

Centro de Informatización Universitaria

Facultad 1

Universidad de las Ciencias Informáticas



**Herramienta para la migración de datos desde el Sistema Integrado de Gestión Bibliotecaria hacia el Repositorio institucional de la Universidad de las Ciencias Informáticas**

Trabajo de diploma para optar por el título de  
Ingeniero en Ciencias Informáticas

**Autor:**

Adrian Maren Dinza

**Tutores:**

Ing. Raudel González Echenique

Ing. Luis Carlos Álvarez Fernández

La Habana, junio de 2012

“Año 54 de la Revolución”

### **Declaración de autoría**

Declaro ser el autor del presente trabajo de diploma y reconozco a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo.

Para que así conste firmo la presente a los días \_\_\_\_ del mes \_\_\_\_ del año \_\_\_\_.

\_\_\_\_\_  
Adrian Maren Dinza

\_\_\_\_\_  
Ing. Raudel González Echenique

\_\_\_\_\_  
Ing. Luis Carlos Álvarez Fernández

---

*" Muchos me dirán aventurero, y lo soy, solo que de un tipo diferente y de los que ponen el pellejo para demostrar sus verdades. "*

*Che.*



### **Dedicatoria**

A mi madre por haberme traído al mundo, a mis hermanos Adriel y Asiel que siempre los tengo conmigo en el corazón.

A mi tía Ali por haber sabido ser como una madre, a la primita que más quiero Laura, a mi abuelo y a mi tío Ale.

A una persona en especial que ya no se encuentra entre nosotros a mi mami Pura.

## **Agradecimientos**

Agradezco a la Universidad de las Ciencias Informáticas, por darme la oportunidad de formarme en ella durante los últimos cinco años y hacer de mis sueños una realidad.

A mis padres, a mis tíos, a mi novia y a su familia que supo acogerme como si fuera una más de ella.

A todos los amigos que hice durante el trascurso por la universidad.

A todos los que de una forma u otra, me han ayudado durante la realización del trabajo: amigos, compañeros, en fin gracias a todos.

### Resumen

La gestión de la información científico-técnica y académica de las instituciones es uno de los desafíos que enfrenta la sociedad actual. La Universidad de las Ciencias Informáticas cuenta con un Sistema Integrado de Gestión Bibliotecaria que se encarga de la adquisición, conservación y divulgación de un gran volumen de obras y bibliografías que responden a una amplia gama de géneros. Además, dicha institución cuenta con un Repositorio institucional para la gestión y concentración en un mismo sitio de la información científico-técnica generada en la universidad. Sin embargo el Sistema Integrado de Gestión Bibliotecaria almacena sus datos en formato Marc21 y a su vez el Repositorio institucional en Dublin Core.

El objetivo del presente trabajo es desarrollar una herramienta que facilite el proceso de migración de datos desde el formato Marc21 hacia Dublin Core, para lograr que parte de la información científico-técnica que se catalogue en el Sistema Integrado de Gestión Bibliotecaria pueda ser importada hacia el Repositorio institucional sin necesidad de volver a pasar por un proceso de catalogación.

**Palabras clave:** dublin core, migración de datos, repositorio institucional, sistema integrado de gestión bibliotecaria.

**Índice**

Introducción .....1

Capítulo 1: Fundamentación teórica .....4

1.1 Sistema Integrado de Gestión Bibliotecaria.....4

1.2 Repositorio institucional .....5

1.3 Migración de datos.....5

1.3.1 Procesos de migración de datos .....6

1.4 MARC21 .....7

1.5 Dublin Core .....8

1.6 Herramientas que implementan la migración de datos.....11

1.7 Estructura de carpeta para importar en el Repositorio institucional .....13

1.8 Herramientas de desarrollo .....14

1.8.1 Gestor de base de datos .....14

1.8.2 Servidor web.....14

1.9 Metodología de desarrollo .....15

1.10 Lenguajes a utilizar.....16

1.10.1 Perl.....16

1.10.2 HTML.....16

1.10.3 CSS.....16

1.11 Herramientas CASE.....17

1.12 Herramientas de automatización del proceso de desarrollo .....17

1.12.1 Eclipse .....17

1.12.2 EPIC .....	18
1.12.3 Control de versiones .....	18
1.13 Conclusiones parciales .....	18
Capítulo 2: Descripción de la solución .....	19
2.1 Propuesta del sistema .....	19
2.2 Modelo de dominio.....	19
2.3 Requerimientos .....	20
2.3.1 Requerimientos funcionales.....	20
2.3.2 Requerimientos no funcionales .....	21
2.4 Modelo de casos de uso del sistema.....	24
2.4.1 Especificación de casos de uso .....	25
2.5 Análisis.....	27
2.5.1 Diagramas de clases del análisis .....	28
2.6 Diseño .....	29
2.6.1 Descripción de la arquitectura .....	29
2.6.2 Modelo-Vista-Controlador (MVC) .....	30
2.6.3 Diagramas de clases del diseño.....	31
2.6.4 Diagramas de colaboración .....	32
2.6.5 Modelo entidad-relación .....	34
2.7 Conclusiones parciales .....	36
Capítulo 3: Implementación y prueba .....	37
3.1 Diagrama de componentes .....	37



3.2 Diagrama de despliegue .....	39
3.3 Pruebas .....	40
3.3.1 Modelos de pruebas .....	40
3.3.2 Pruebas de caja negra .....	40
3.3.3 Resultados de las pruebas .....	43
3.4 Conclusiones parciales .....	43
Conclusiones generales.....	44
Recomendaciones .....	45
Referencias bibliográficas.....	46
Bibliografía consultada .....	48
Glosario de términos.....	50
Anexos.....	52

**Índice de figuras**

Figura 1 Diagrama de procesos de la migración de datos.....6

Figura 2 Ejemplo de estructura de carpeta para importación.....13

Figura 3 Diagrama del modelo de dominio.....20

Figura 4 Diagrama de casos de uso.....25

Figura 5 Diagrama de clases del análisis del CU Exportar registros .....28

Figura 6 Diagrama de clases del análisis del CU Generar estructura de carpeta .....29

Figura 7 Funcionamiento del MVC .....30

Figura 8 Diagrama de clases CU Exportar registros.....31

Figura 9 Diagrama de clases CU Generar estructura de carpeta.....32

Figura 10 Diagrama de colaboración CU Exportar registros .....33

Figura 11 Diagrama de colaboración CU Generar estructura de carpeta .....34

Figura 12 Modelo entidad-relación .....35

Figura 13 Diagrama de componentes .....38

Figura 14 Diagrama de despliegue.....39

Figura 15 Gráfica de las iteraciones de los casos de pruebas.....43

## Índice de tabla

Tabla 1 Especificación de requerimientos .....	21
Tabla 2 Actor del sistema .....	24
Tabla 3 Especificación del CU Exportar registros .....	25
Tabla 4 Especificación del CU Generar estructura de carpeta .....	26
Tabla 5 Caso de prueba Exportar registros.....	41
Tabla 6 Caso de prueba Generar estructura de carpeta.....	42

## Introducción

Uno de los mayores retos que enfrentan en la actualidad las instituciones científicas y académicas es el almacenamiento de la información científica generada por sus investigadores. Universidades, instituciones científicas y centros de desarrollo constituyen una fuente inagotable de conocimientos.

En la Universidad de las Ciencias Informáticas (UCI), universidad de nuevo tipo que tiene en su modelo de formación como procesos sustantivos la docencia, la producción y la investigación, el almacenamiento de la bibliografía digital corre a cargo de la biblioteca de la institución.

La biblioteca cuenta con un Sistema Integrado de Gestión Bibliotecaria (SIGB), a través del cual se manejan de forma digital los procesos que se desarrollan en la misma: adquisición, circulación, catalogación y manejo de los usuarios. El SIGB cuenta con un catálogo en línea en el cual se pueden realizar búsquedas sobre el material bibliográfico con que cuenta la institución, tanto en formato duro (aunque en muchos casos cuentan con una versión digital) como en formato digital.

En el año 2010, en un trabajo de diploma realizado por estudiantes de la Facultad 10, se dan los primeros pasos para la implantación de un Repositorio institucional en la UCI, convirtiéndose en una realidad en el año 2011. El Repositorio institucional tiene como objetivo almacenar en formato digital la información científico-técnica generada por los investigadores (estudiantes y profesores) de la universidad. En febrero de 2012 el Repositorio contaba con 5492 archivos almacenados, principalmente trabajos de diploma, tesis de maestría, doctorado y artículos científicos presentados en eventos de la universidad.

Parte de la documentación manejada por la biblioteca, al constituir investigaciones de personal de la universidad y contar con versiones en formato duro y digitales se encuentran accesibles desde el catálogo en línea del SIGB y en el Repositorio institucional. Este último es el que almacena el archivo en formato digital.

Una vez que los documentos llegan a la biblioteca son catalogados por el personal especializado de la biblioteca. Los metadatos del documento se almacenan en formato Marc21, y pasan a estar disponibles en el catálogo en línea. Sin embargo, el proceso de catalogación se debe repetir al insertar el documento digital en el Repositorio institucional, el cual utiliza el Dublin Core como formato de descripción de metadatos.

Solamente en el año 2011 se procesaron más de 800 documentos científicos generados en la universidad,

para lo cual fue necesario catalogar los mismos en el SIGB y en el Repositorio institucional.

De la situación antes expuesta se origina el siguiente **problema científico**: ¿Cómo evitar el doble procesamiento de la documentación científica generada en la Universidad de las Ciencias Informáticas en los sistemas de gestión de la biblioteca?

Se plantea como el **objeto de estudio** el proceso de migración de datos bibliográficos, de donde se deriva el **campo de acción** como la migración de datos desde el Sistema Integrado de Gestión Bibliotecaria hacia los Repositorios institucionales.

Como **idea a defender** se parte de la siguiente premisa: la implementación de una herramienta para la migración de datos permitirá lograr la migración de los registros bibliográficos que se encuentran en el Sistema Integrado de Gestión Bibliotecaria hacia el Repositorio institucional de la Universidad de las Ciencias Informáticas.

De esta manera, se define como **objetivo general**, desarrollar una herramienta que permita la migración de registros bibliográficos desde el Sistema Integrado de Gestión Bibliotecaria hacia el Repositorio Institucional para evitar el doble procesamiento de la documentación científica generada en la Universidad de las Ciencias Informáticas.

Para lograr el objetivo general se proponen las siguientes **tareas**:

- Revisión y análisis bibliográfico del formato de descripción bibliográfica MARC21.
- Revisión y análisis bibliográfico del formato de descripción bibliográfica Dublin Core.
- Análisis de las tecnologías actuales a utilizar para el desarrollo de la migración de datos.
- Implementación de una herramienta para la migración de datos de MARC21 a Dublin Core.
- Ejecución de pruebas funcionales a través del método de caja negra a la herramienta desarrollada.

Como **métodos teóricos** se utilizan:

- Histórico-lógico: En el trabajo de diploma este método se usa con el objetivo de estudiar lo referente a los sistemas que implementan la migración de datos, ya sea tanto el ámbito nacional como internacional y específicamente en la Universidad de las Ciencias Informáticas, para obtener mayor conocimiento sobre el tema. Este método ayuda a dar cumplimiento al estudio del estado del arte relacionado con la migración de datos.

Como **método empírico**:

- Análisis de documentos, se basa en la revisión de documentos utilizados durante la investigación.

Este trabajo se estructura en 3 capítulos, en los cuales se expresan los elementos de la investigación, así como las características y requisitos a tener en cuenta en el desarrollo de la herramienta para la migración de datos y las pruebas a realizar a la propia herramienta.

**Capítulo 1:** Fundamentación teórica. En este capítulo se presentan los conceptos fundamentales para la comprensión de lo referente a las principales características de las herramientas de migración. Además, se describen las principales herramientas y tecnologías a utilizar para el desarrollo de la misma.

**Capítulo 2:** Descripción de la solución. En este capítulo se describen los objetivos fundamentales de la organización y el flujo actual de los procesos involucrados en el campo de acción. Abarca los elementos relacionados con la informatización de la cual serán objeto dichos procesos. Se describe de forma general la propuesta del sistema y cómo debe funcionar, además de los diferentes requisitos con los que debe contar. Además, describe la propuesta del análisis y diseño de la herramienta.

**Capítulo 3:** Implementación y prueba. En este capítulo se detalla el modelo de despliegue de la herramienta y los componentes que lo conforman. Además, se describen las pruebas de unidades realizadas que permitan validar la solución y se exponen ejemplos de pruebas ejecutadas a la herramienta que demuestran la integridad y aceptabilidad de la misma.

## Capítulo 1: Fundamentación teórica

En este capítulo se realiza un estudio sobre alguna de las herramientas usadas para la migración de datos existentes en el mundo. Se presentan y analizan los conceptos esenciales relacionados con el tema y cómo se lleva a cabo la migración de registros bibliográficos, para así definir una estrategia a la hora de desarrollar una migración de datos.

### 1.1 Sistema Integrado de Gestión Bibliotecaria

Es un conjunto organizado de recursos humanos que utilizan dispositivos y programas informáticos, adecuados a la naturaleza de los datos que deben procesar, para realizar procesos y facilitar los servicios que permiten alcanzar los objetivos de una institución bibliotecaria. Permite almacenar de forma organizada el conocimiento humano contenido en todo tipo de materiales bibliográficos para satisfacer las necesidades informativas, recreativas y de investigación de los usuarios [1].

