuario realizar una búsqueda avanzada en los contenidos del sitio.

Escenario	Descripción	Palabras clave 1	Palabras clave 2	Palabras clave 3	Tipo de contenido	Autor	Rango de fechas	Respuesta del sistema.	Flujo central.
EC 12.1 Se realiza la búsqueda satisfactoriamente.	El usuario introduce una o más palabras clave en el cuadro de texto de la búsqueda avanzada.	V (cualquier conjunto de caracteres, más de tres)	NA	NA	NA	NA	NA	Realiza una búsqueda general en los contenidos del portal.	1 Clic en menú Buscar. 2 Clic pestaña búsqueda avanzada. 3 Introducir una o varias palabras clave en el cuadro de texto de la búsqueda avanzada. 4 Clic en el botón de buscar.
EC 12.2	El usuario	V	V	V	NA	NA	NA	Realiza una	1 Clic en

CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA

Se realiza la búsqueda satisfactoriamente. Se usa combinación AND/OR.	introduce una o más palabras clave en el cuadro de texto de la búsqueda avanzada. Las combina con palabras en el otro cuadro usando AND/OR.	(cualquier conjunto de caracteres, más de tres)	(cualquier conjunto de caracteres, más de tres)	(cualquier conjunto de caracteres, más de tres)				búsqueda general en los contenidos del portal.	menú Buscar. 2 Clic pestaña búsqueda avanzada. 3 Introducir una o varias palabras clave en el cuadro de texto de la búsqueda avanzada. 4 Clic en el botón de buscar.
EC 12.3 Se realiza la búsqueda satisfactoriamente. Se filtra por tipo de contenido.	El usuario introduce una o más palabras clave en el cuadro de texto de la búsqueda avanzada. Escoge uno o más tipo de contenido.	V (cualquier conjunto de caracteres, más de tres)	NA	NA	V (una o más casillas seleccionadas)	NA	NA	Realiza una búsqueda general en los contenidos del portal.	1 Clic en menú Buscar. 2 Clic pestaña búsqueda avanzada. 3 Introducir una o varias palabras clave en el cuadro de texto de la búsqueda avanzada. 4 Clic en el botón de buscar.
EC 12.4 Se	El usuario introduce	V (cual	NA	NA	NA	V (un	NA	Realiza una búsqueda	1 Clic en menú

CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA

realiza la búsqueda satisfactoriamente. Se filtra por autor.	una o más palabras clave en el cuadro de texto de la búsqueda avanzada . Escoge el autor.	quier conjunto de caracteres , más de tres)				autor seleccionado)		general en los contenidos del portal.	Buscar. 2 Clic pestaña búsqueda avanzada . 3 Introducir una o varias palabras clave en el cuadro de texto de la búsqueda avanzada . 4 Clic en el botón de buscar.
EC 12.5 Se realiza la búsqueda satisfactoriamente. Se filtra por rango de fechas.	El usuario introduce una o más palabras clave en el cuadro de texto de la búsqueda avanzada . Escoge un rango de fechas.	V (cualquier conjunto de caracteres , más de tres)	NA	NA	NA	NA	V (un rango de fechas seleccionado)	Realiza una búsqueda general en los contenidos del portal.	1 Clic en menú Buscar. 2 Clic pestaña búsqueda avanzada . 3 Introducir una o varias palabras clave en el cuadro de texto de la búsqueda avanzada . 4 Clic en el botón de buscar.
EC 12.6 Campo vacío.	El usuario deja de insertar la	I (nulo)	I (nulo)	I (nulo)	NA	NA	NA	El sistema emite un mensaje:	1 Clic en menú Buscar.

CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA

	palabra clave.							“Por favor, escriba algunas palabras clave.”	2 Clic pestaña búsqueda avanzada . 3 Introducir una o varias palabras clave en el cuadro de texto de la búsqueda avanzada . 4 Clic en el botón de buscar.
EC 12.7 Palabra corta.	El usuario inserta la palabra clave con menos de 3 caracteres.	I (cualquier conjunto de caracteres, menos de tres)	I (cualquier conjunto de caracteres, menos de tres)	I (cualquier conjunto de caracteres, menos de tres)	NA	NA	NA	El sistema emite un mensaje: “Debe incluir al menos una palabra clave positiva con 3 o más caracteres.”	1 Clic en menú Buscar. 2 Clic pestaña búsqueda avanzada . 3 Introducir una o varias palabras clave en el cuadro de texto de la búsqueda avanzada . 4 Clic en el botón de buscar.
EC 12.8 No se encuentran	El usuario inserta una palabra	V (cualquier conjunto	NA	NA	NA	NA	NA	El sistema emite un mensaje: “Su	1 Clic en menú Buscar. 2 Clic

CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA

resultados.	clave que no se encuentra .	nto de caracteres , más de tres)						búsqueda no produjo resultados. Verifique que las palabras estén bien escritas. No encierre la frase entre comillas si desea buscar por cada palabra individualmente. Fabulosa búsqueda generalmente devolverá más resultados que "fabulosa búsqueda". Considere relajar su búsqueda con OR. Fabulosa OR búsqueda generalmente devolverá más resultados que fabulosa búsqueda."	pestaña búsqueda avanzada . 3 Introducir una o varias palabras clave en el cuadro de texto de la búsqueda avanzada . 4 Clic en el botón de buscar.
EC 12.9 Rango de fechas incorrecto.	El usuario introduce el rango de fechas incorrectamente (no cronológico)	NA	NA	NA	NA	NA	I (primera fecha más reciente)	El sistema emite un mensaje: "El rango de fecha está incorrecto,	1 Clic en menú Buscar. 2 Clic pestaña búsqueda avanzada

CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA

	amente).					que la segunda)	escoger la primera fecha menor o igual que la segunda.”	. 3 Introducir una o varias palabras clave en el cuadro de texto de la búsqueda avanzada . 4 Clic en el botón de buscar.
--	----------	--	--	--	--	-----------------	---	--

Tabla 15: Caso de prueba Realizar búsqueda avanzada.

No.	Nombre del campo.	Clasificación.	Valor Nulo.	Descripción.
1	Palabras clave 1.	Campo de texto.	Sí. (Solo si alguno de los otros dos campos de palabras clave no es nulo)	En este campo se insertarán las palabras a buscar. No debe ser menor de tres caracteres.
2	Palabras clave 2.	Campo de texto.	Sí. (Solo si alguno de los otros dos campos de palabras clave no es nulo)	En este campo se insertarán las palabras a buscar. No debe ser menor de tres caracteres.
3	Palabras clave 3.	Campo de texto.	Sí. (Solo si alguno de los otros dos campos de palabras clave no es nulo)	En este campo se insertarán las palabras a buscar. No debe ser menor de tres caracteres.
4	Tipo de contenido.	Casillas de selección.	Sí.	Esta opción está desactivada por defecto, se puede activar haciendo clic en una o varias casillas.
5	Autor.	Lista de selección.	Sí.	Esta opción está desactivada por defecto, se

CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA

				puede activar haciendo clic en una de las opciones de la lista.
6	Rango de fechas.	Listas de selección.	Sí.	Esta opción está desactivada por defecto, se puede activar haciendo clic en una de las opciones de cada una de las listas.

Tabla 16: Descripción de las variables del CP Realizar búsqueda avanzada.

Resultados de las pruebas de funcionalidad

Se realizaron 4 iteraciones, donde se identificaron diez no conformidades (ver tabla 17), las cuales fueron disminuyendo considerablemente con cada iteración realizada. En cada iteración se muestra la cantidad de No conformidades, las que se le dieron respuesta (Cerradas) y las que por un motivo justificado no pueden ser resueltas (No proceden). Los resultados arrojados por el sistema al final de cada prueba aplicada han sido satisfactorios, indicando la calidad y validez que posee el mismo.

Iteración	No Conformidades	Cerrada (s)	No procede (n)
1ra	10	5	5
2da	5	4	1
3ra	1	1	0
4ta	0	-	-

Tabla 17: Iteraciones de las pruebas de funcionalidad.

Después de terminada la implementación del módulo de búsqueda este se le integró al portal y se seleccionaron 15 usuarios para que se conectaran a la aplicación durante 4 días. Los dos primeros días sin el módulo y los dos últimos días con el módulo incluido.

A continuación se muestra una tabla con el comportamiento de las variables durante los cuatro días.

Variables	Día 1	Día 2	Día 3	Día 4
Tiempo promedio para encontrar la información.	10 min	8 min	30 seg	30 seg
Información irrelevante encontrada (por	85	89	3	2.5

CAPÍTULO 4: IMPLEMENTACIÓN Y PRUEBA

ciento).				
Consultas promedio a los contenidos.	30	25	40	60
Uso de filtros en la búsqueda (por ciento).	0	0	53	92
Uso de la búsqueda avanzada (por ciento).	0	0	20	26

Tabla 18: Comportamiento de las variables de la hipótesis.

Como se puede apreciar en la tabla anterior después de los cuatro días de probado el módulo se logró facilitar la búsqueda de información a los usuarios y estos usaron los filtros para recuperar información por lo que la hipótesis planteada es válida.

4.4 Conclusiones parciales

En este capítulo se han desarrollado los aspectos que muestran la implementación del sistema. Se realizó el diagrama de despliegue así como el diagrama de componentes, con lo que se mostraron las relaciones entre los elementos de *hardware*, y además el conjunto de componentes del sistema (módulos y ficheros de implementación) y cómo se relacionan entre ellos. Se efectuaron pruebas, las unitarias arrojaron resultados relevantes, ya que permitieron validar el código fuente desarrollado durante la etapa de implementación del sistema. Igualmente se realizaron pruebas de funcionalidad (en estas últimas se detectaron diez no conformidades que fueron resueltas), teniendo como objetivo, estas pruebas: comprobar el correcto funcionamiento del sistema y certificando que el mismo cumple los requisitos establecidos.

CONCLUSIONES

Con el objetivo de dar cumplimiento al objetivo general y a la problemática planteada en este trabajo, se han llevado a cabo cada una de las tareas que fueron trazadas al comienzo del mismo:

- Se revisó la bibliografía actualizada para generar el marco teórico-conceptual. Con ello se logró conocer el estado del arte de las herramientas, librerías, técnicas y modelos utilizados para la RI así como las herramientas, tecnologías y metodologías para el desarrollo del módulo.
- Se analizaron las herramientas y bibliotecas actuales utilizadas para la recuperación de información y se pudo concluir que se integraría al proyecto el motor de búsqueda Apache Solr que utiliza como base la biblioteca Lucene, para permitir la búsqueda facetada.
- Se estudiaron las técnicas y modelos de RI; los lenguajes y tipos de búsqueda que existen. Además de los portales educativos, sus tendencias presentes y las tecnologías para su desarrollo actual. Con este estudio se obtuvo una visión global de cómo funcionan internamente las herramientas y bibliotecas de búsqueda especialmente las seleccionadas, así como las tendencias en la creación de portales para un mejor entendimiento de cómo desarrollar el módulo para el portal educativo.
- Se elaboraron los artefactos según la metodología RUP los cuales permitieron una mejor comprensión de cómo se debía desarrollar el módulo cumpliendo con los requisitos de para su creación y despliegue.
- Se implementaron los requerimientos definidos, para lo cual se desarrolló el módulo de búsqueda apoyado en los diagramas creados como parte de los flujos de trabajo análisis y diseño e implementación. Este módulo hace uso de la búsqueda facetada, integrando la herramienta Apache Solr a la arquitectura del portal, además de permitir la búsqueda avanzada.
- Se realizaron pruebas al sistema a través de las pruebas unitarias, que certificaron la utilización del código fuente en JavaScript, y las pruebas de funcionalidad con las que se detectaron y posteriormente erradicaron un conjunto de diez deficiencias. Estas pruebas permitieron entregar un módulo con la calidad requerida para ser integrado al portal.