Permite informatizar las operaciones bibliotecarias más comunes e integra en un único sistema todos los procesos de gestión bibliotecaria. Típicamente abarca la catalogación, circulación, reportes, gestión de usuarios, consulta y adquisición de materiales. Aunque la tendencia de los desarrolladores de SIGB en los últimos años ha consistido en ampliar su oferta con nuevos productos o módulos más o menos integrados en sus sistemas de automatización como son sistemas de resolución de enlaces, meta-buscadores, gestores de recursos electrónicos, nuevas interfaces o portales web de consulta.

El SIGB presenta las siguientes características:

- Es un sistema de código abierto (Open Source) distribuido bajo una licencia GPL de uso que permite disponer del código fuente.
- Un moderno y completo Sistema de Información Bibliográfico (SIB) que incluye todas las funciones de gestión de una biblioteca dentro de un mismo sistema.
- Se integra fácilmente con otras aplicaciones a través de servicios web.
- Diseñado con una arquitectura abierta de tres capas.
- Posibilidad de integración en portales institucionales.
- Posibilidad de utilizar gestor documental o biblioteca virtual.
- Multibiblioteca que permite gestionar las funciones de varias sedes con políticas centralizadas o descentralizadas.

- No tiene límite en cuanto al número de registros que se puedan manejar dentro del sistema.
- Se adapta a todo tipo de bibliotecas.
- Maneja la estructura de tesoro/autoridades.
- Protocolo Z39.50 para transferencia de información.

## 1.2 Repositorio institucional

Un Repositorio institucional consiste en un conjunto de servicios prestados por la universidad a los miembros de su comunidad. Para recopilar, gestionar, difundir y preservar la producción documental generada en la institución, cualquiera que sea su tipología a través de la creación de una colección digital organizada, abierta e interoperable para garantizar visibilidad e impacto [2].

Es esencial un compromiso organizativo para la administración de estos materiales digitales, incluyendo la preservación a largo plazo cuando sea necesario, así como la organización y acceso o su distribución [3].

Entre los beneficios que aportan los Repositorios institucionales se encuentran:

- Permiten a la institución elevar la visibilidad de sus investigaciones al ofrecer acceso a las obras de sus miembros.
- Las organizaciones científicas y académicas de todo el mundo obtienen ventajas de un acceso más fácil y rápido a las investigaciones desarrolladas en otras instituciones.
- La inclusión de un trabajo científico aumenta la proyección del perfil del autor a escala mundial y el impacto de la investigación realizada.
- Permiten realizar estudios de producción científica e impacto de la investigación.
- Favorecen la preservación de las investigaciones producidas.

## 1.3 Migración de datos

Muchas circunstancias existen cuando una organización necesita migrar las aplicaciones o las bases de datos. Puede ser que sea tan simple como una mejora a una nueva versión del sistema. O puede ser que implique cambiar a una nueva base de datos o aplicación. Después de una fusión o de una adquisición, a menudo se retiran las aplicaciones redundantes, pero los datos tienen que ser preservados en el sistema de supervivencia [4].

La migración de datos tiene su fundamento en la ampliación de un sistema de gestión de base. En este



contexto, se tratan de exportar los datos a un nuevo sistema con mayor capacidad o más funciones adicionales. Estos cambios llevan consigo una adaptación de todos los datos de una base de datos a otra. Por tanto, siempre que se producen cambios de un sistema de gestión a otro, se habla inevitablemente de los procesos de migración de datos.

Con frecuencia las operaciones de exportación de datos requieren de un tratamiento especial, resultando muy costosas de realizar. Esto sucede cuando primeramente el formato de exportación de datos no coincide con el de importación de la nueva aplicación. Es típico encontrarse con una incompatibilidad de los formatos de salida y entrada entre dos aplicaciones, en algo tan básico como el formato. El segundo problema es la codificación de los caracteres. En muchos casos se tratan de importar datos codificados con normas ISO 8859-1, cuando las tablas de la base de datos de destino están configuradas para formatos UTF-8 UNICODE. Dicho de otra forma, es necesario que los datos tengan características de codificación para que puedan ser identificados correctamente y evitar una transformación inadecuada de los caracteres especiales. El tercer problema es cuando la extensión del formato de archivo está bien definida, pero no lo está la codificación interna del formato. Dicho de otra forma, la información contenida en el archivo de exportación, no está correctamente definida o delimitada. Esto provoca que el sistema de destino al que se pretende exportar todos los datos, no diferencie correctamente qué información corresponde a qué campo descriptivo o etiqueta.

### 1.3.1 Procesos de migración de datos

La migración de datos está regida por varios procesos los cuales son [5]:

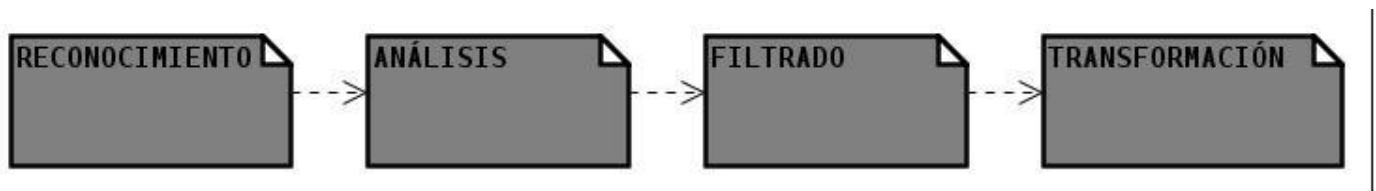


Figura 1 Diagrama de procesos de la migración de datos

1. Reconocimiento: proceso en el cual se identifica la estructura de los datos en la fuente de datos.
2. Análisis: proceso donde se define una estructura para almacenar de forma óptima los datos de la fuente, de modo que puedan realizarse los siguientes procesos de migración, usualmente este proceso se mezcla con el de reconocimiento.

3. Filtrado: proceso opcional que se realiza solamente si se quiere aplicar un filtro a los datos, por ejemplo, reducir su contenido tomando en cuenta cierto factor.
4. Transformación: en este proceso se realizan las transformaciones requeridas para que los datos tomen el formato deseado.

## 1.4 MARC21

MARC21 es la edición actual del Formato MARC. Se trata de una norma para la representación y comunicación electrónica de la información bibliográfica y relacionada [6].

Esta información bibliográfica puede ser: títulos, nombres, temas, notas, información sobre publicación. Además, contiene elementos de datos para los siguientes tipos de materiales:

- Libros.

Ítem de texto de tipo monográfico, tales como libros encuadernados, electrónicos o micro-formas.

- Seriadas.

Ítem de texto con un patrón recurrente de publicación, publicaciones periódicas, diarios, anuarios.

- Archivos de computación.

Software, datos numéricos, multimedia en computación, sistemas o servicios en línea. Otros tipos de recursos electrónicos se codifican según su aspecto más significativo, tales como textos ("libros" o "seriadas"), cartográfico ("mapas").

- Mapas.

Todo tipo de materiales cartográficos, incluyendo mapas en hojas y globos en forma impresa, manuscritos, electrónicos y micro-formas.

- Música.

Partituras musicales impresas y manuscritas.

- Registros sonoros.

Registros sonoros no musicales y registros sonoros musicales.

- Material visual.

Imágenes y objetos, medios que se proyectan, películas, gráficos de dos dimensiones, objetos tridimensionales, objetos reales.

- Materiales mixtos.

Fundamentalmente colecciones de archivo y manuscritos con material de formas mixtas.

Los campos de datos se agrupan en bloques de acuerdo con el primer carácter de la etiqueta, la función de los datos dentro del registro está determinada generalmente por el bloque al que pertenece [7].

0XX Información de control, números de identificación y clasificación.

1XX Asientos principales.

2XX Títulos y párrafo del título (título, edición, pie de imprenta).

3XX Descripción física.

4XX Menciones de serie.

5XX Notas.

6XX Campos de acceso temático.

7XX Asientos secundarios diferentes de los de materias y series; campos ligados.

8XX Asientos secundarios de series, existencias, etc.

9XX Reservados para implementación local.

También se le dan significado a los campos a partir de los dos últimos caracteres de la etiqueta, aunque existen algunas excepciones.

X00 Nombres personales.

X10 Nombres corporativos.

X11 Nombres de reuniones.

X30 Títulos uniformes.

X40 Títulos bibliográficos.

X50 Términos temáticos.

X51 Nombres geográficos.

## 1.5 Dublin Core

El formato Dublin Core es, juntamente con el formato MARC, el formato de metadatos más ampliamente utilizado en el desarrollo de Repositorios de objetos digitales. La iniciativa tenía como objetivo principal, a mediados de los años noventa, elaborar una propuesta de etiquetas para el etiquetado de recursos digitales en la web. El resultado fue la elaboración de una propuesta de 15 etiquetas (ISO 15836-2003) las cuales fueron pronto complementadas con calificadores, que permitían añadir más detalle y precisión a la

descripción de los recursos. Si bien su uso se ha extendido gracias principalmente a su fácil aplicación y a su potencial como formato de intercambio de metadatos entre aplicaciones, su planteamiento inicial general ha propiciado posteriormente la elaboración de distintos perfiles de aplicación orientados a la descripción de tipos de contenidos específicos [8].

Las implementaciones de Dublin Core usan generalmente XML. Dublin Core se define por ISO en su norma ISO 15836 del año 2003, y la norma NISO Z39.85-2007.

Uno de los objetivos de Dublin Core es describir a través de metadatos objetos digitales como audio, texto, e incluso imágenes. Los campos de Dublin Core describen al objeto digital guardado en algún medio de almacenamiento lo que permite el acceso al objeto por medio de la URL.

Las características que postulan a Dublin Core como un sistema para ser adoptado en la descripción de recursos electrónicos se desglosan a continuación [8]:

- Simpleza.

El Dublin Core está diseñado para ser utilizado tanto por personas sin conocimientos de catalogación, como por catalogadores con experiencia en el uso de modelos de descripción de recursos tradicionales.

- Interoperabilidad Semántica.

En el ámbito de Internet, los distintos métodos para describir metadatos, interfieren con la capacidad de buscar traspasando los límites de los formatos. Promover el uso de un conjunto de descriptores que ayude a unificar y estandarizar el contenido de los datos, aumenta la posibilidad de interoperabilidad semántica de las disciplinas.

- Flexibilidad.

Aunque es motivado inicialmente por la necesidad de la descripción de recursos por los autores-generadores de cada sitio, el Dublin Core tiene suficiente flexibilidad como para codificar adicionalmente, estructuras con semántica más elaboradas que puedan ser aplicadas en la descripción de recursos tradicionales.

- Modularidad de metadatos en la web.

La amplia diversidad de metadatos en la web, requiere de un medioambiente que soporte la coexistencia de muchos desarrollos independientes y la existencia de paquetes de metadatos. El núcleo de Dublin Core apunta específicamente a la recuperación de recursos, pero se desarrollan distintos campos para ser utilizados en otros ámbitos.

- Arquitectura de metadatos para el web.

El consorcio World Wide Web (W3C) es el foro principal en los estándares para la web, y ha comenzado recientemente a centrarse en la configuración de una arquitectura de metadatos para la web. El Marco de Descripción de Recurso o RDF, se está desarrollando para sustentar diversas necesidades de metadatos que tienen proveedores y vendedores de información. Los representantes del núcleo de Dublin Core están implicados activamente en el desarrollo de esta configuración, aportando la perspectiva de la biblioteca digital, en el desarrollo de esta arquitectura.

El sistema Dublin Core está constituido por quince metadatos básicos, los cuales se dividen en tres apartados: contenido, propiedad intelectual y aplicación.

En contenido se puede encontrar:

- Título: se refiere al título que lleva por nombre el documento.
- Materia: en este campo se hace referencia a los diversos temas que puede contener el material.
- Descripción: en este campo se hace un breve resumen sobre el contenido del objeto digital.
- Fuente: es como una pequeña ficha bibliográfica que se elabora para asentar los datos sobre la procedencia del documento original.
- Relación: este campo tiene que ver con el material principal u objetos de su misma referencia, ya sea una colección, una serie, un documento, etcétera.
- Cobertura: este campo se refiere al proyecto o sitio donde estará resguardada la información, aquí pueden anotarse fechas, zonas geográficas.

En propiedad intelectual se encuentran los siguientes campos:

- Creador: aquí se anota el autor intelectual de la obra o documento original.
- Editor: este campo se refiere al sitio o colección responsable, a la que está adscrito el material.
- Colaborador: en este campo se anotan, si es que se da el caso, el nombre u organización que contribuyó a la creación del material, que no se especificó en la parte de Creador.
- Derechos: se anota en este campo el nombre o la institución a la cual pertenece el material y lo facilitó.

En aplicación se encuentran los restantes campos:

- Fecha: se anota la fecha de elaboración del registro.

- **Formato:** en este campo se registra el tipo de extensión con que se presenta el objeto digital, ya sea HTML, JPG, GIF o PDF.
- **Identificador:** se refiere a la dirección electrónica de origen a la que está adscrito el material. Para ello se utilizan las siglas URL.
- **Tipo:** aquí se menciona la presentación que tiene el objeto digital, ya sea como texto, audio, vídeo, etcétera.
- **Lenguaje:** en este campo se establecen las siglas correspondientes al idioma en que se presenta la publicación.

### 1.6 Herramientas que implementan la migración de datos

En el mundo actualmente existen varias herramientas para proveer datos a partir de colecciones de metadatos almacenados en distintos tipos de base datos. Los lenguajes de programación empleados para desarrollar estas herramientas son en su mayoría PHP, Java o Perl, así como tecnologías web ASP o JSP. Entre los desarrolladores de estas herramientas se encuentran el Laboratorio de Investigación en Bibliotecas Digitales (DLRL) del Instituto Politécnico de Virginia (Virginia Tech), así como el Laboratorio de Tecnologías Interactivas y Cooperativas (ICT), el Centro de investigación en Tecnologías de Información y Automatización (CENTIA) de la Universidad de Las Américas Puebla, México [9].