Para la construcción y desarrollo del sistema se utilizaron las herramientas y tecnologías seleccionadas como son: lenguajes de programación: HTML, JavaScript, PHP y CSS; además del CMS Drupal. Para el acceso a datos, el gestor de base de datos PostgreSQL; la herramienta de modelado utilizada fue el

CONCLUSIONES

Visual Paradigm. Al mismo tiempo, se desarrolló el sistema haciendo uso de la metodología de desarrollo RUP.

Por lo anteriormente expuesto se puede concluir que los objetivos propuestos para el presente trabajo han sido cumplidos satisfactoriamente. Se incluyen además una serie de recomendaciones que se deben tener en cuenta para el trabajo futuro.

RECOMENDACIONES

Después de terminado el desarrollo del sistema se detectaron algunas carencias que pueden incluirse como posibles mejoras al mismo. A continuación se detallan las recomendaciones realizadas al sistema:

1. Integración de la búsqueda avanzada con la utilización de la herramienta Apache Solr.
2. Creación de otros filtros que permitan al usuario simplificar más su búsqueda.
3. Mejoras en la presentación de los resultados y de los filtros.

REFERENCIAS BIBLIOGRÁFICAS

1. **Glosario de Internet y World Wide Web.** Definición de Portal. [En línea] 2007. <http://www.definicion.org/portal>.
2. **Lujo Aliaga Libely, Zenoyda y Cedeño Galindo, Victoria.** Tesis "DISEÑO DE UNA ESTRATEGIA DE DIRECCIÓN PARA LA IMPLEMENTACIÓN DE PORTALES EDUCATIVOS". La Habana : Universidad de las Ciencias Informáticas, 2007. pág. 87.
3. **BARRIO, F. G.** *Los Portales Educativos como fuentes de Recursos Materiales.* s.l. : ICONO 14, 2006.
4. *Modern Information Retrieval.* **Baeza-Yates, Ricardo.** 1999.
5. **Universidad de las Ciencias Informáticas.** Sitio UCI de Tesis. *Documentación sobre Metodología de la Investigación.* [En línea] [Citado el: 2 de noviembre de 2012.] http://tesis.uci.cu/viewpage.php?page_id=21.
6. **Pinto Molina, Maria.** BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN. [En línea] Universidad de Granada, 13 de abril de 2011. [Citado el: 23 de noviembre de 2012.] http://www.mariapinto.es/e-coms/recu_infor.htm.
7. **Claudia, Salaverry Sayán.** Monografías.com. *Recuperación de Información en Internet.* [En línea] [Citado el: 21 de Noviembre de 2012.] http://www.Recuperación%20de%20Información%20en%20Internet%20-%20Monografias_com.htm.
8. **Parra Zambrano, Eduardo.** Monografías.com. *Comunidades Virtuales de Aprendizaje.* [En línea] [Citado el: 23 de noviembre de 2012.] <http://www.monografias.com/trabajos82/comunidades-virtuales-de-apre>.
9. **Singhal, Amit.** «Modern Information Retrieval: A Brief Overview». s.l. : Bulletin of the IEEE Computer Society Technical 24, 2001, págs. pp. 35–43.
10. **Colectivo de Autores.** *Librerías Lucene y dotLucene para Recuperación de Información. Estudio y desarrollo de casos prácticos.* s.l. : Departamento de Informática y Automática Universidad de Salamanca, 2007.
11. **Mercader Guevara, Alberto.** Tesis "Implementación de una herramienta de recuperación de información dentro de la Informática Jurídica.". La Habana : Universidad de las Ciencias Informáticas, 2010. pág. 67.
12. **Castillo Duvergel, Yoannia y Arenas Pérez, Raidel.** Tesis: "Implementación de algoritmos de recomendaciones de imágenes para textos noticiosos.". La Habana : Universidad de las Ciencias Informáticas, 2009. pág. 72.
13. **Valle, Mireldis García del.** Tesis *Metodología para la evaluación de Sistemas de Recuperación de Información Web en la Universidad de ciencias Informáticas.* Universidad de las Ciencias Informáticas. La Habana : s.n., 2009. pág. 108.
14. **De Seta, Leonardo.** Dos Ideas. *Apache solr: una introduccion.* [En línea] 11 de junio de 2010. [Citado el: 6 de noviembre de 2012.] <http://www.dosideas.com/noticias/java/913-apache-solr-una-introduccion.html>.
15. **Fernández Luna, Juan Manuel, y otros.** Empleo de motores de búsqueda de código abierto para la recuperación de información vertical. Enero-Junio de 2009. Vol. 3, 1-2.
16. **PostgresSQL.** Sitio oficial de PostgresSQL. [En línea] 2011. [Citado el: 23 de noviembre de 2012.] <http://www.postgresql.org/docs/8.4/static/textsearch-dictionaries.html>.
17. **Pérez Leiva, Yidielis y Bello González, Sergio.** Tesis *Implementación de una nueva versión del módulo Búsqueda para el Repositorio de Objetos de Aprendizaje RHODA.* Universidad de las Ciencias Informáticas. La Habana : s.n., 2011. pág. 80.
18. **Widmayer, Peter y Ottmann, Thomas.** *Algorithmen und Datenstrukturen.* cuarta edición. s.l. :

REFERENCIAS BIBLIOGRÁFICAS

Spektrum Akademischer Verlag, 2002. ISBN 3-8274-1029-0.

19. **Lara Sierra, Joaquin.** Educación Virtual . *¿Qué son las Comunidades Virtuales?* . [En línea] 2007. [Citado el: 23 de noviembre de 2012.] <http://educacionvirtuall.blogspot.com/2007/04/qu-son-las-comunidades-virtuales.html>.
20. **Buytaert, Dries.** Página del autor de Drupal. [En línea] 1999. <http://buytaert.net/>.
21. **R. P. I. Carrillo, A. Rodríguez.** Metodología de Desarrollo del Software. [En línea] 2008. [Citado el: 30 de octubre de 2012.] <http://www.google.com.cu/url?sa=t&rct=j&q=Metodolog%C3%ADa+de+Desarrollo+del+Software.+2008..>
22. **Universidad de las Ciencias Informáticas.** EVA. *Conferencia #1. Introducción a la Ingeniería de Software, ISW 1.* [En línea]
23. **Alvarez, Miguel Angel.** Qué es HTML 5. [En línea] 14 de Octubre de 2009. [Citado el: 13 de Mayo de 2013.] <http://www.desarrolloweb.com/articulos/que-es-html5.html>.
24. **Hernández Carballo, Sandy.** Tesis: "Propuesta para la recuperación de información proveniente de Internet para el Sistema de Información de Mercadotecnia de Albet S.A.". La Habana : Universidad de las Ciencias Informáticas, 2012. pág. 105.
25. **PHP.** Sitio Oficial de PHP. [En línea] [Citado el: 12 de octubre de 2012.] <http://www.php.net/>.
26. **Eguíluz Pérez, Javier.** *Introducción a CSS.*
27. **Object Management Group, Inc.** Sitio Oficial de UML. [En línea] 1997. [Citado el: 13 de octubre de 2012.] <http://www.uml.org/>.
28. **Pereira, M.** UML. [En línea] 2007. [Citado el: 30 de octubre de 2012.] <http://es.scribd.com/doc/2080534/UML>.
29. **Sitio Oficial de Netbeans.** Sitio Oficial de Netbeans. [En línea] [Citado el: 12 de octubre de 2012.] <https://netbeans.org/>.
30. **Sitio web del Proyecto Apache.** Proyecto Apache. [En línea] <http://www.apache.org/>.
31. **WORSLEY, John C. y DRAKE, Joshua D.** Practical PostgreSQL. s.l. : O'Reilly Media, Inc., 2011.
32. **Pressman, Roger S.** *Ingeniería de Software, un enfoque práctico. Quinta edición.* s.l. : McGraw-Hill Companie, 2002. ISBN: 8448132149.
33. **Universidad Yacambu.** "MODELADO DE SISTEMAS DE INFORMACIÓN". [En línea] [Citado el: 16 de Mayo de 2013.] <http://www.oocities.org/es/monsalvelaura/fase2/analisis.html>.
34. **Gil Rodríguez, Fran.** *Experto de Drupal 7. Nivel Avanzado.* s.l. : Forcontu , 2011. págs. 33-47. ISBN-13.
35. **Álvarez Valdés, Ana María y Montes de Oca Rodríguez, Osmany.** Tesis: "*Desarrollo del módulo Ejercicios de la Colección Multisaber en su versión multiplataforma*". Universidad de Ciencias Informáticas. La Habana : s.n., 2010. pág. 89.
36. **Ortega Aponte, Graviel y Guerra Reyes, Yailemi.** Tesis: *BluePub Sender: Sistema de Envío de Publicidad vía Bluetooth.* La Habana : Universidad de las Ciencias Informáticas, 2012. pág. 117.
37. **Pérez Mariñán, Martín.** "Patrones de Diseño (Design Patterns)". [En línea] [Citado el: 9 de Marzo de 2013.] eupla.org/apuntes/doc_download/293-patrones.html.
38. **Martín, Liber Matos.** Se soluciona programando. *POO y patrones de diseño en DRUPAL* . [En línea] 2010. [Citado el: 9 de Marzo de 2013.] <http://jitcode.blogspot.com/2010/06/poo-y-patrones-de-diseno-en-drupal.html>.
39. **JACOBSON, Ivar, RUMBAUGH, James y BOOCH, Grady.** "El proceso unificado de desarrollo". s.l. : Addison Wesley, 2000, 9.
40. **Pressman, R. S.** *Ingeniería del Software, un enfoque práctico.* Habana : s.n., 2005.
41. **Huerta Carralero, Ismary.** Tesis: "*Propuesta de un proceso para la realización de pruebas a*

REFERENCIAS BIBLIOGRÁFICAS

aplicaciones de software.". La Habana : Universidad de las Ciencias Informáticas., 2009. pág. 98.

42. **Rodríguez Jiménez, Arianna.** *Tesis: "Proceso de evaluación de la documentación para el servicio de Pruebas funcionales a nivel de Sistema que ofrece el Grupo de Calidad de FORTES."*. La Habana : Universidad de las Ciencias Informáticas, 2011. pág. 72.