A continuación se muestran algunas de estas herramientas proveedoras de datos:

- Open Archives in a BOX (OAIB).

Archivos Abiertos en una Carpeta (OAIB) es una sencilla aplicación para exportar metadatos almacenados en un sistema de gestión de bases de datos relacionales como Microsoft Access, Oracle o MySQL. Se creó en 2001 por DLT (Tecnologías de Bibliotecas Digitales) en Estados Unidos. La herramienta es fácil de instalar sin necesidad de conocimientos de programación, ya que no es necesario hacer ningún tipo de implementación por parte del usuario. Al configurar el sistema se crea un fichero de configuración en XML que establece cómo acceder a la base de datos y como traducir los “verbos” en consultas SQL, así como traducir las respuestas a estas consultas en SQL en registros de metadatos.

- Virginia Tech OAI (VTOAI).

Es una aplicación en Perl enfocada a la generación del esqueleto o estructura principal de un servidor a partir de los metadatos almacenados en bases de datos relacionales multitablas. Es un producto del laboratorio de investigación en bibliotecas digitales de Virginia Tech y la Biblioteca Digital del Laboratorio

de Investigación (DLRL).

Es una herramienta que no requiere instalación ni compilación, ya que solo es necesario copiar los guiones Perl y contar con un servidor web con capacidades para utilizar guiones de Interfaz de Entrada Común (CGI). Las extensiones, configuraciones y contenedores se especifican empleando XML Schema que es un lenguaje de esquema utilizado para describir la estructura y las restricciones de los contenidos de los documentos XML, y cuenta con composición de código para componentes o bibliotecas de componentes separados.

➤ Proveedor de Datos PHP (PHP OAI Data Provider).

Es un desarrollo de la Universidad de Oldenburg, Alemania. Incluye la compresión de respuestas en XML lo que reduce significativamente el volumen de datos a transferir. Es una aplicación de configuración sencilla y se puede conectar con varias bases de datos. Además, puede instalarse en plataformas Linux, Unix y Windows, requiere tecnologías Apache y PHP y soporta bases de datos relacionales como Oracle o MySQL.

### **1.6.1 Resultados del estudio de las herramientas para la migración de datos**

La herramienta para la migración de datos propuesta en este trabajo de diploma debe ser capaz de seleccionar uno o varios registros, además debe realizar la migración de Marc 21 hacia Dublin Core de los registros seleccionados. Una vez migrados los registros, la herramienta debe permitir seleccionar los registros migrados y además seleccionar los documentos asociados a estos registros bibliográficos para generar una carpeta por cada registro con la estructura adecuada para que puedan ser importadas hacia el Repositorio institucional.

A partir del estudio de estas herramientas se determina que no cumplen con los requisitos para la migración de registros bibliográficos del SIGB hacia el Repositorio institucional. Sin embargo, se determina que el lenguaje Perl permite un manejo adecuado de los formatos Marc21 y Dublin Core, cuentan con módulos para la conversión de uno a otro. Además, el estudio de estas herramientas permite alcanzar un mejor dominio sobre el tema de migraciones de registros bibliográficos, lo cual contribuye a un mejor desarrollo de la herramienta propuesta.

## 1.7 Estructura de carpeta para importar en el Repositorio institucional

A un Repositorio normalmente le llega el momento de realizar una importación masiva de registros, sin tener que recurrir al método de carga de uno en uno de la interfaz de usuario. El método estándar de importación es mediante el formato simple de archivo y el uso por líneas de comandos. El comando import tratará de importar una estructura de directorio-subdirectorios y ficheros con una estructura predeterminada [10].

La estructura para importar registros al Repositorio institucional de la Universidad de las Ciencias Informáticas está formada por un directorio, dentro del cual se encuentra una carpeta que contiene varios ficheros.

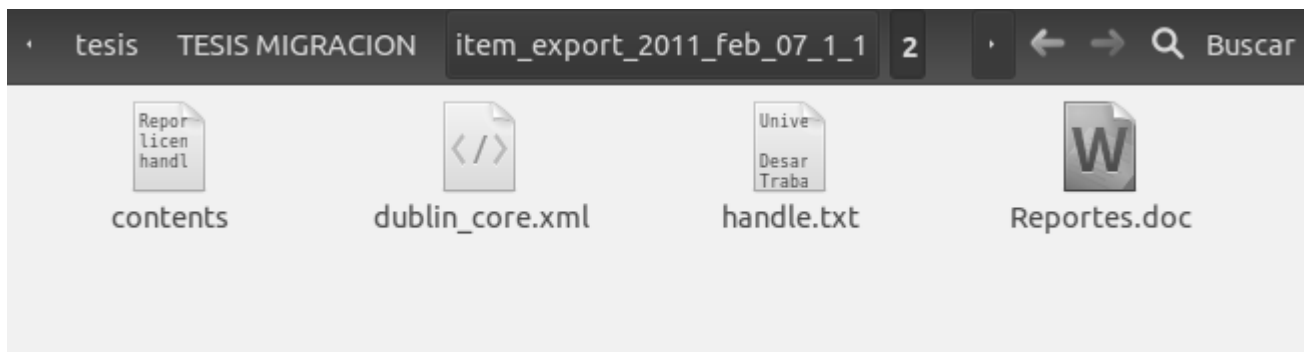


Figura 2 Ejemplo de estructura de carpeta para importación

El fichero **contents** enumera los ficheros que contiene el subdirectorio, junto con posibles indicaciones del bundle. Los bundles son agrupaciones de ficheros dentro del registro, que separan los diversos tipos de ficheros.

Además, dentro de cada carpeta va incluido un archivo **dublin\_core.xml**, que contiene los metadatos del registro bibliográfico que se desea importar hacia el Repositorio institucional. El fichero **handle** es utilizado por el Repositorio institucional como un medio de asignación de identificadores únicos globales a los archivos.



## 1.8 Herramientas de desarrollo

En este trabajo se seleccionan una serie de herramientas que son indispensables para la implementación de la solución propuesta, entre las cuales está el o los lenguajes que se utiliza para el desarrollo de la herramienta, el gestor de base de datos y el servidor web.

### 1.8.1 Gestor de base de datos

El gestor de base de datos que se utiliza es MySQL en su versión 5.1, ya que es un sistema de gestión de bases de datos relacional, licenciado bajo la GPL de la GNU. Su diseño multihilo le permite soportar una gran carga de forma muy eficiente. Fue creada por la empresa sueca MySQL AB, que mantiene el derecho de copia (copyright) del código fuente del servidor SQL, así como también de la marca.

MySQL está implementado bajo software libre, es uno de los gestores de bases de datos más usados en el mundo del software libre, debido a su gran rapidez y facilidad de uso. Esta gran aceptación es debido, en parte, a que existen infinidad de librerías y otras herramientas que permiten su uso a través de gran cantidad de lenguajes de programación, además de su fácil instalación y configuración.

Las principales características que posee este gestor de bases de datos son las siguientes [11]:

- Aprovecha la potencia de sistemas multiprocesador, gracias a su implementación multihilo.
- Soporta gran cantidad de tipos de datos para las columnas.
- Dispone de API's en gran cantidad de lenguajes (C, C++, Java, PHP).
- Gran portabilidad entre sistemas.
- Soporta hasta 32 índices por tabla.
- Logra mantener un buen nivel de seguridad en los datos.

### 1.8.2 Servidor web

El servidor web escogido para la implementación de la herramienta para la migración de datos es Apache por su excelencia, robustez y su estabilidad. Apache es una muestra de que el trabajo voluntario y cooperativo dentro de Internet es capaz de producir aplicaciones de calidad profesional difíciles de igualar.

El servidor Apache es un servidor web de código abierto para plataformas Unix (BSD, GNU/Linux, etc.), Microsoft Windows, Macintosh y otras, que implementa el protocolo HTTP/1.1 y la noción de sitio virtual. Es el servidor HTTP más usado (56 % de los servidores Web usan Apache) [12].

La licencia Apache es una de las licencias BSD, no es GPL. Esta licencia te permite modificar el código fuente (incluso productos propietarios) siempre que les reconozcas su trabajo.

El servidor web Apache posee las siguientes características [13]:

- Es posible ejecutarlo en una multitud de Sistemas Operativos, lo que lo hace prácticamente universal.
- Apache es una tecnología gratuita de código fuente abierta. El hecho de ser gratuita es importante pero, no tanto como que se trate de código fuente abierto.
- Apache es un servidor altamente configurable de diseño modular. Es muy sencillo ampliar las capacidades del servidor web Apache. Actualmente existen muchos módulos para Apache que son adaptables a este y se encuentran disponibles para ser instalados.
- Apache es muy utilizado con lenguajes de programación como Perl, PHP y otros lenguajes script. También trabaja con java teniendo todo el soporte que se necesita para tener páginas dinámicas.
- Apache permite personalizar la respuesta ante los posibles errores que se puedan dar en el servidor. Es posible configurar Apache para que ejecute un determinado script cuando ocurra un error en concreto.

## 1.9 Metodología de desarrollo

El desarrollo de un software debe regirse por una guía que indique en cada momento de su ciclo de vida quien debe hacer qué, cómo y cuándo. El análisis y diseño del software en construcción es de suma importancia y es desarrollado con diferentes enfoques según la metodología de desarrollo de software usada [14].

Luego de un estudio realizado de las principales metodologías de desarrollo que existen en el mundo se decidió utilizar RUP (Rational Unified Process). Se escoge el uso de esta metodología ya que ella implementa buenas prácticas actuales de ingeniería del software, además posee un desarrollo iterativo e incremental del software, una administración del software, un uso de arquitecturas basadas en requerimientos, verificación de la calidad del software y un control efectivo de cambios.

Además RUP describe cómo obtener requerimientos, organizarlos y documentarlos. También es capaz de rastrear y documentar decisiones, captar y documentar requerimientos del negocio. RUP guía en la configuración de un ambiente de proceso apropiado a cada proyecto.

## **1.10 Lenguajes a utilizar**

### **1.10.1 Perl**

Se escoge este lenguaje ya que la herramienta se va a encontrar incluida como un módulo en el SIGB. El código fuente del SIGB se encuentra en lenguaje Perl. Además, Perl es un lenguaje de programación muy utilizado para construir aplicaciones CGI para la web. Es un lenguaje de programación muy práctico para extraer información de archivos de texto y generar informes a partir del contenido de los ficheros. Perl es un lenguaje extensible a partir de otros lenguajes, ya que desde Perl se pueden hacer llamadas a subprogramas escritos en otros lenguajes. Perl está inspirado a partir de lenguajes como C, SH, AWK y SED (algunos provenientes de los sistemas Unix), pero está enfocado a ser más práctico y fácil que estos últimos.

No establece ninguna filosofía de programación concreta. No se puede decir que sea orientado a objetos, modular o estructurado aunque soporta directamente todos estos paradigmas; su punto fuerte son las labores de procesamiento de textos y archivos.

### **1.10.2 HTML**

El HTML (lenguaje de marcación de hipertexto), es el lenguaje de marcas de texto utilizado normalmente en la World Wide Web (W.W.W). Fue creado en 1986 por el físico nuclear Tim Berners-Lee, el cual tomó dos herramientas preexistentes: el concepto de hipertexto (conocido también como link o ancla) el cual permite conectar dos elementos entre sí y el SGML (Lenguaje Estándar de Marcación General) el cual sirve para colocar etiquetas o marcas en un texto que indique como debe verse. HTML no es propiamente un lenguaje de programación como C++, Visual Basic, etc., sino un sistema de etiquetas. HTML no presenta ningún compilador, por lo tanto, algún error de sintaxis que se presente no lo detectará y se visualizará en la forma que lo entienda.

El entorno para trabajar HTML es simplemente un procesador de texto, como el que ofrecen los sistemas operativos Windows (Bloc de notas), UNIX (el editor vi o ed.) o el que ofrece MS Office (Word).

### **1.10.3 CSS**

Las hojas de estilo en cascada, es un mecanismo simple que describe cómo se va a mostrar un documento en la pantalla, o cómo se va a imprimir, o incluso cómo va a ser pronunciada la información

presente en ese documento a través de un dispositivo de lectura. Esta forma de descripción de estilos ofrece a los desarrolladores el control total sobre el estilo y el formato de sus documentos.

CSS se utiliza para dar estilo a documentos HTML y XML, separando el contenido de la presentación. Los estilos definen la forma de mostrar los elementos HTML y XML. CSS permite a los desarrolladores web controlar el estilo y el formato de múltiples páginas web al mismo tiempo.

## **1.11 Herramientas CASE**

Las herramientas CASE (Computer Aided Software Engineering) son aplicaciones informáticas con el propósito de aumentar la productividad en el desarrollo de software, reduciendo el tiempo y los costos. Estas herramientas apoyan el modelado y diseño en todas las fases del ciclo de vida del software [15].

Después de analizar las herramientas que existen se opta por elegir a Visual Paradigm para el desarrollo de la aplicación.

### **1.11.1 Visual Paradigm**

Visual Paradigm para UML es una herramienta profesional que soporta el ciclo de vida del software: análisis y diseño orientados a objetos, construcción, pruebas y despliegue. Este software de modelado permite la rápida construcción de aplicaciones de alta calidad, reduciendo los costos. Acelera el desarrollo de aplicaciones, sirviendo de intermediario visual entre arquitectos, analistas y diseñadores de software, mediante un ambiente de modelado superior que posibilita un trabajo más fácil y dinámico [16].

Esta herramienta permite diseñar bases de datos y generar sus esquemas. Es integrable con varios entornos de desarrollo como NetBeans, Microsoft Visual Studio y Eclipse, facilitando una mayor rapidez en el análisis, captura, plan, desarrollo, comprobación y despliegue de los requisitos. Es multiplataforma y cuenta con una versión libre para la comunidad. Es capaz de generar la documentación del sistema en los formatos PDF, HTML y el formato de documentos de Microsoft Word. Además, permite importar proyectos de otras herramientas de modelado como Rational Rose, Erwin y Microsoft Visio. Soporta la revisión ortográfica, brindando sugerencias para los idiomas: inglés y español.

## **1.12 Herramientas de automatización del proceso de desarrollo**

### **1.12.1 Eclipse**

Eclipse es una plataforma de programación, usada para crear entornos integrados de desarrollo. Cuenta

con una característica ventajosa, y es la de permitir la integración con plugins de programación para proporcionar todas sus funcionalidades al frente de la plataforma a diferencia de otros entornos monolíticos donde las funcionalidades están todas incluidas, sean de interés o no para el usuario. Uno de estos plugins es el EPIC, destinado para la programación en el lenguaje Perl.

### **1.12.2 EPIC**

EPIC (Editor de Perl y Entorno de Desarrollo Integrado para Eclipse) es un Entorno de Desarrollo Integrado (IDE) de código abierto (con redactor y depuración de código incluidos) basado en la plataforma de Eclipse, compatible con Linux. Para la escritura de programas CGI o proyectos de Perl robustos con gran cantidad de módulos, EPIC es el más recomendable y extensible IDE disponible actualmente, gracias a la integración con todas las características principales y las convenciones de la Interfaz de Usuario Gráfica de Eclipse [17].

Ventajas del uso de EPIC:

- Fácil instalación en la plataforma Eclipse.
- Entre las principales funciones soportadas se encuentran el resaltado de sintaxis, autocompletado de funciones y métodos, soporte para documentación de Perl, plantillas de código, expresiones regulares.
- Herramientas de documentación para funciones de Perl incorporadas
- Los errores de sintaxis son detectados y explicados mientras se edita un archivo en tiempo real.

### **1.12.3 Control de versiones**

Para el control de versiones se seleccionó el Subversion y un plugin de Eclipse llamado Subclipse, ya que al estar integrado al Eclipse permite trabajar sobre las copias de trabajo como si fueran proyectos normales de Eclipse y desde ahí se pueden realizar todas las operaciones normales.

### **1.13 Conclusiones parciales**

No existe una herramienta que cumpla con las funcionalidades para la migración de registros bibliográficos desde un SIGB hacia un Repositorio institucional. Sin embargo, el estudio de las existentes brinda una panorámica de las herramientas de migración de datos y contribuye a definir las características que debe poseer la herramienta a desarrollar.

### **Capítulo 2: Descripción de la solución**

En este capítulo se presentan las características que debe cumplir la herramienta a desarrollar, además se realiza el diseño de la propuesta de solución y se enumerarán los requisitos funcionales y los requisitos no funcionales.

#### **2.1 Propuesta del sistema**

El desarrollo de la herramienta para la migración de datos desde el SIGB hacia el Repositorio institucional debe de ser de una gran utilidad para los administradores del SIGB, ya que lo librará de todo trabajo manual a la hora de migrar registros bibliográficos desde el formato MARC21 hacia el formato Dublin Core. La herramienta va a estar integrada al SIGB en forma de módulo. Va a contar con 2 funcionalidades principales. La primera funcionalidad es que permite realizar una búsqueda de registros bibliográficos para realizar la migración de datos sobre los registros obtenidos. La segunda funcionalidad es que una vez migrados estos registros permiten generar una carpeta para cada registro con la estructura adecuada para que puedan ser importadas hacia el Repositorio institucional.

#### **2.2 Modelo de dominio**

El modelo de dominio puede utilizarse para capturar y expresar el entendimiento ganado en un área bajo análisis como paso previo al diseño de un sistema. Es utilizado por el analista como un medio para comprender el sector de negocios al cual el sistema va a servir [18].

El modelo de dominio puede ser tomado como el punto de partida para el diseño del sistema. Cuando se realiza la programación orientada a objetos, el funcionamiento interno del software va a imitar en alguna medida a la realidad, por lo que el mapa de conceptos del modelo de dominio constituye una primera versión del sistema.

El objetivo de los modelos de dominio es ayudar a comprender los conceptos que utilizan los usuarios, los conceptos con los que trabajan y con los que debe trabajar la aplicación.

A continuación se muestra el diagrama del modelo de dominio:

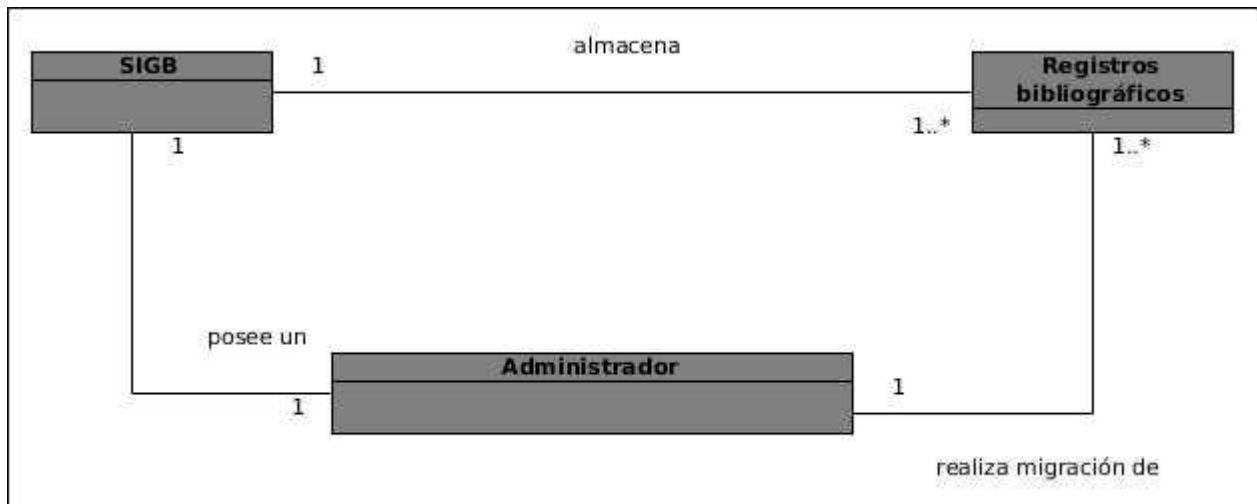


Figura 3 Diagrama del modelo de dominio

## 2.3 Requerimientos

Antes de comenzar cualquier proyecto es necesario tener en cuenta los objetivos y funcionalidades que debe cumplir dicho proyecto para comenzar su implementación. En la fase de requerimientos es donde se detectan estas funcionalidades. Los requerimientos expresan las condiciones o capacidades que deben cumplir los sistemas para poder lograr el desarrollo final del proyecto. Los requisitos se catalogan en funcionales y no funcionales.

Los requerimientos funcionales no son más que funcionalidades que el sistema debe ser capaz de realizar, características que debe cumplir el sistema y los no funcionales son una serie de características requeridas por el sistema.

En este trabajo de diploma se han identificado una serie de requerimientos los cuales debe cumplir la herramienta una vez desarrollada.

### 2.3.1 Requerimientos funcionales

- RF. 1 Seleccionar registros a exportar.
- RF. 2 Exportar registros bibliográficos seleccionados.
- RF. 3 Exportar todos los registros bibliográficos existentes.
- RF. 4 Mostrar registros migrados.
- RF. 5 Seleccionar archivos y documentos.

RF. 6 Generar estructura de carpeta.

### 2.3.2 Requerimientos no funcionales

- Interfaz gráfica.

RNF. 1 Interfaz sencilla donde se muestre solamente lo necesario.

RNF. 2 Mostrar mensaje de error en la introducción de datos de una forma sencilla y explicativa, la entrada de datos incorrecta será detectada por la herramienta para la migración de datos.

RNF. 3 Mostrar todos los textos y mensajes en español.

- Usabilidad.

RNF. 4 El sistema deberá tener una interfaz gráfica uniforme a través del mismo, incluyendo pantallas, menús y opciones.

- Confiabilidad.

RNF. 5 Se debe mantener la integridad de los registros originales en el proceso de migración de datos.

- Software.

La PC donde se ejecutará la herramienta debe cumplir ciertos requisitos los cuales son:

RNF. 6 GNU/Linux.

RNF. 7 Un intérprete de Perl 5.0.

RNF. 8 La PC del cliente debe contar con un navegador web.

### 2.3.3 Especificación de requerimientos

Tabla 1 Especificación de requerimientos

Numero	Nombre	Descripción	Complejidad	Prioridad para cliente
RF1	Seleccionar registros a exportar.	El sistema debe permitir seleccionar los registros que se deseen importar.	Baja	Baja
<b>Prototipo</b>				
Ver anexo 1.				



	<b>Campos</b>	<b>Tipos de Datos</b>	<b>Reglas o Restricciones</b>
	No procede.	No procede.	Debe seleccionar al menos un registro.
	<b>Observaciones</b>		
RF2	Exportar registros bibliográficos seleccionados.	El sistema debe ser capaz de permitir exportar los registros seleccionados por el administrador.	Alta  Alta
<b>Prototipo</b>			
Ver anexo 2.			
	<b>Campos</b>	<b>Tipos de Datos</b>	<b>Reglas o Restricciones</b>
	No procede.	No procede.	No procede.
	<b>Observaciones</b>		
RF3	Exportar todos los registros bibliográficos existentes.	El sistema debe permitir exportar todos los registros existentes.	Alta.  Alta.
<b>Prototipo</b>			
Ver anexo 3.			
	<b>Campos</b>	<b>Tipos de Datos</b>	<b>Reglas o Restricciones</b>

	No procede.	No procede.	No procede.
	<b>Observaciones</b>		
RF4	Mostrar registros migrados.	El sistema debe ser capaz de mostrar todos los registros que fueron migrados.	Alta.
	<b>Prototipo</b>		
	Ver anexo 6.		
	<b>Campos</b>	<b>Tipos de Datos</b>	<b>Reglas o Restricciones</b>
	No procede.	No procede.	No procede.
	<b>Observaciones</b>		
RF5	Seleccionar archivos y documentos.	El sistema debe ser capaz de permitir al administrador seleccionar los archivos y los documentos.	Alta.
	<b>Prototipo</b>		
	Ver anexo 4.		
	<b>Campos</b>	<b>Tipos de Datos</b>	<b>Reglas o Restricciones</b>
	No procede.	No procede.	No procede.
	<b>Observaciones</b>		
RF6	Generar estructura de carpeta.	El sistema debe ser capaz de permitirle al	Alta.

		administrador generar una carpeta para cada registro bibliográfico con la estructura adecuada.		
<b>Prototipo</b>				
Ver anexo 5.				
<b>Campos</b>		<b>Tipos de Datos</b>	<b>Reglas o Restricciones</b>	
No procede.		No procede.	No procede.	
<b>Observaciones</b>				

## 2.4 Modelo de casos de uso del sistema

Después de haber realizado una descripción de las funcionalidades que debe cumplir la herramienta se pasa a la elaboración del diagrama de casos de uso mediante el lenguaje de modelado UML. Antes se debe identificar los actores que intervendrán en el diagrama. Los actores se definen como los roles que puede tener un usuario.

En esta herramienta interactúa un actor del sistema quien se define a continuación:

Tabla 2 Actor del sistema

Actores	Justificación
Administrador	El administrador del sistema, es el único encargado para hacer uso de la herramienta, ya que posee los conocimientos necesarios.