43. **Comunidad de Drupal.** Sitio Oficial de Drupal. *QUnit.* [En línea] 2 de Septiembre de 2009. [Citado el: 9 de Mayo de 2013.] <http://drupal.org/project/qunit>.

BIBLIOGRAFÍA

1. Eguíluz Pérez, Javier. Introducción a CSS.
2. Sitio web del Proyecto Apache. Proyecto Apache. [En línea] <http://www.apache.org/>.
3. Gil Rodríguez, Fran. Experto de Drupal 7. Nivel Avanzado. s.l. : Forcontu , 2011. ISBN-13.
4. Pressman, R. S. Ingeniería del Software, un enfoque práctico. Habana : s.n., 2005.
5. Comunidad de Drupal. Sitio Oficial de Drupal. [En línea] <http://drupal.org>.
6. Drupal Api. Api reference. [En línea] <http://drupaleros.uci.cu/drupalapi/api/drupal>.
7. La Comunidad UCI de Drupal. Drupaleros. [En línea] <http://drupaleros.uci.cu/>.

ANEXOS

Anexo 1: Definición de Casos de Uso del Sistema

CU-3	Filtrar por boletín.
Actor	Usuario.
Descripción	El caso de uso le permite al usuario filtrar la búsqueda por los números de boletines.
Referencia	RF-3, CU-1

Tabla 1: Definición de CU Filtrar por boletín.

CU-4	Filtrar por avisos.
Actor	Usuario.
Descripción	El caso de uso le permite al usuario filtrar la búsqueda por avisos.
Referencia	RF-4, CU-1

Tabla 2: Definición de CU Filtrar por avisos.

CU-5	Filtrar por eventos.
Actor	Usuario.
Descripción	El caso de uso le permite al usuario filtrar la búsqueda por eventos.
Referencia	RF-5, CU-1

Tabla 3: Definición de CU Filtrar por eventos.

CU-6	Filtrar por recursos educativos.
Actor	Usuario registrado.
Descripción	El caso de uso le permite al usuario registrado y con los permisos apropiados filtrar la búsqueda por recursos educativos.
Referencia	RF-6, CU-1

Tabla 4: Definición de CU Filtrar por recursos educativos.

CU-7	Filtrar por clasificación de eventos.
Actor	Usuario.
Descripción	El caso de uso le permite al usuario filtrar los eventos por autor, fecha de envío o fecha de modificación.
Referencia	RF-7, CU-5

Tabla 5: Definición de CU Filtrar por clasificación de eventos.

CU-8	Filtrar por clasificación de avisos.
Actor	Usuario.
Descripción	El caso de uso le permite al usuario filtrar los avisos por autor, fecha de envío o fecha de modificación.
Referencia	RF-8, CU-4

Tabla 6: Definición de CU Filtrar por clasificación de avisos.

CU-9	Filtrar por clasificación de boletín.
Actor	Usuario.
Descripción	El caso de uso le permite al usuario filtrar los números de boletines por autor, año, mes, número, fecha de envío o fecha de modificación.
Referencia	RF-9, CU-3

Tabla 7: Definición de CU Filtrar por clasificación de boletín.

CU-11	Filtrar por clasificación de recurso educativo.
Actor	Usuario registrado.
Descripción	El caso de uso le permite al usuario registrado y con los permisos apropiados filtrar los recursos educativos por autor, fecha de envío o fecha de modificación.
Referencia	RF-11, CU-6

Tabla 8: Definición de CU Filtrar por clasificación de recurso educativo.**Anexo 2: Descripción de Casos de Uso del Sistema**

Caso de uso	
CU-3	Filtrar por boletín.
Propósito	Filtrar por los números de boletines la búsqueda que se va a realizar.
Actores: Usuario.	
Resumen: El caso de uso se inicia cuando el usuario selecciona la opción que le permite filtrar la búsqueda que va a realizar por los números de boletines. El sistema restringe la búsqueda a los números de boletines.	
Referencias	RF-3, CU-1
Acción del actor	Respuesta del sistema

1 El caso de uso se inicia cuando el actor selecciona la opción de filtrar por los números de boletines.	2 Se filtra la búsqueda para cuando se vaya a realizar solo busque en los números de boletines.
3 Puede ahora realizar una búsqueda filtrada solo en los números de boletines. (Para mantener los filtros seleccionados debe dar clic en la casilla para mantener los filtros anteriores.)	4 El CU termina.
Puntos de extensión	
Línea 3 el usuario puede restringir más la búsqueda. Ver CU-9.	

Tabla 1: Descripción del Caso de Uso Filtrar por boletín.

Caso de uso	
CU-4	Filtrar por avisos.
Propósito	Filtrar por avisos la búsqueda que se va a realizar.
Actores: Usuario.	
Resumen: El caso de uso se inicia cuando el usuario selecciona la opción que le permite filtrar la búsqueda que va a realizar por avisos. El sistema restringe la búsqueda a los avisos solamente.	
Referencias	RF-4, CU-1
Acción del actor	Respuesta del sistema
1 El caso de uso se inicia cuando el actor selecciona la opción de filtrar por avisos.	2 Se filtra la búsqueda para cuando se vaya a realizar solo busque en los avisos.
3 Puede ahora realizar una búsqueda filtrada solo en los avisos. (Para mantener los filtros seleccionados debe dar clic en la casilla para mantener los filtros anteriores.)	4 El CU termina.
Puntos de extensión	
Línea 3 el usuario puede restringir más la búsqueda. Ver CU-8.	

Tabla 2: Descripción del Caso de Uso Filtrar por avisos.

Caso de uso	
CU-5	Filtrar por eventos.
Propósito	Filtrar por eventos la búsqueda que se va a realizar.

Actores: Usuario.	
Resumen: El caso de uso se inicia cuando el usuario selecciona la opción que le permite filtrar la búsqueda que va a realizar por eventos. El sistema restringe la búsqueda a los eventos solamente.	
Referencias	RF-5, CU-1.
Acción del actor	Respuesta del sistema
1 El caso de uso se inicia cuando el actor selecciona la opción de filtrar por eventos.	2 Se filtra la búsqueda para cuando se vaya a realizar solo busque en los eventos.
3 Puede ahora realizar una búsqueda filtrada solo en los eventos. (Para mantener los filtros seleccionados debe dar clic en la casilla para mantener los filtros anteriores.)	4 El CU termina.
Puntos de extensión	
Línea 3 el usuario puede restringir más la búsqueda. Ver CU-7.	

Tabla 3: Descripción del Caso de Uso Filtrar por eventos.

Caso de uso	
CU-6	Filtrar por recursos educativos.
Propósito	Filtrar por recursos educativos la búsqueda que se va a realizar.
Actores: Usuario registrado.	
Resumen: El caso de uso se inicia cuando el usuario registrado selecciona la opción que le permite filtrar la búsqueda que va a realizar por recursos educativos. El sistema restringe la búsqueda a los recursos educativos solamente.	
Referencias	RF-6, CU-1.
Acción del actor	Respuesta del sistema
1 El caso de uso se inicia cuando el actor selecciona la opción de filtrar por recursos educativos.	2 Se filtra la búsqueda para cuando se vaya a realizar solo busque en los recursos educativos.

3 Puede ahora realizar una búsqueda filtrada solo en los recursos educativos. (Para mantener los filtros seleccionados debe dar clic en la casilla para mantener los filtros anteriores.)	4 El CU termina.
Puntos de extensión	
Línea 3 el usuario puede restringir más la búsqueda. Ver CU-11.	

Tabla 4: Descripción del Caso de Uso Filtrar por recursos educativos.

Caso de uso	
CU-7	Filtrar por clasificación de eventos.
Propósito	Filtrar la búsqueda realizada en los eventos, por ciertos campos de los eventos.
Actores: Usuario.	
Resumen: El caso de uso se inicia cuando el usuario selecciona la opción que le permite restringir la búsqueda realizada en los eventos por autor, fecha de envío o fecha de modificación. Además de poder escoger otro tipo de contenido que quiera filtrar. El sistema restringe la búsqueda según la opción escogida o busca en el otro tipo de contenido seleccionado.	
Referencias	RF-7, CU-5.
Acción del actor	Respuesta del sistema
1 El caso de uso se inicia cuando el actor después de haber realizado una búsqueda inicial filtrada por eventos desea restringir más la búsqueda y selecciona una de las opciones presentadas: <ul style="list-style-type: none"> ✓ Autor. ✓ Fecha de envío. ✓ Fecha de modificación. O puede buscar por otro tipo de contenido.	2 Se restringe la búsqueda en los eventos según la opción escogida. O busca en el tipo de contenido seleccionado.
3 Ahora puede buscar en la nueva restricción escogida. (Para mantener los filtros seleccionados debe dar clic en la casilla para mantener los filtros anteriores.)	4 El CU termina.

Tabla 5: Descripción del Caso de Uso Filtrar por clasificación de eventos.

CU-8	Filtrar por clasificación de avisos.	
Propósito	Filtrar la búsqueda realizada en los avisos, por ciertos campos de los avisos.	
Actores: Usuario.		
Resumen: El caso de uso se inicia cuando el usuario selecciona la opción que le permite restringir la búsqueda realizada en los avisos por autor, fecha de envío o fecha de modificación. Además de poder escoger otro tipo de contenido que quiera filtrar. El sistema restringe la búsqueda según la opción escogida o busca en el otro tipo de contenido seleccionado.		
Referencias	RF-8, CU-4	
Acción del actor	Respuesta del sistema	
1 El caso de uso se inicia cuando el actor después de haber realizado una búsqueda inicial filtrada por avisos desea restringir más la búsqueda y selecciona una de las opciones presentadas: <ul style="list-style-type: none"> ✓ Autor. ✓ Fecha de envío. ✓ Fecha de modificación. O puede buscar por otro tipo de contenido.	2 Se restringe la búsqueda en los avisos según la opción escogida. O busca en el tipo de contenido seleccionado.	
3 Ahora puede buscar en la nueva restricción escogida. (Para mantener los filtros seleccionados debe dar clic en la casilla para mantener los filtros anteriores.)	4 El CU termina.	

Tabla 6: Descripción del Caso de Uso Filtrar por clasificación de avisos.

Caso de uso	
CU-9	Filtrar por clasificación de boletín.
Propósito	Filtrar la búsqueda realizada en los boletines, por ciertos campos de los boletines.
Actores: Usuario.	
Resumen: El caso de uso se inicia cuando el usuario selecciona la opción que le permite restringir la búsqueda realizada en los números de boletines por autor, año, mes, número, fecha de envío o fecha de modificación. Además de poder escoger otro tipo de contenido que quiera filtrar. El sistema restringe la búsqueda según la opción escogida o busca en el otro tipo de contenido seleccionado.	
Referencias	RF-9, CU-3.