Una vez identificados los actores del sistema se pasa a la elaboración del diagrama de casos de uso del sistema. A continuación se muestra el diagrama de casos de uso del sistema:

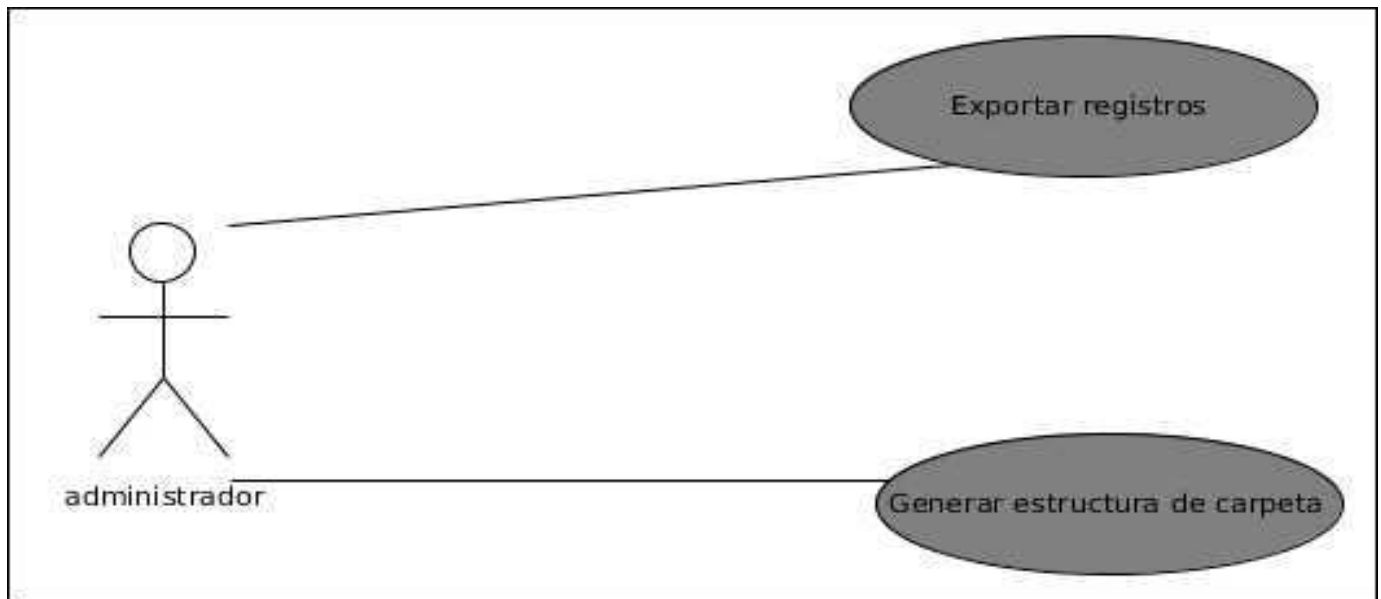


Figura 4 Diagrama de casos de uso

### 2.4.1 Especificación de casos de uso

CU1.Exportar registros:

Tabla 3 Especificación del CU Exportar registros

<b>Objetivo</b>	El administrador pretende migrar los registros existentes en el sistema, los cuales se encuentran en MARC21 hacia el formato Dublin Core.
<b>Actores</b>	Administrador.
<b>Resumen</b>	El caso de uso se inicia cuando el administrador selecciona los registros que desea migrar, la herramienta brinda la opción de seleccionar algunos mediante una búsqueda o seleccionarlos todos. Se selecciona la opción de exportarlos donde la herramienta ofrece la opción hacia dónde desea descargarlos.
<b>Complejidad</b>	Alta.
<b>Prioridad</b>	Crítico.
<b>Precondiciones</b>	Usuario autenticado.
<b>Flujo de eventos</b>	
<b>Flujo básico &lt;Curso normal de eventos&gt;</b>	

Actor	Sistema
1. Buscar registros.	2. Muestra el listado de los registros.
	3. Brinda la opción de seleccionar registros. 4. Brinda la opción de seleccionar todos los registros.
5. Selecciona registros.	6. Procede a migrar los registros seleccionados. 7. Termina el caso de uso.
<b>Flujo alternativo 1: "Selecciona todos los registros"</b>	
Actor	Sistema
1. Selecciona todos los registros.	2. Procede a migrar todos los registros. 7. Termina el caso de uso.
<b>Flujo alternativo 2: "No se encontraron registros"</b>	
Actor	Sistema
1. Buscar registros.	2. Muestra un mensaje donde especifica que no se encontraron los registros deseados.
<b>Requisitos no funcionales</b>	RNF.1, RNF.2, RNF.3.

CU2.Generar estructura de carpeta:

Tabla 4 Especificación del CU Generar estructura de carpeta

<b>Objetivo</b>	Con este caso de uso se pretende dejar lista la carpeta con toda la estructura correcta para que pueda ser importada por el Repositorio institucional.
<b>Actores</b>	Administrador.
<b>Resumen</b>	Primeramente se selecciona el registro, después se obtiene el documento asociado a dicho registro, luego se obtiene el handle y por último se obtiene el contents.
<b>Complejidad</b>	Alta.
<b>Prioridad</b>	Secundario.
<b>Precondiciones</b>	CU Exportar registros.
<b>Flujo de eventos</b>	

<b>Flujo básico &lt;Curso normal de eventos&gt;</b>	
<b>Actor</b>	<b>Sistema</b>
1. Buscar registro.	2. Muestra la ruta hacia el registro.
3. Selecciona el registro.	4. Muestra el registro seleccionado.
5. Buscar documento.	6. Muestra la ruta hacia el documento.
7. Selecciona el documento.	8. Muestra el documento seleccionado.
9. Selecciona subir archivos.	10. Muestra la opción de generar carpeta.
11. Selecciona generar carpeta.	12. Genera carpeta. 13. Termina el caso de uso.
<b>Flujo alterno 1: "Cancelar generar estructura de carpeta"</b>	
<b>Actor</b>	<b>Sistema</b>
1. Cancela generar estructura de carpeta.	2. Retorna al inicio del caso de uso.
<b>Requisitos no funcionales</b>	RNF.1, RNF.2, RNF.3

## 2.5 Análisis

El análisis del sistema es uno de los flujos de trabajos realizados durante el proceso de software, este se desarrolla fundamentalmente dentro de la fase de elaboración. El análisis consiste en obtener una visión del sistema, de modo que solamente se interesa por los requisitos funcionales [19].

En el modelo de análisis se utilizan varias clases las cuales son:

CI\_<Nombre de la clase>: Modelan la interacción entre los usuarios y el sistema.

CC\_<Nombre de la clase>: Encapsulan el comportamiento de cada caso de uso y coordinan el trabajo de las clases interfaz y entidad.

CE\_<Nombre de la clase>: Modelan toda la información del sistema que posee una vida larga y que puede ser persistente.

### 2.5.1 Diagramas de clases del análisis

CU1.Exportar registros:

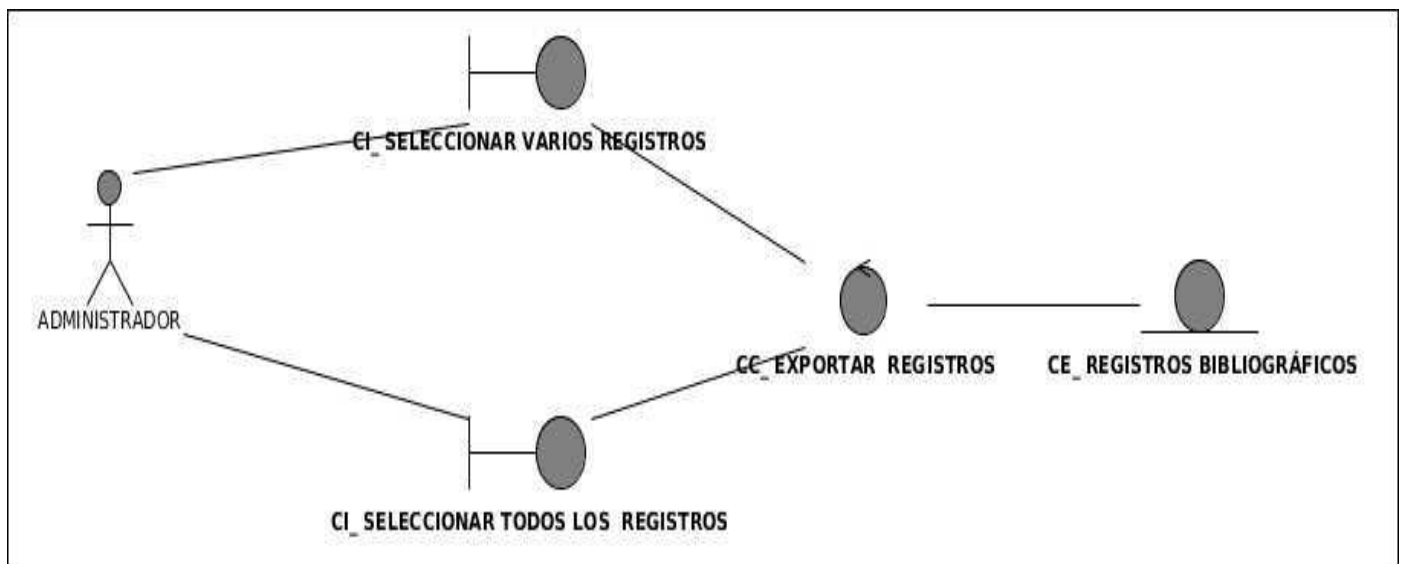


Figura 5 Diagrama de clases del análisis del CU Exportar registros

CU2.Generar estructura de carpeta:

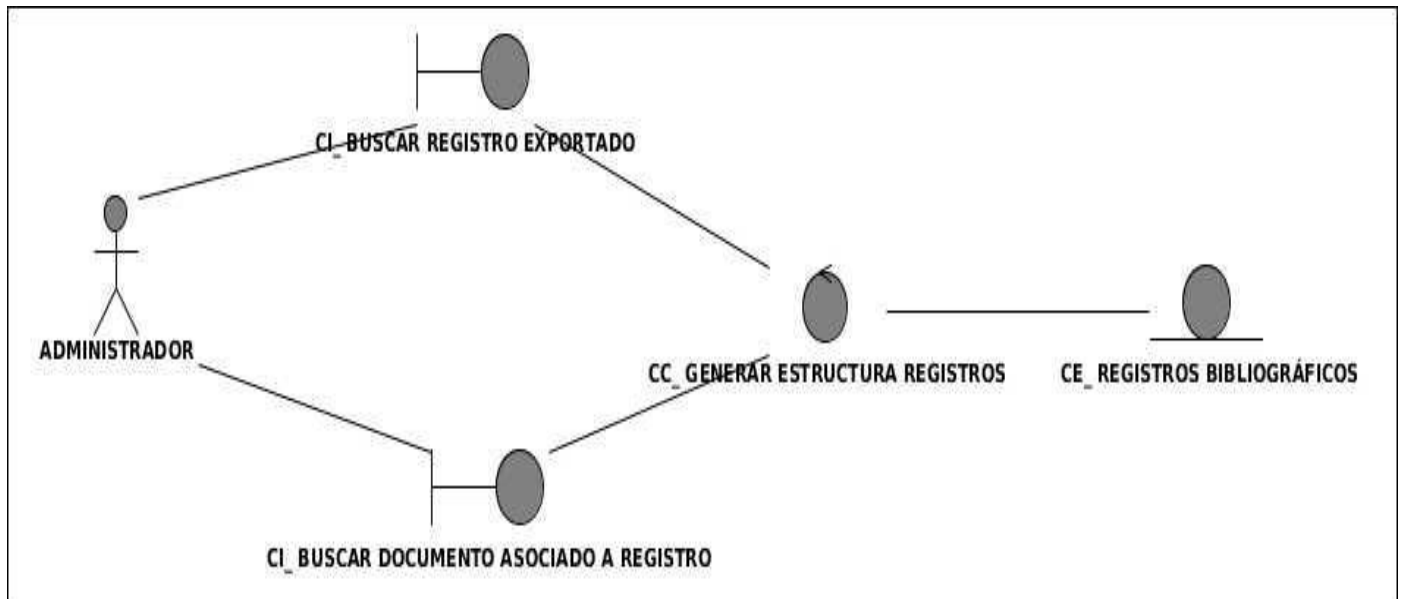


Figura 6 Diagrama de clases del análisis del CU Generar estructura de carpeta

## 2.6 Diseño

En la fase de diseño se modela el sistema, de manera que soporte todos los requerimientos, tanto funcionales como no funcionales. Uno de sus propósitos fundamentales es: “Crear una entrada apropiada y un punto de partida para actividades de implementación subsiguientes capturando los requerimientos o subsistemas individuales, interfaces y clases”, se dice que el diseño es un esquema a la implementación [19].

### 2.6.1 Descripción de la arquitectura

Una arquitectura de software, también denominada arquitectura lógica, consiste en un conjunto de patrones y abstracciones coherentes que proporcionan el marco de referencia necesario para guiar la construcción del software de un sistema de información. La arquitectura de software establece los fundamentos para trabajar en una línea común que permite alcanzar los objetivos y necesidades del sistema de información [20].



### 2.6.2 Modelo-Vista-Controlador (MVC)

El MVC es un patrón de arquitectura de software que separa los datos de una aplicación, la interfaz de usuario, y la lógica de control en tres componentes distintos. El patrón MVC se ve frecuentemente en aplicaciones web [21].

A continuación se muestra un diagrama con el funcionamiento del patrón arquitectónico MVC [22].

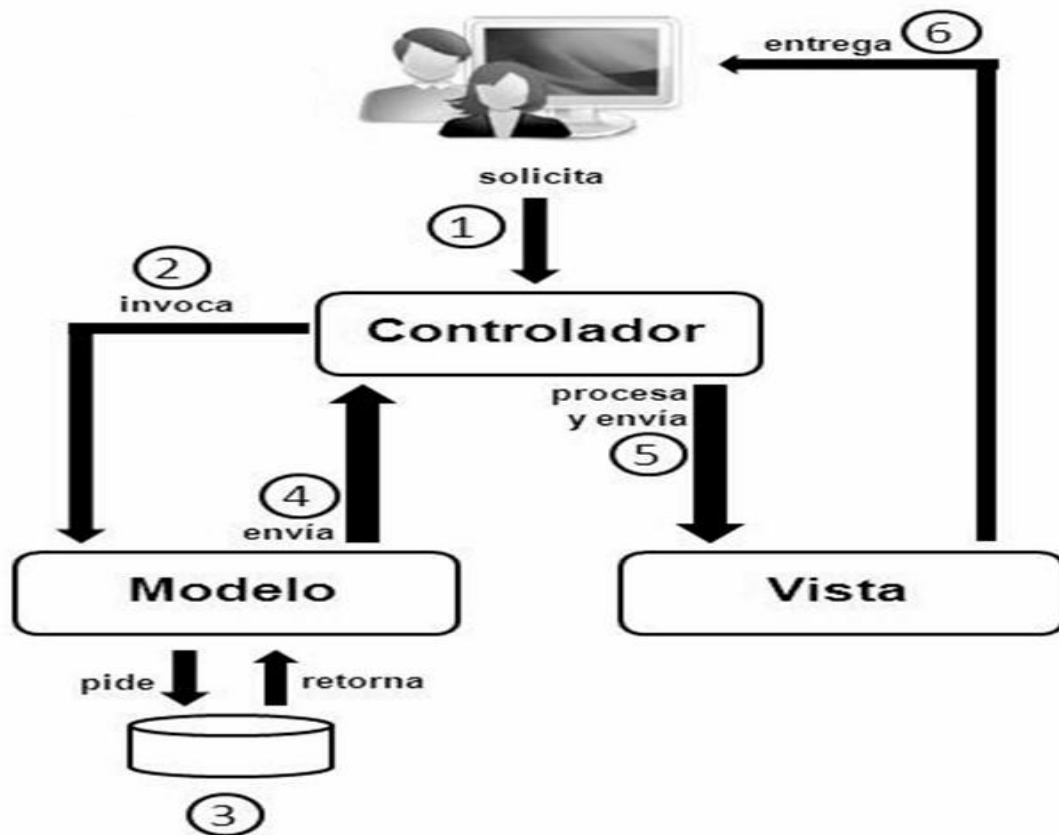


Figura 7 Funcionamiento del MVC

El modelo es el componente encargado del acceso a datos, modela los datos y el comportamiento detrás de los procesos del negocio, también se realizan las consultas a la base de datos, se ejecutan los cálculos de los procesos de negocio y se procesan las diferentes órdenes. Se encapsulan los datos y los comportamientos que son independientes de la presentación.