Acción del actor	Respuesta del sistema
<p>1 El caso de uso se inicia cuando el actor después de haber realizado una búsqueda inicial filtrada por los números de boletines desea restringir más la búsqueda y selecciona una de las opciones presentadas:</p> <ul style="list-style-type: none"> ✓ Autor. ✓ Año. ✓ Mes. ✓ Número. ✓ Fecha de envío. ✓ Fecha de modificación. <p>O puede buscar por otro tipo de contenido.</p>	<p>2 Se restringe la búsqueda en los boletines según la opción escogida. O busca en el tipo de contenido seleccionado.</p>
<p>3 Ahora puede buscar en la nueva restricción escogida. (Para mantener los filtros seleccionados debe dar clic en la casilla para mantener los filtros anteriores.)</p>	<p>4 El CU termina.</p>

Tabla 7: Descripción del Caso de Uso Filtrar por clasificación de boletín.

Caso de uso	
CU-11	Filtrar por clasificación de recurso educativo.
Propósito	Filtrar la búsqueda realizada en los recursos educativos, por ciertos campos de los recursos educativos.
Actores: Usuario registrado.	
Resumen: El caso de uso se inicia cuando el usuario registrado selecciona la opción que le permite restringir la búsqueda realizada en los recursos educativos por autor, fecha de envío o fecha de modificación. Además de poder escoger otro tipo de contenido que quiera filtrar. El sistema restringe la búsqueda según la opción escogida o busca en otro tipo de contenido seleccionado.	
Referencias	RF-11, CU-6.
Acción del actor	Respuesta del sistema

<p>1 El caso de uso se inicia cuando el actor después de haber realizado una búsqueda inicial filtrada por los recursos educativos desea restringir más la búsqueda y selecciona una de las opciones presentadas:</p> <ul style="list-style-type: none"> ✓ Autor. ✓ Fecha de envío. ✓ Fecha de modificación. <p>O puede buscar por otro tipo de contenido.</p>	<p>2 Se restringe la búsqueda en los recursos educativos según la opción escogida. O busca en el tipo de contenido seleccionado.</p>
<p>3 Ahora puede buscar en la nueva restricción escogida. (Para mantener los filtros seleccionados debe dar clic en la casilla para mantener los filtros anteriores.)</p>	<p>4 El CU termina.</p>

Tabla 8: Descripción del Caso de Uso Filtrar por clasificación de recurso educativo.

Anexo 3: Diagramas de Clases de Análisis

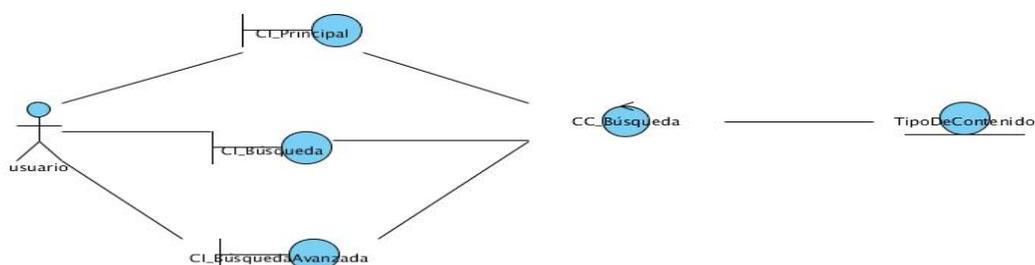


Figura 1: Diagrama de Clases de Análisis del CU Realizar búsqueda avanzada.

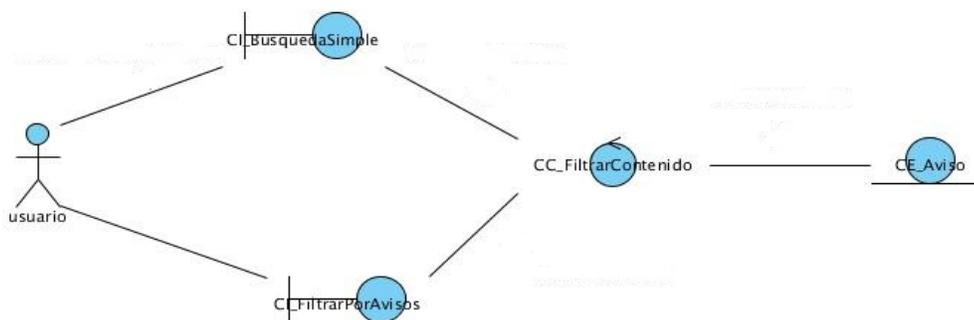


Figura 2: Diagrama de Clases de Análisis del CU Filtrar por avisos.

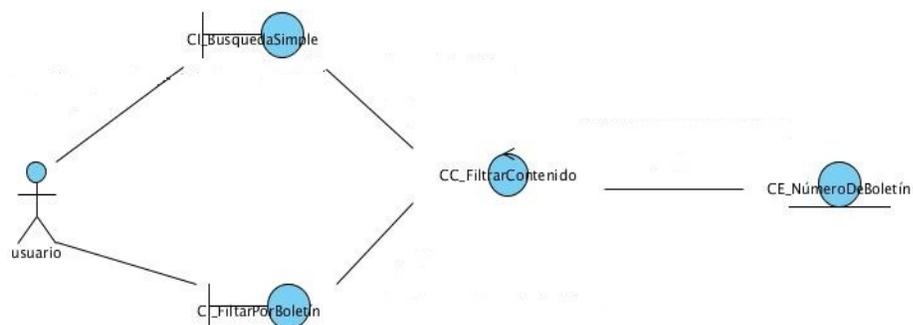


Figura 3: Diagrama de Clases de Análisis del CU Filtrar por boletín.

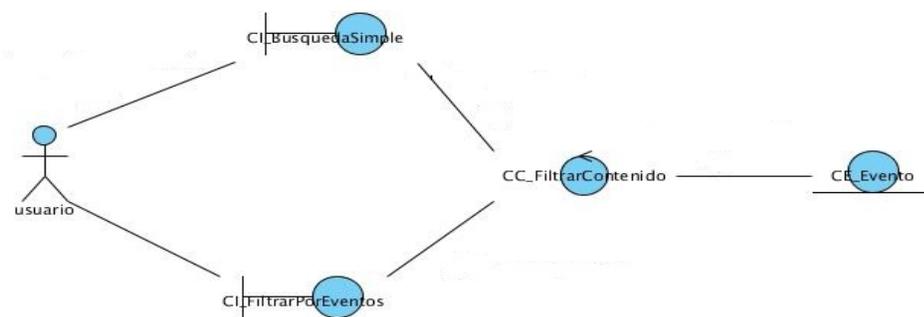


Figura 4: Diagrama de Clases de Análisis del CU Filtrar por eventos.

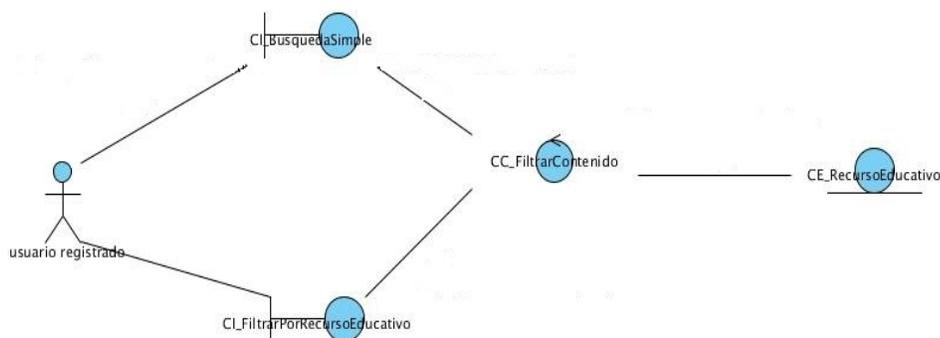


Figura 5: Diagrama de Clases de Análisis del CU Filtrar por recurso educativo.

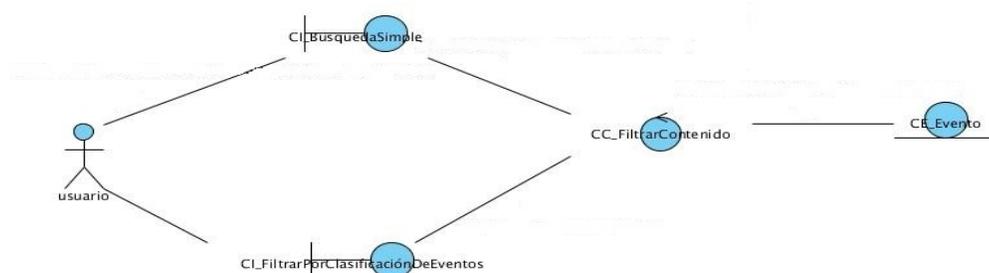


Figura 6: Diagrama de Clases de Análisis del CU Filtrar por clasificación de eventos.

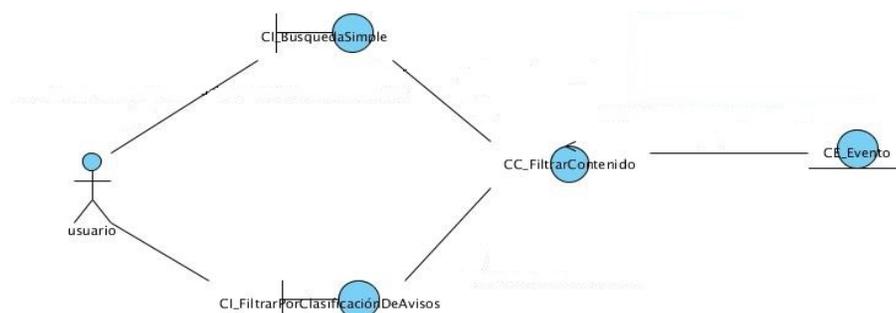


Figura 7: Diagrama de Clases de Análisis del CU Filtrar por clasificación de avisos.

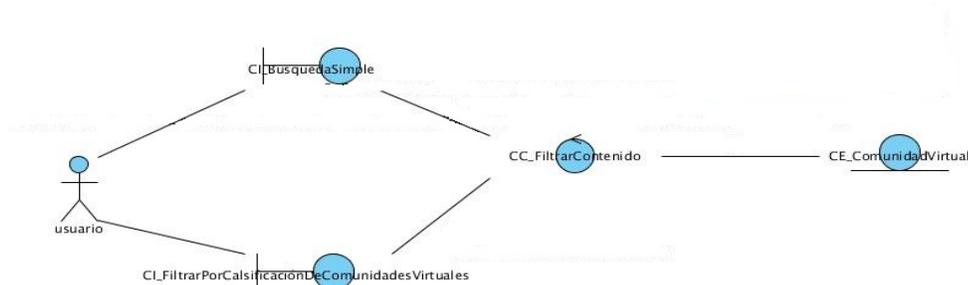


Figura 8: Diagrama de Clases de Análisis del CU Filtrar por clasificación de comunidades virtuales.

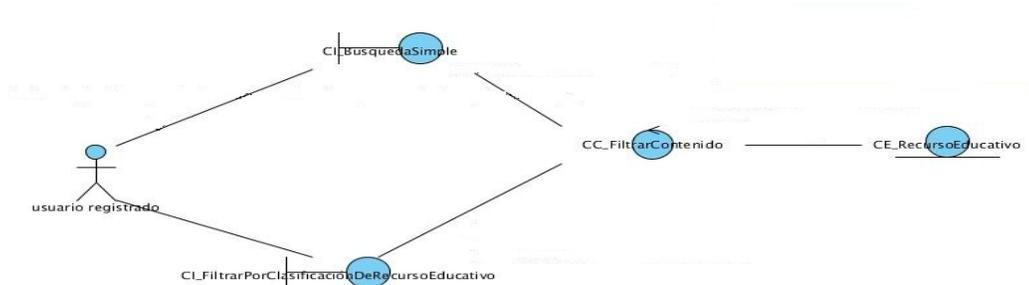


Figura 9: Diagrama de Clases de Análisis del CU Filtrar por clasificación de recurso educativo.

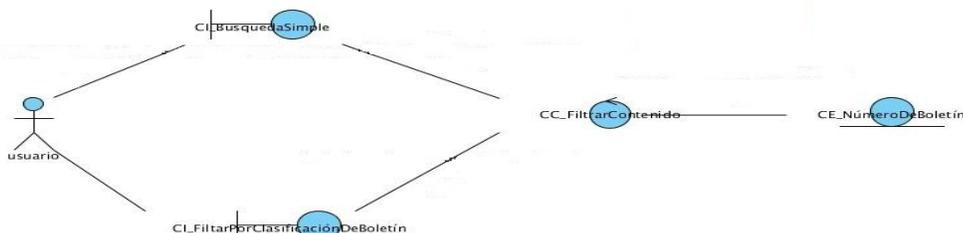


Figura 10: Diagrama de Clases de Análisis del CU Filtrar por clasificación de boletín.

Anexo 4: Diagramas Colaboración

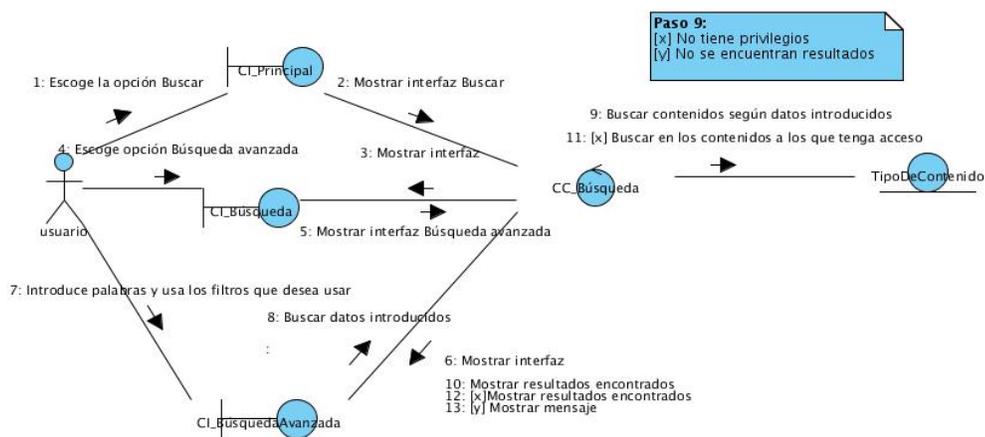


Figura 1: Diagrama de Colaboración del CU Realizar búsqueda avanzada.

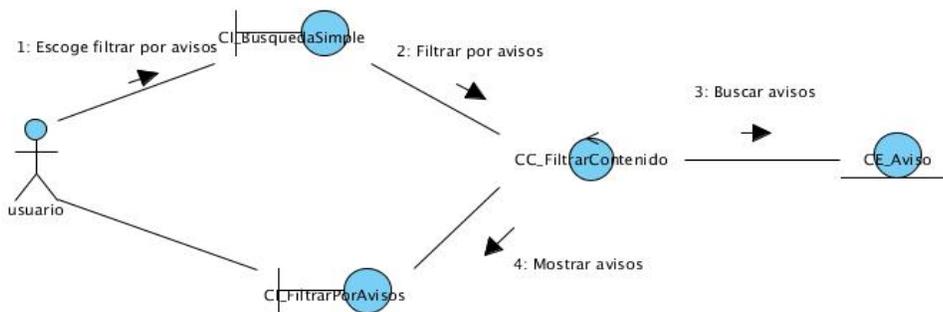


Figura 2: Diagrama de Colaboración del CU Filtrar por avisos.

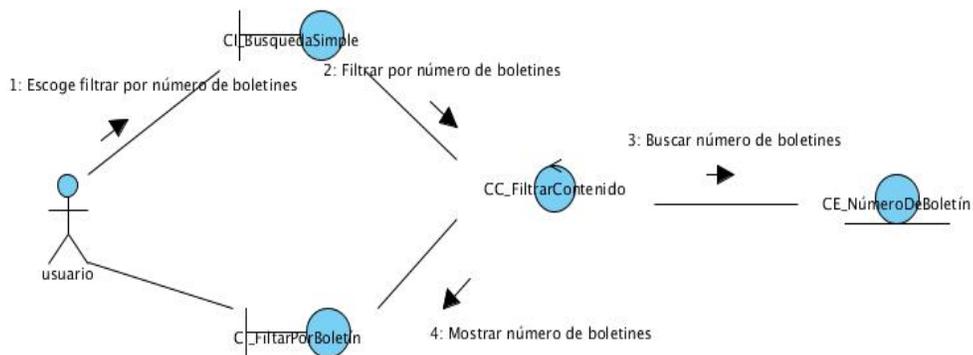


Figura 3: Diagrama de Colaboración del CU Filtrar por boletín.

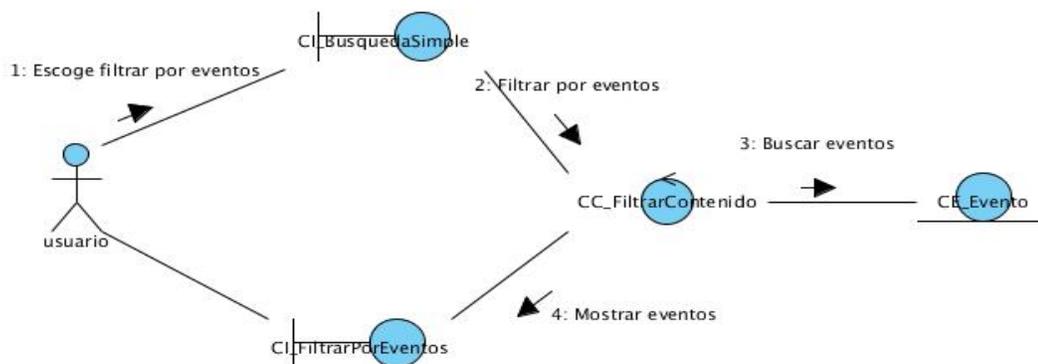


Figura 4: Diagrama de Colaboración del CU Filtrar por eventos.

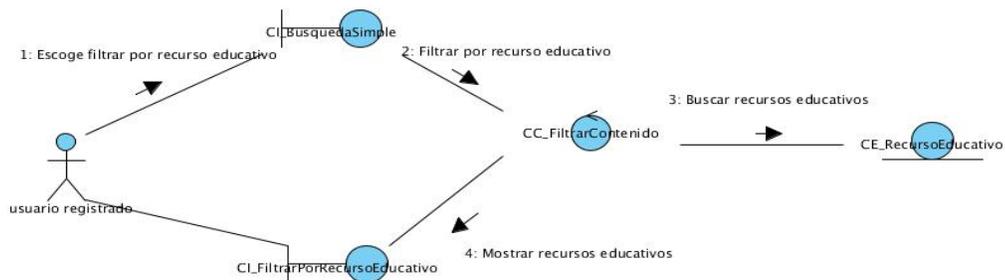


Figura 5: Diagrama de Colaboración del CU Filtrar por recurso educativo.

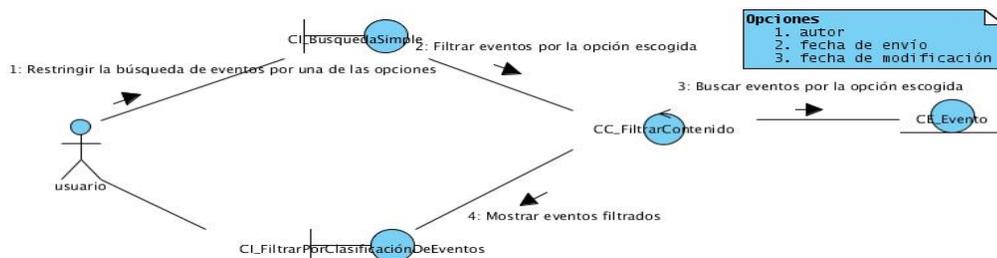


Figura 6: Diagrama de Colaboración del CU Filtrar por clasificación de eventos.

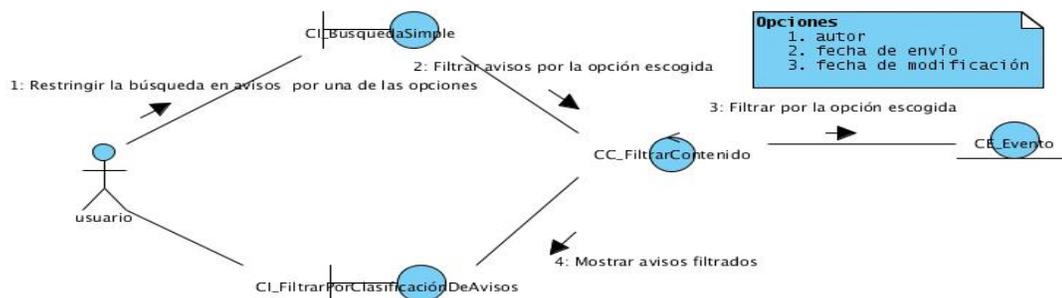


Figura 7: Diagrama de Colaboración del CU Filtrar por clasificación de avisos.

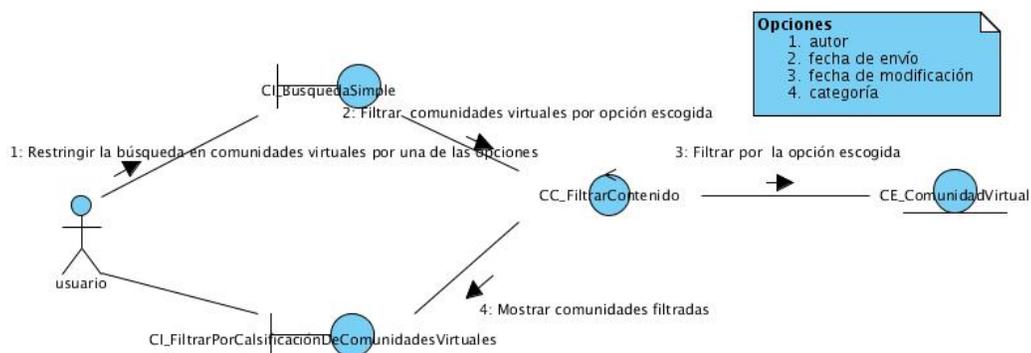


Figura 8: Diagrama de Colaboración del CU Filtrar por clasificación de comunidades virtuales.