En la vista se despliega la información de acuerdo con el tipo de cliente, así como los resultados de la lógica del negocio (modelo). Una vista puede ser una página web o una parte de una página.

Sirve como una conexión lógica entre la interacción del usuario y los servicios de negocio disponibles. En otras palabras, el controlador actúa como intermediario entre el modelo, la vista y cualquier otro recurso necesario para generar una página.

El funcionamiento de estos componentes se lleva a cabo de la siguiente manera:

Primeramente el controlador recibe una orden y decide quién la lleva a cabo en el modelo. Una vez que el modelo termina sus operaciones devuelve el flujo al controlador y este envía el resultado a la vista, mostrándose el resultado al usuario.

### 2.6.3 Diagramas de clases del diseño

CU1.Exportar registros:

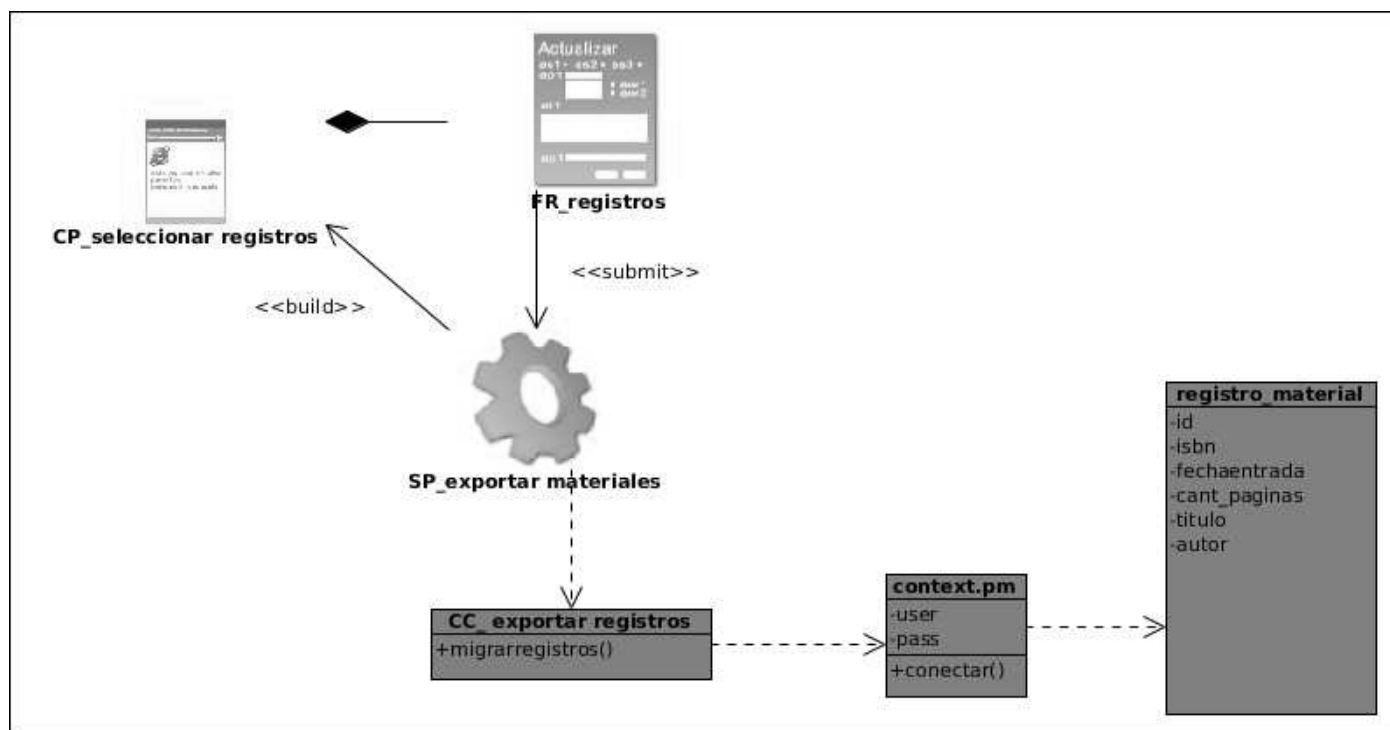


Figura 8 Diagrama de clases CU Exportar registros

CU2. Generar estructura de carpeta:

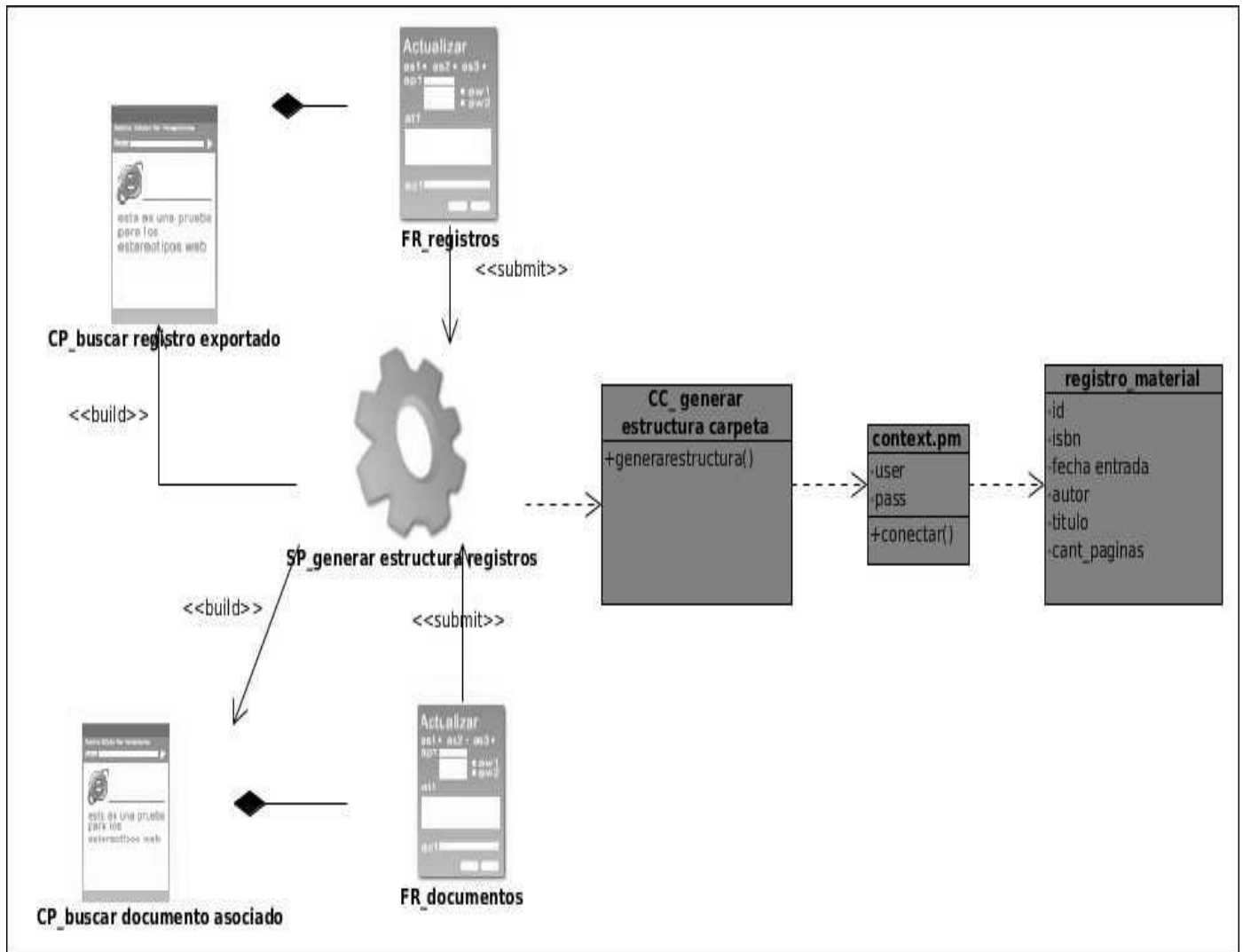


Figura 9 Diagrama de clases CU Generar estructura de carpeta

### 2.6.4 Diagramas de colaboración

CU1. Exportar registros:

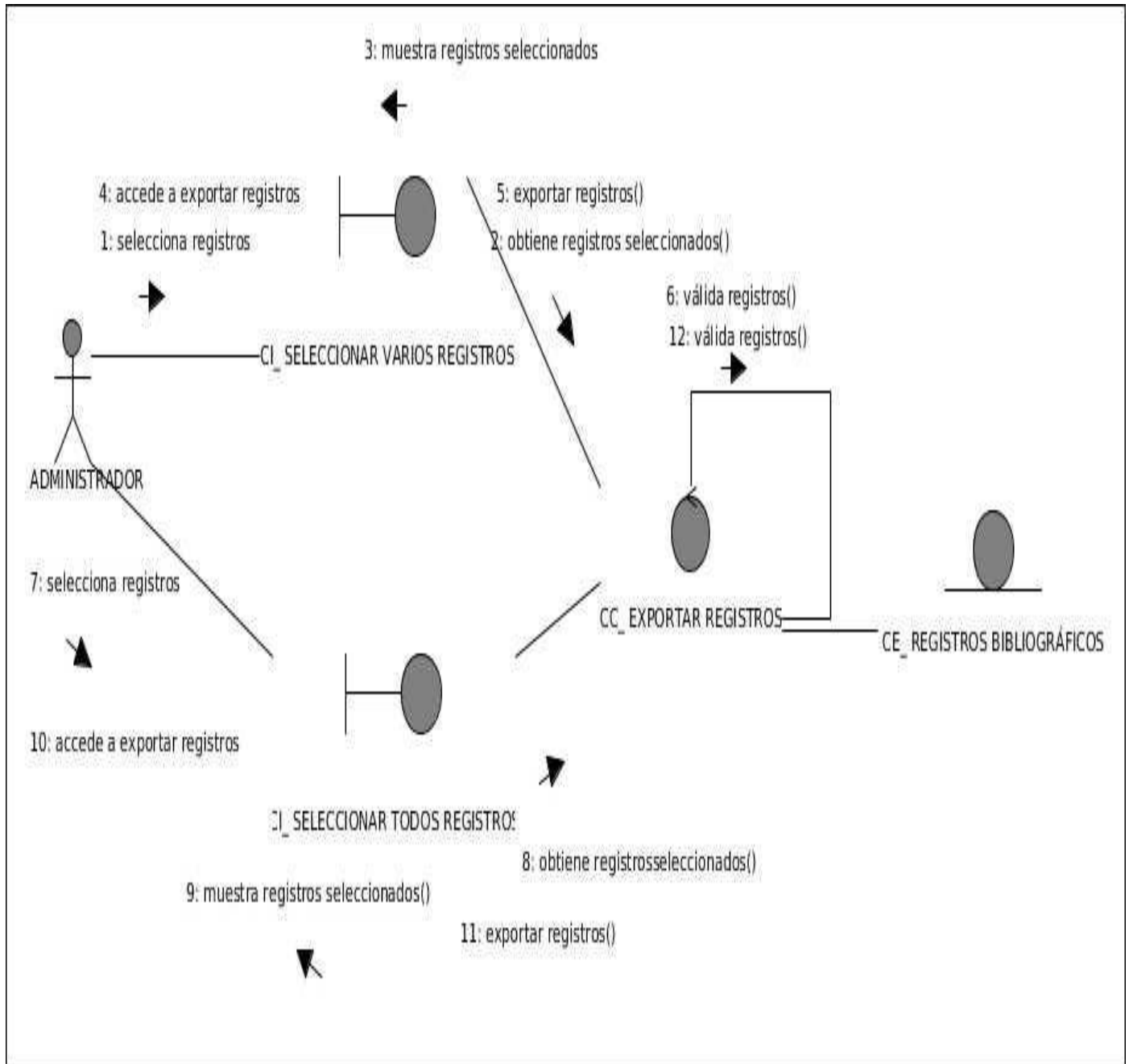


Figura 10 Diagrama de colaboración CU Exportar registros

CU2. Generar estructura de carpeta:

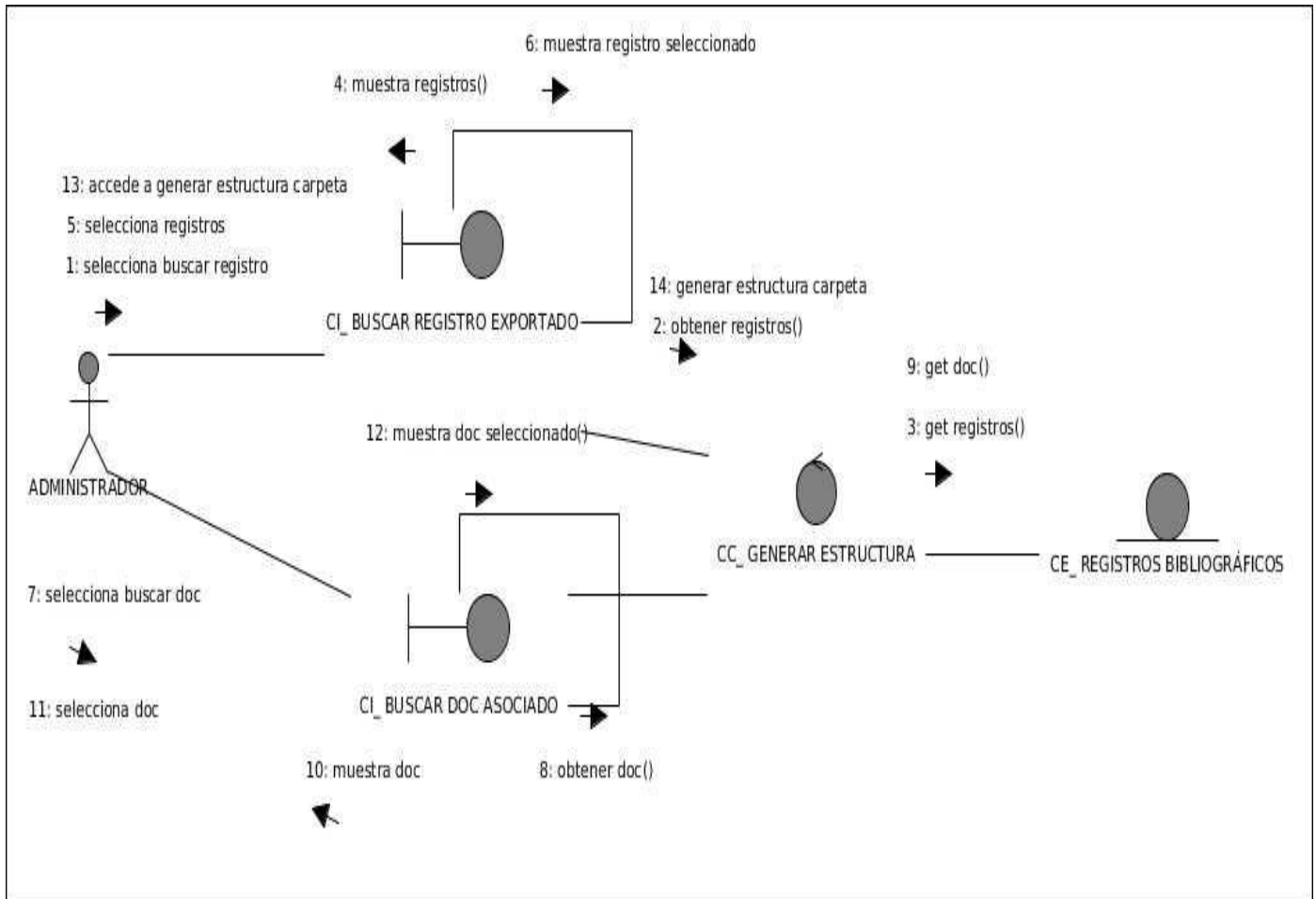


Figura 11 Diagrama de colaboración CU Generar estructura de carpeta

2.6.5 Modelo entidad-relación

El modelo entidad-relación es un modelo conceptual o teórico que se utiliza para el modelado de datos de un sistema de información. Estos modelos expresan entidades relevantes para dicho sistema así como sus interrelaciones y propiedades, están basados además en una percepción de un mundo real que consiste en un grupo de elementos básicos llamados entidades y relaciones.

A continuación se representa el modelo entidad-relación de la base de datos, en la cual se reflejan las relaciones existentes entre cada tabla.

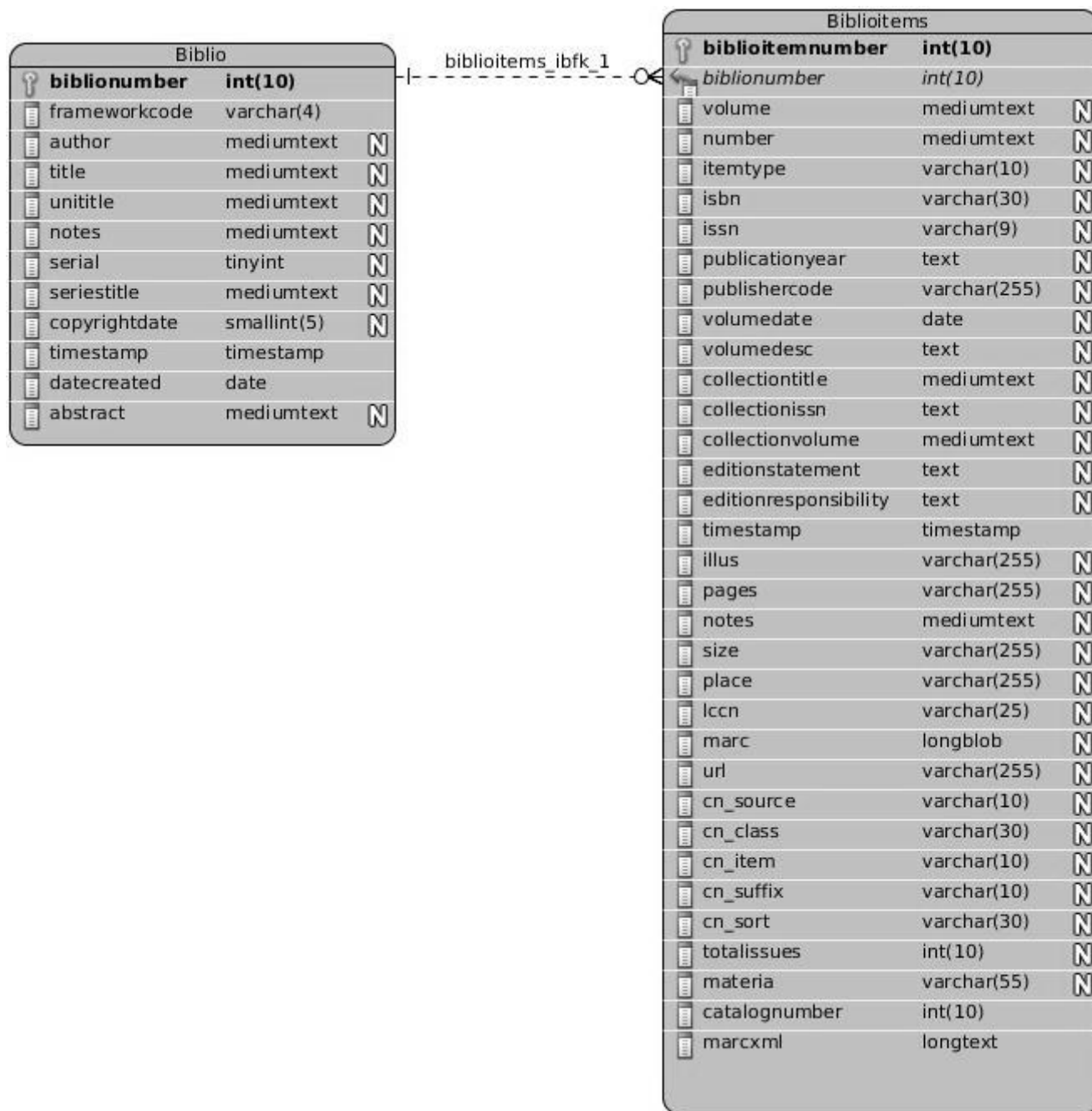


Figura 12 Modelo entidad-relación

### **2.7 Conclusiones parciales**

En este capítulo se dejan las puertas abiertas para el posterior desarrollo de la propuesta de solución planteada, obteniéndose a partir de un análisis las funcionalidades que debe tener la herramienta, las cuales se representaron mediante el diagrama de casos de uso describiéndose paso a paso todas las acciones del actor y las respuestas del sistema. Teniendo en cuenta todas estas características se puede pasar a desarrollar la herramienta poniendo en práctica el cumplimiento de los requisitos tanto funcionales como no funcionales.

### Capítulo 3: Implementación y prueba

Para iniciar la implementación se comienza con el resultado del diseño y se implementa el sistema en término de componentes: ficheros de código fuente, scripts, ficheros de código binario, ejecutables y similares. La mayor parte de la arquitectura del sistema es capturada durante el diseño, siendo el objetivo general de la implementación, desarrollar la arquitectura y el sistema en conjunto. El modelo de implementación describe como se organizan los componentes de acuerdo con los mecanismos de estructuración y modularización disponibles en el entorno de implementación y en el lenguaje de programación utilizado, y como depende un componente de otro. Durante las pruebas se verifica el resultado de la implementación probando cada construcción, incluyendo tanto construcciones internas, como intermedias, así como las versiones finales del sistema a ser entregadas a terceros [23].

#### 3.1 Diagrama de componentes

Los diagramas de componentes describen los elementos físicos del sistema y su relación, mostrando las dependencias lógicas entre componentes de software. El diagrama de componentes hace parte de la vista física de un sistema, la cual modela la estructura de implementación de la aplicación por sí misma, su organización en componentes y su despliegue en nodos de ejecución. La vista de implementación se representa con los diagramas de componentes [24].

A continuación se muestra el diagrama de componentes:



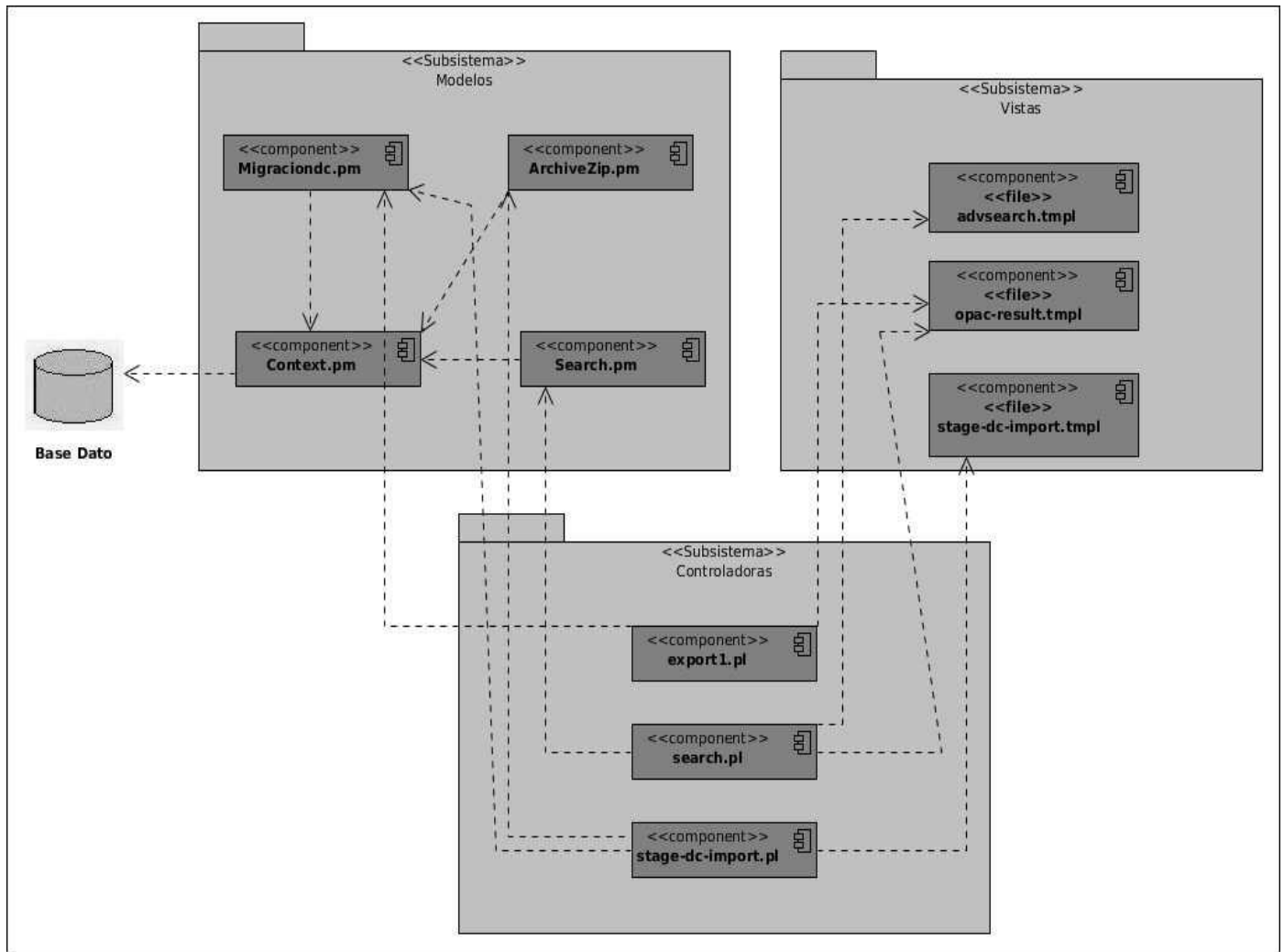


Figura 13 Diagrama de componentes

El subsistema Vistas contiene los componentes necesarios para la interacción del usuario con la herramienta de migración, los cuales son manejados por el subsistema de Controladoras. El subsistema Controladoras, es el controlador de las actividades de la herramienta, este contiene los ficheros de código fuente, los cuales interactúan con los demás subsistemas coordinando las acciones del software. El subsistema Modelo es el encargado de la interacción con la base de datos, para de esta forma gestionar la información con la que trabaja la herramienta.