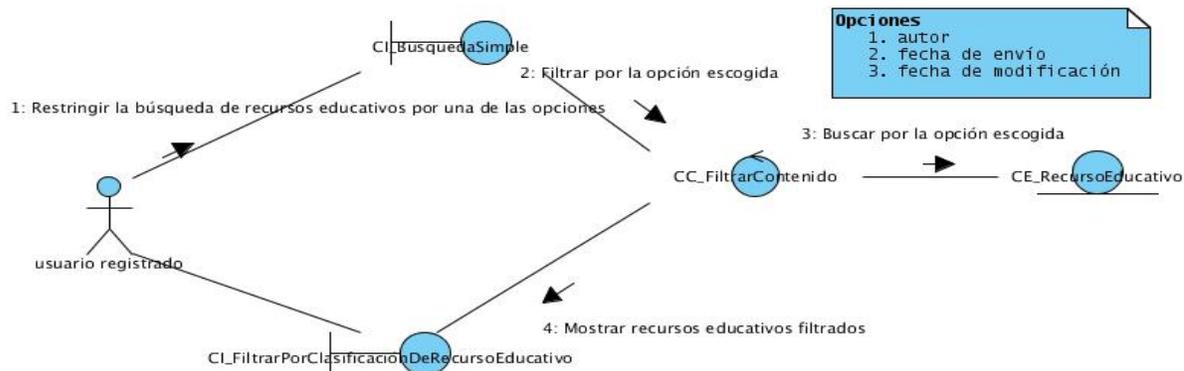


Figura 9: Diagrama de Colaboración del CU Filtrar por clasificación de recurso educativo.

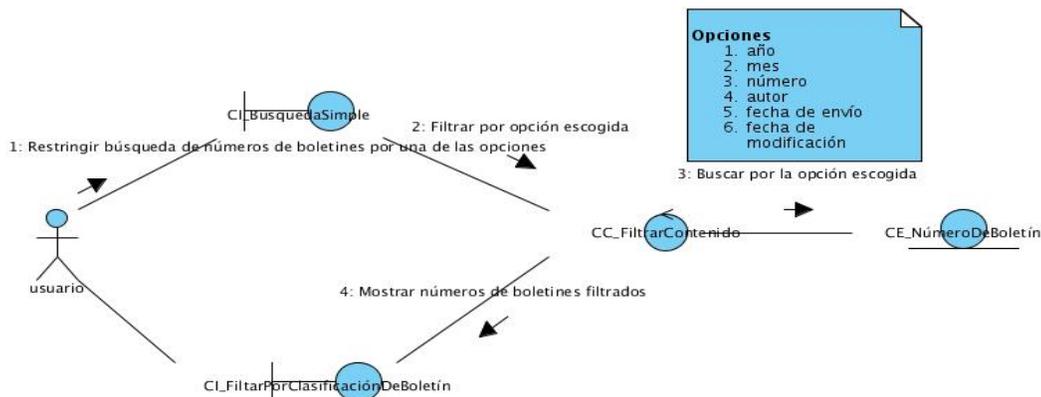


Figura 10: Diagrama de Colaboración del CU Filtrar por clasificación de boletín.

Anexo 5: Diagramas de Clases de Diseño Web

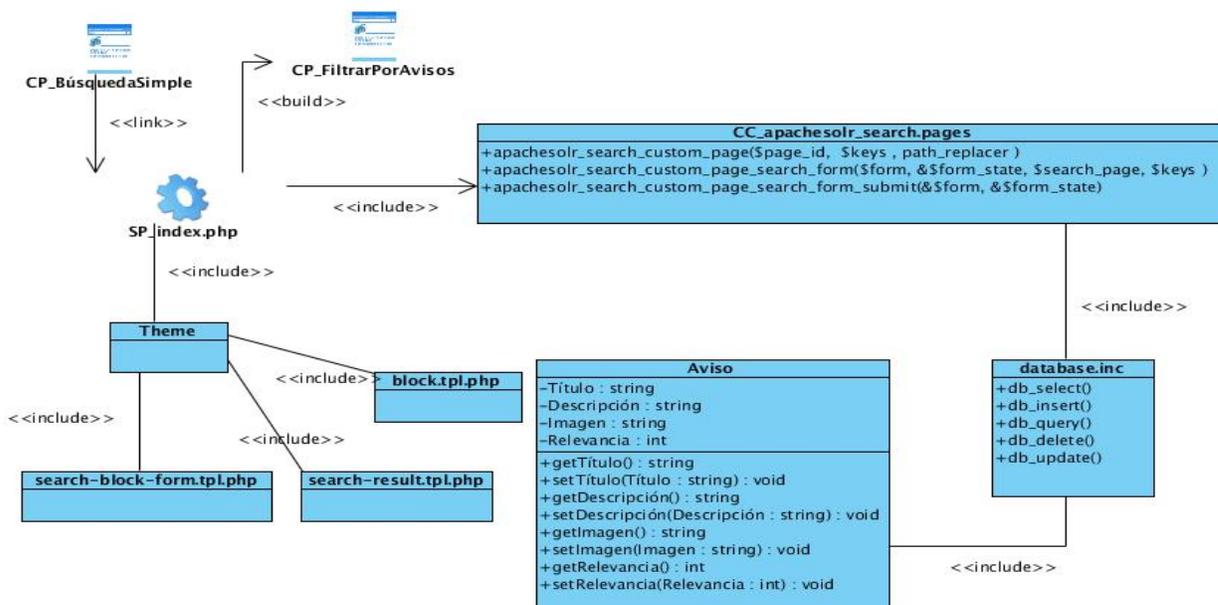


Figura 1: Diagrama de Clases del CU Filtrar por avisos.

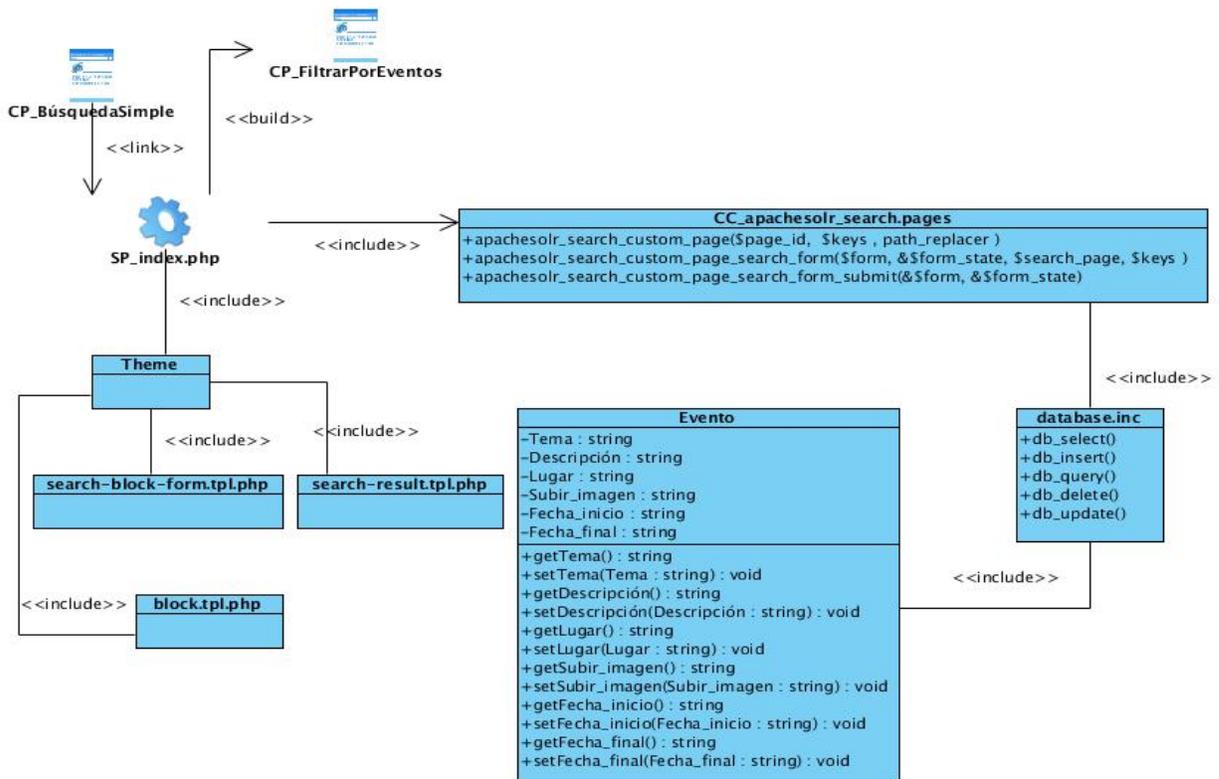


Figura 2: Diagrama de Clases del CU Filtrar por eventos.

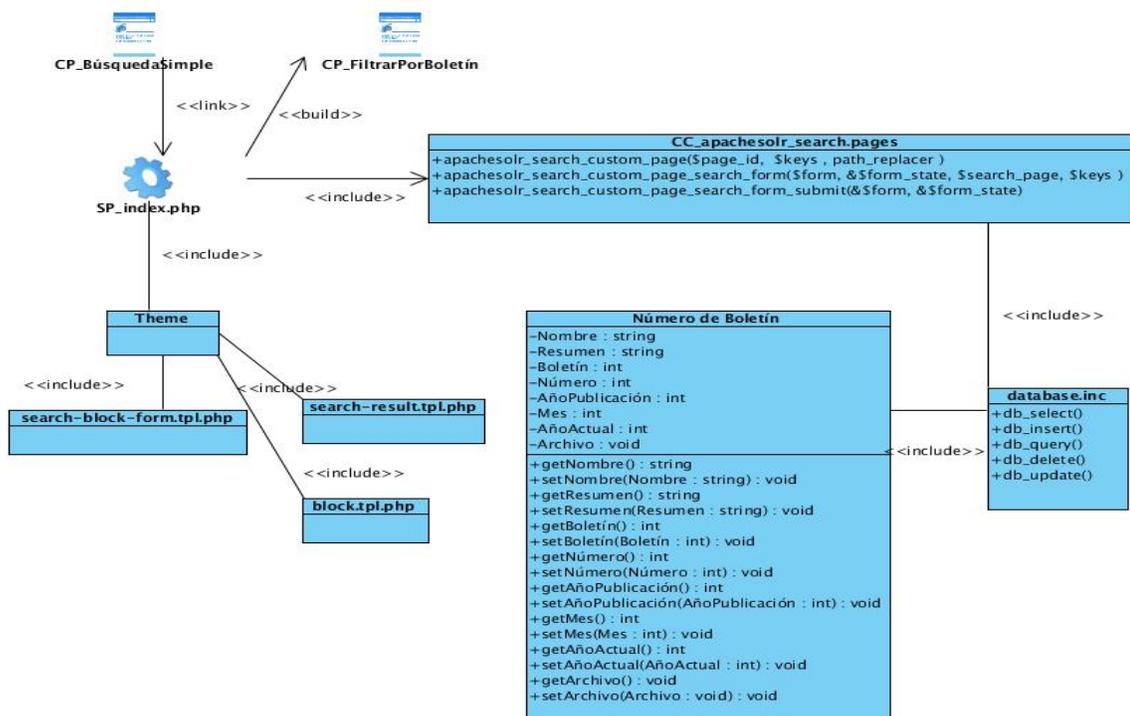


Figura 3: Diagrama de Clases del CU Filtrar por boletín.

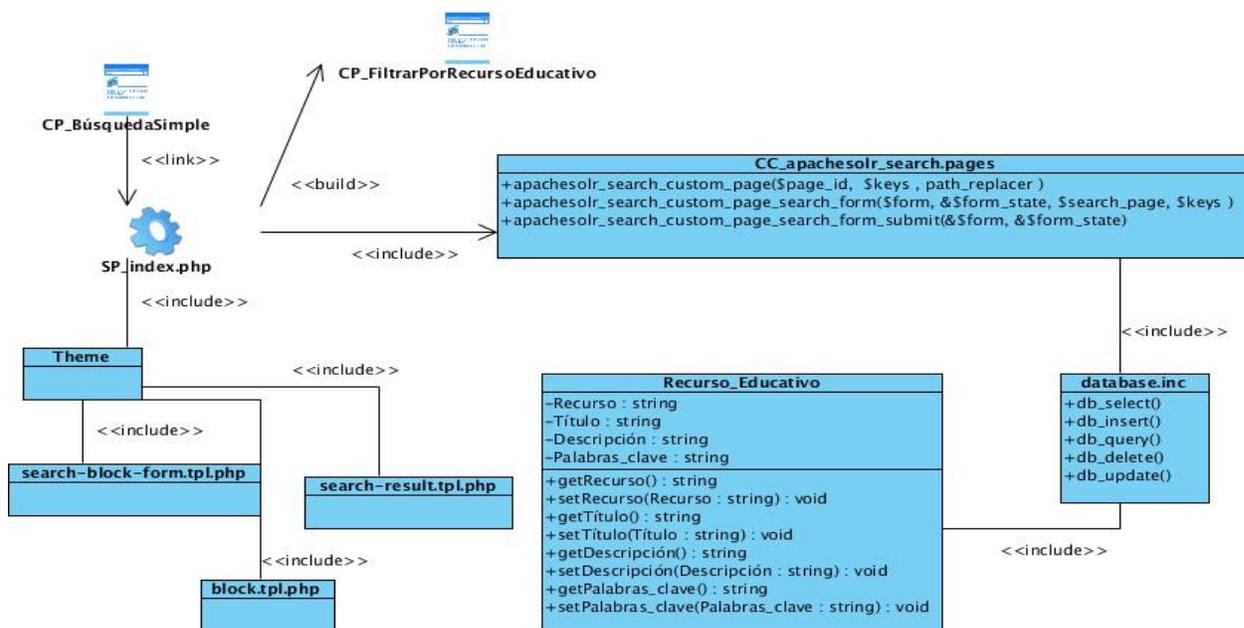


Figura 4: Diagrama de Clases del CU Filtrar por recurso educativo.

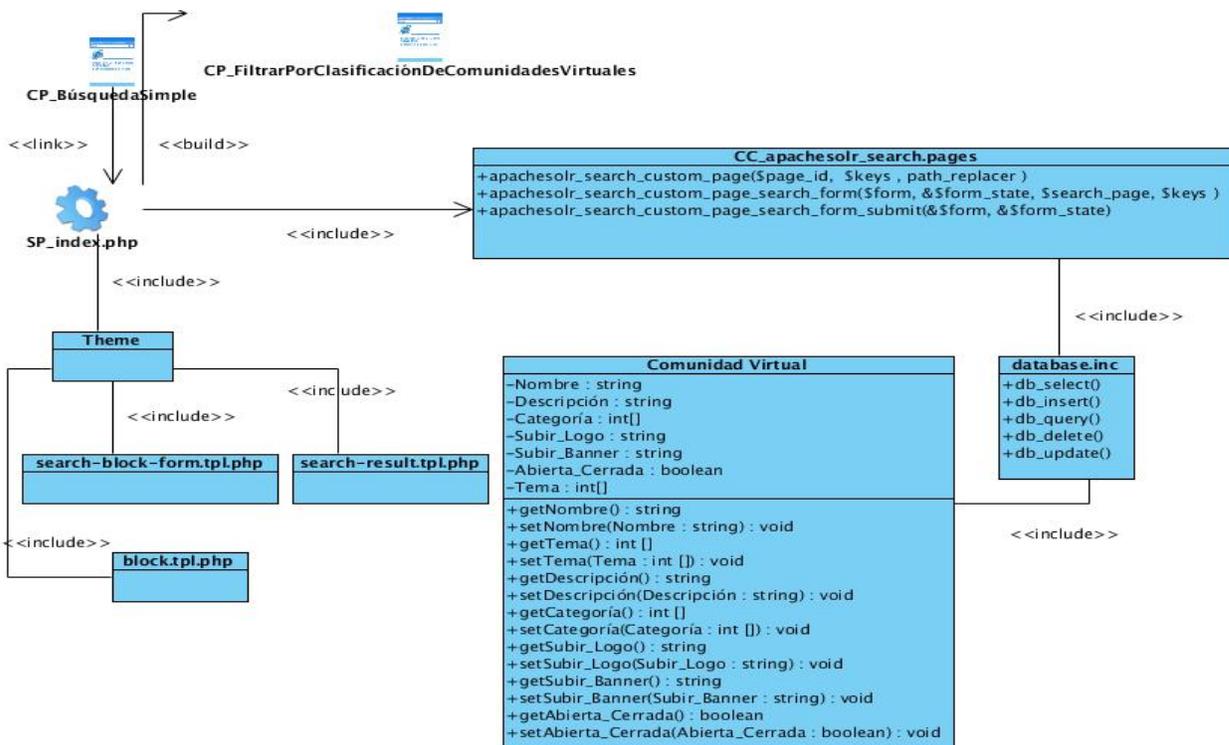


Figura 5: Diagrama de Clases del CU Filtrar por clasificación de comunidades virtuales.

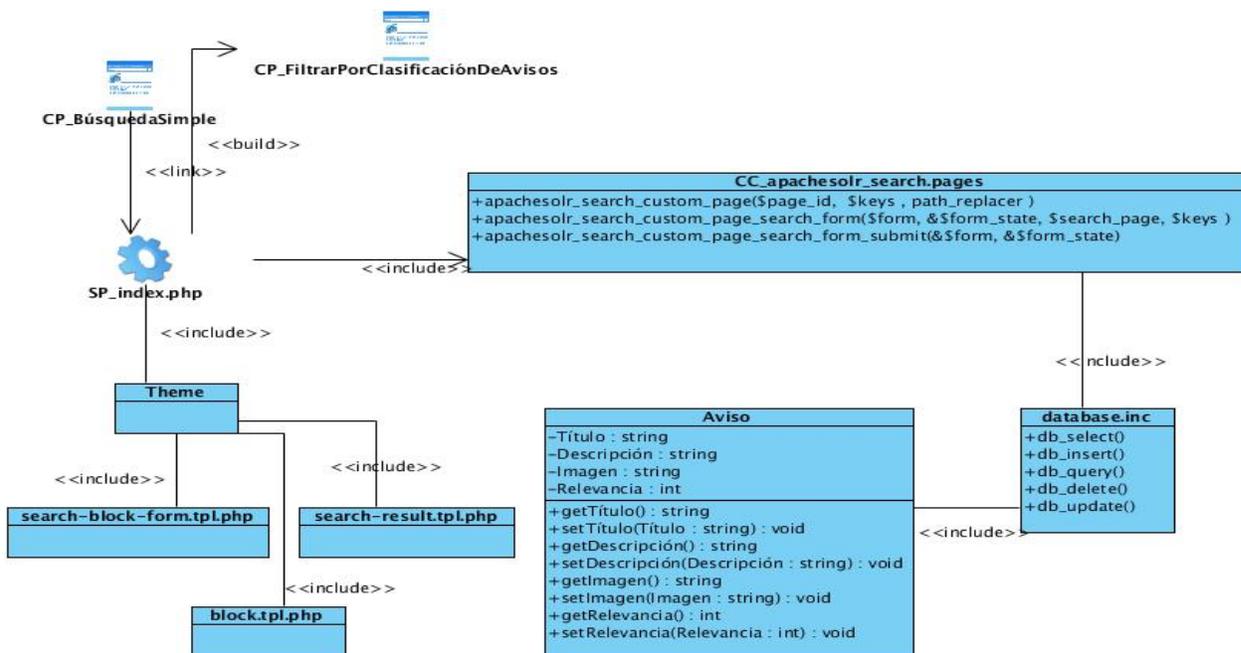


Figura 6: Diagrama de Clases del CU Filtrar por clasificación de avisos.

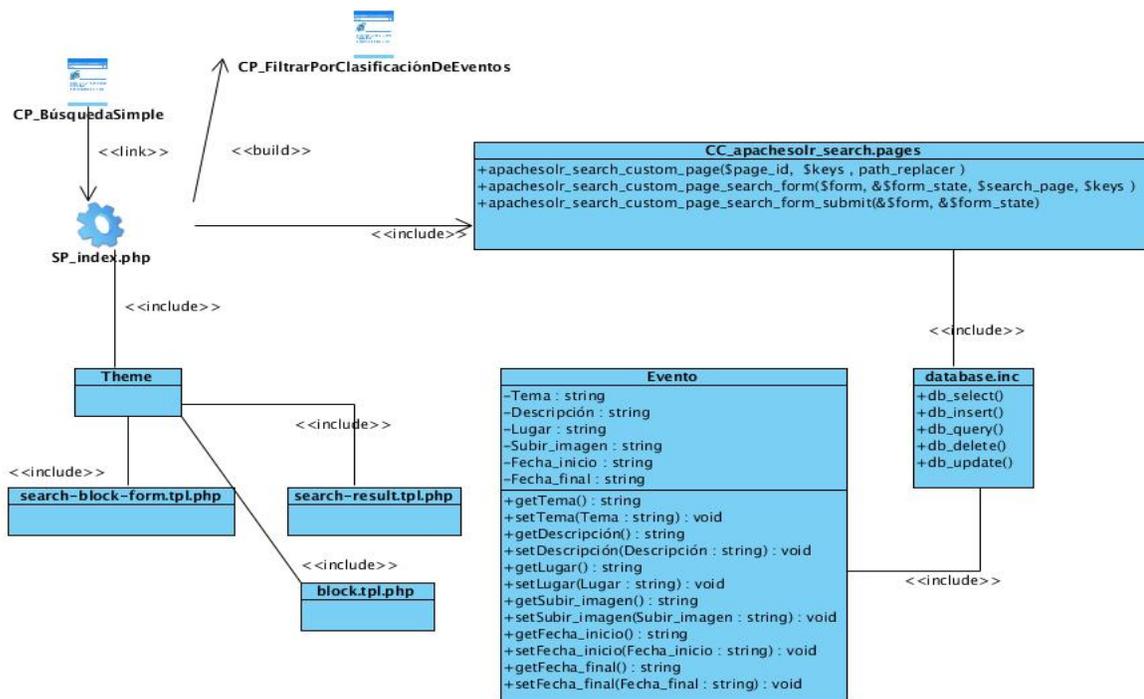


Figura 7: Diagrama de Clases del CU Filtrar por clasificación de eventos.

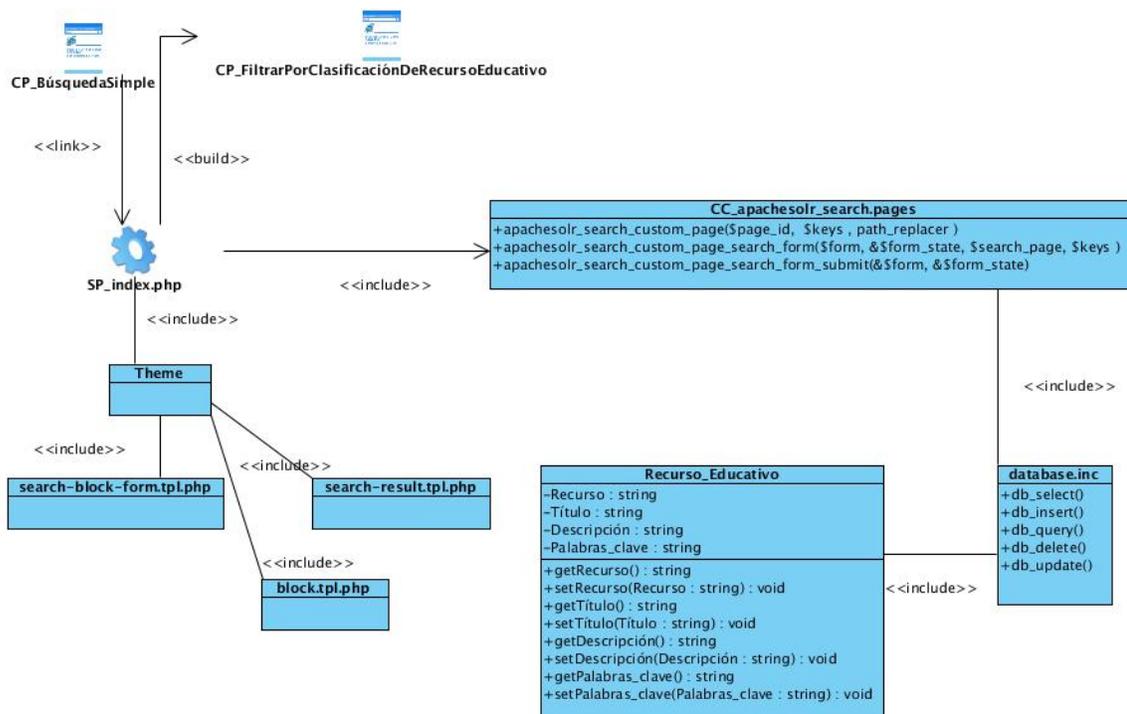


Figura 8: Diagrama de Clases del CU Filtrar por clasificación de recurso educativo.

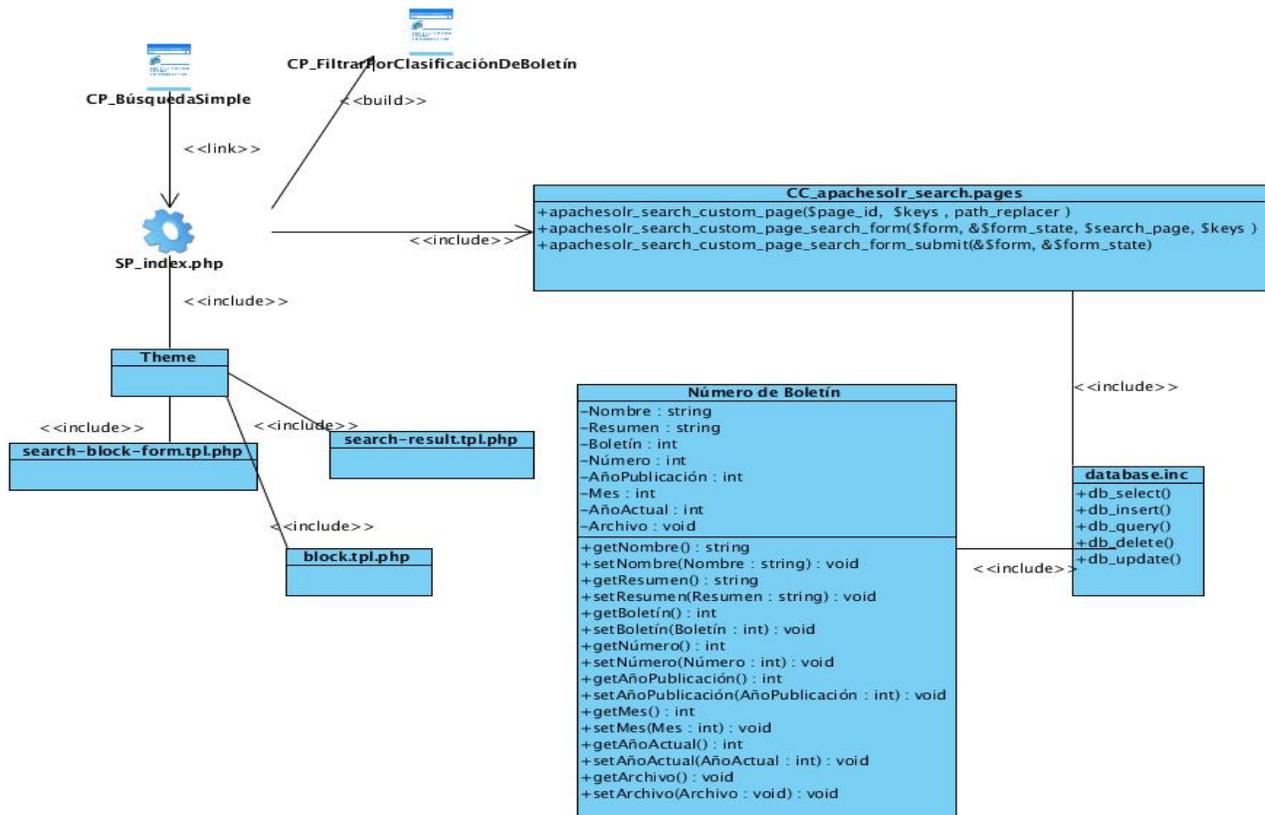


Figura 9: Diagrama de Clases del CU Filtrar por clasificación de boletín.

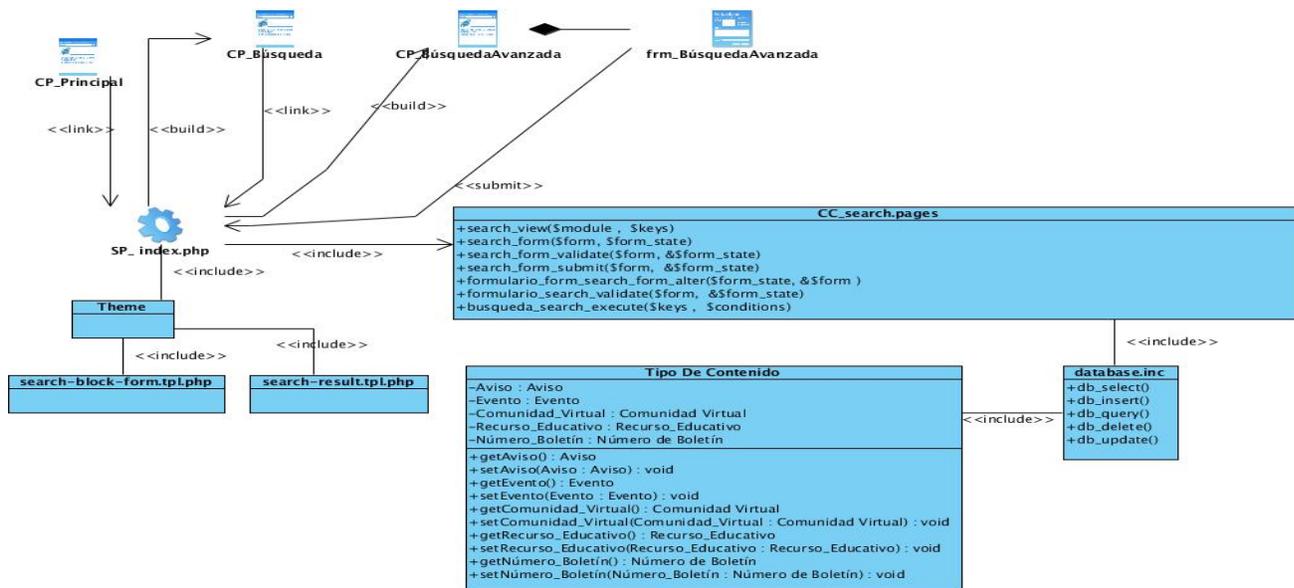


Figura 10: Diagrama de Clases del CU Realizar búsqueda avanzada.

GLOSARIO

- **Bases de Datos:** Un conjunto de información homogénea que mantiene una estructura ordenada y que toda ella se encuentra relacionada.
- **Hook (gancho):** Drupal utiliza estas funciones que permiten la comunicación entre los distintos módulos del sistema.
- **Internet:** Es una red de cómputo a nivel mundial que agrupa a distintos tipos de redes usando un mismo protocolo de comunicación. Los usuarios en Internet pueden compartir datos, recursos y servicios.
- **Índices:** Listado de términos normalizados que representan el contenido de un recurso. Algunos tipos son: índice de materias, índice alfabético, índice KWIC, índice KWOC.
- **Operadores:** En un lenguaje de interrogación, los operadores son los encargados de expresar las relaciones que mantienen entre sí los términos que pueden definir las necesidades informativas del usuario.
- **Palabras clave (Keywords):** Término significativo en lenguaje natural que representa el contenido del documento. En la búsqueda de información esta opción es esencial ya que permite acotar y precisar información.
- **Servlets:** Son objetos que corren dentro y fuera del contexto de un contenedor de *servlets* (ejemplo: Apache Tomcat) y extienden su funcionalidad. La palabra *servlet* deriva de otra anterior, *applet*, que se refería a pequeños programas que se ejecutan en el contexto de un navegador *web*. El uso más común de los *servlets* es generar páginas *web* de forma dinámica a partir de los parámetros de la petición que envíe el navegador *web*.
- **Tesauros:** Es un listado terminológico controlado sobre un área o ámbito de conocimiento que mantiene entre sí relaciones semánticas y genéricas. Su principal característica es que los términos están ordenados jerárquicamente, permitiendo la precisión terminológica en la búsqueda de información.