### 3.2 Diagrama de despliegue

El modelo de despliegue es un modelo de objetos que describe la distribución física del sistema en términos de cómo se distribuye la funcionalidad entre los nodos de cómputo. El modelo de despliegue se utiliza como entrada fundamental en las actividades de diseño e implementación debido a que la distribución del sistema tiene una influencia principal en su diseño. Cada nodo representa un recurso de cómputo, normalmente, un procesador o un dispositivo hardware similar. Los nodos poseen relaciones que representan medios de comunicación entre ellos, tales como Internet, Intranet, bus y similares.

El modelo de despliegue puede describir diferentes configuraciones de red, incluidas las configuraciones para pruebas y simulación. La funcionalidad (los procesos) de un nodo se define por los componentes que se distribuyen sobre ese nodo. El modelo de despliegue en sí mismo representa una correspondencia entre la arquitectura de software y la arquitectura del sistema (el hardware).

A continuación se muestra el diagrama de despliegue:

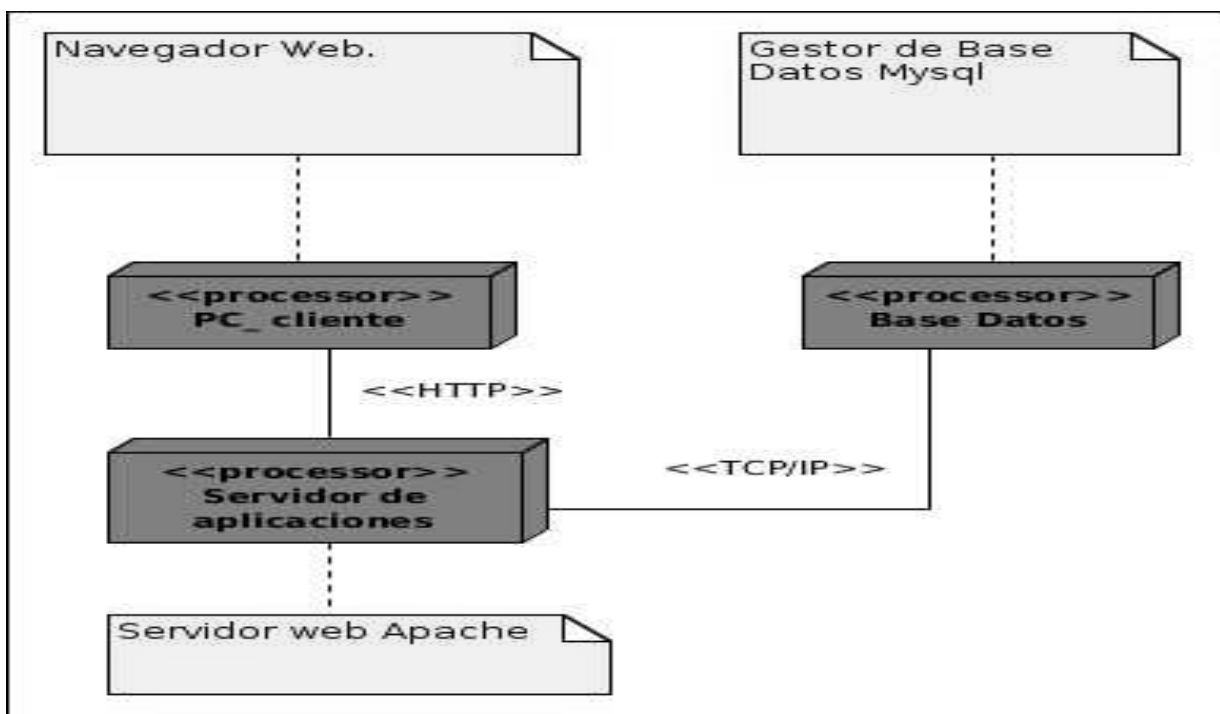


Figura 14 Diagrama de despliegue

### 3.3 Pruebas

#### 3.3.1 Modelos de pruebas

Las pruebas son una actividad en la cual un sistema o componente es ejecutado bajo condiciones o requerimientos especificados, los resultados son observados y registrados, y una evaluación es hecha de algún aspecto del sistema o componente. La prueba de software es un elemento crítico para la garantía de la calidad del software y representa una revisión final de las especificaciones del diseño y de la codificación. Existen dos métodos de pruebas fundamentales: los métodos de caja negra y los de caja blanca.

#### 3.3.2 Pruebas de caja negra

Una prueba de caja negra verifica el comportamiento observable externamente del sistema, o sea, se refiere a las pruebas que se llevan a cabo sobre la interfaz del software. Pretenden demostrar que las funciones del software son operativas, que la entrada se acepta de forma adecuada y que se produce un resultado correcto, así como que la integridad de la información externa se mantiene. Especifica como probar un caso de uso o un escenario específico de un caso de uso. Dentro de las pruebas de caja negra se encuentran las pruebas de sistema. A continuación se muestra el resultado de este tipo de prueba realizada a la aplicación [25].

En este trabajo se realizan pruebas funcionales mediante el método de caja negra aplicando la técnica de partición de equivalencia, la cual permite examinar los valores válidos e inválidos de las entradas existentes en el software, descubre de forma inmediata una clase de errores que, de otro modo, requerirían la ejecución de muchos casos antes de detectar el error genérico.

En las siguientes tablas se muestra de manera resumida la descripción de los casos de prueba desarrollados para cada caso de uso definido, especificando la información de entrada, los resultados obtenidos una vez ejecutado el caso de prueba y las condiciones que deben cumplirse mientras este se ejecuta.

Para que se pueda ejecutar el caso de uso Exportar registros, deben existir en el catálogo del SIGB los registros bibliográficos a los que se les pretende realizar la migración de datos.

Tabla 5 Caso de prueba Exportar registros

Nombre del caso de uso	Exportar registros	
Entradas	Resultados	Flujo central
El administrador accede a buscar registros y no especifica un criterio de búsqueda válido en campo de búsqueda. (¿,=,>)	El sistema le muestra un mensaje:  No se encontraron registros.	El administrador accede a realizar una migración de datos por lo que realiza una búsqueda de los registros bibliográficos a los cuales pretende realizar el proceso y en el campo de búsqueda no especifica variables válidas.
El administrador accede a buscar registros e introduce un criterio de búsqueda válido.	El sistema le muestra todos los resultados de la búsqueda.	El administrador accede a realizar una migración de datos por lo que realiza una búsqueda de los registros bibliográficos a los cuales pretende realizar el proceso y obtiene una lista de registros.
El administrador accede a exportar registros y no selecciona ningún registro.	El sistema le muestra un mensaje:  Debe seleccionar un registro.	El administrador selecciona el botón exportar sin antes haber seleccionado ningún registro de la lista obtenida.
El administrador accede a exportar registros y selecciona al menos un registro.	El sistema exporta los registros seleccionados.	El administrador selecciona uno o varios registros de la lista obtenida y selecciona la opción de exportar.

Tabla 6 Caso de prueba Generar estructura de carpeta

Nombre del caso de uso	Generar estructura de carpeta	
Entradas	Resultados	Flujo central
El administrador accede a seleccionar un archivo no válido. (Doc.)	El sistema le muestra un mensaje:  El archivo no es válido.	El administrador en el campo buscar registro selecciona registro con una extensión no válida.
El administrador accede a seleccionar un archivo. (XML)	El sistema importa el registro.	El administrador en el campo buscar registro selecciona registro.
El administrador accede a seleccionar un documento no válido. (Doc.)	El sistema le muestra un mensaje:  El documento no es válido.	El administrador en el campo buscar archivos selecciona uno o varios documentos con una extensión no válida.
El administrador accede a seleccionar un documento. (Zip, PDF)	El sistema importa el documento.	El administrador en el campo buscar archivos selecciona uno o varios documentos.
El administrador accede a generar estructura de carpeta. (XML, PDF o Zip)	El sistema genera un compactado con las carpetas.	El administrador selecciona la opción generar carpeta una vez que los registros y los documentos fueron subidos al sistema.

### 3.3.3 Resultados de las pruebas

Para la realización de las pruebas se realizaron dos iteraciones hasta llegar al resultado final. En la primera iteración se encontraron 4 no conformidades, cumpliéndose con un 20% de los requisitos definidos en el trabajo de tesis. En la segunda iteración no se identificó ninguna no conformidad, cumpliendo así con el 100% de los requisitos. Las pruebas se realizaron con todos los miembros del proyecto productivo Biblioteca, el cual pertenece al centro CENIA de la Facultad1 en la Universidad de las Ciencias Informáticas. Los resultados de las pruebas fueron satisfactorios cumpliéndose así con lo especificado en los requisitos.

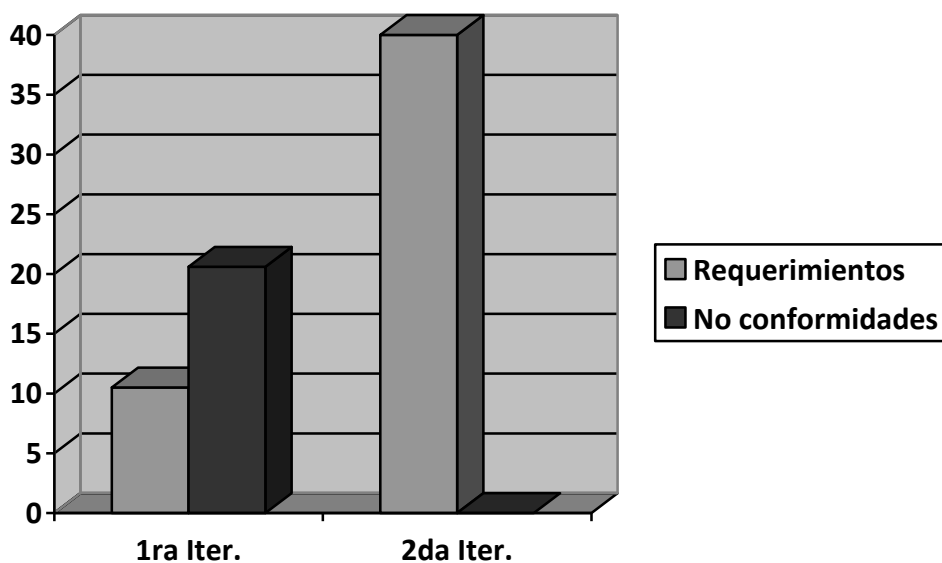


Figura 15 Gráfica de las iteraciones de los casos de pruebas

### 3.4 Conclusiones parciales

En este capítulo se expuso la estructura de la herramienta, mediante los modelos de despliegue y de componentes. Se hizo un análisis de los tipos de prueba que se pueden hacer a la herramienta para comprobar su correcto funcionamiento, realizándole pruebas de caja negra, que aunque no aseguran la no existencia de fallos, si dan una alta confiabilidad, y un nivel de calidad de la aplicación lista para pasar a una etapa de pruebas de aceptación por parte del cliente.

### **Conclusiones generales**

- La herramienta de migración de datos brinda al administrador de los sistemas de gestión bibliotecaria la posibilidad de desarrollar un trabajo más ágil y organizado.
- Las herramientas y tecnologías escogidas se ajustan a los requerimientos y propósitos que se persiguen en el presente trabajo.
- Se especifican las funcionalidades y cualidades que debe cumplir la herramienta para la migración dando paso a un mejor entendimiento de las actividades a informatizar.
- A la herramienta de migración desarrollada se le han realizado las pruebas pertinentes para asegurar su calidad. Las no conformidades encontradas fueron corregidas garantizando que se encuentra lista para su utilización.

### Recomendaciones

Se recomienda:

- Trabajar en el mejoramiento del diseño de las interfaces de usuario con el fin de lograr una mejor apariencia visual.
- Agregar nuevas funcionalidades a la herramienta para posibilitar la migración hacia otros formatos bibliográficos y que permita realizar el proceso de migración de Dublin Core hacia Marc21.
- Agregar a la herramienta la opción de que el usuario defina que campos de metadatos desea migrar.



### Referencias bibliográficas

- [1] **García Melero, L.A y García Camarero, E.** Automatización de bibliotecas. Madrid: Arco/Libros, 1999. [15 febrero 2012].
- [2] **Tránsito Ferreras Fernández.** Repositorio Institucionales, Salamanca, 2009. [17 febrero 2012].
- [3] **Clifford A Lynch.** "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age" ARL, no. 226. [20 febrero 2012].
- [4] **Talend.** Migración de datos [28 mayo 2012]. Disponible en: <http://es.talend.com/solutions-data-integration/data-migration.php>.
- [5] **Carsí, J. A. Ramos, I. Silva, J.** Un Generador Automático de Planes de Migración de Datos. I+D Computación. Departamento de Sistemas Informáticos y Computación, Vol. 1, No. 1, 2002. [2 marzo 2012]. Disponible en: <http://issi.dsic.upv.es/publications/archives/f-1050406290873/No1Art02.pdf>.
- [6] **CROSS, DAVID.** Data Munging with Perl. Manning Publications Co. 2001. [28 mayo 2012].
- [7] **Olaran Múgica, M.** Nuevas tendencias en la catalogación a comienzos del siglo XXI, 2011. [20 marzo 2012]. Disponible en: <http://docupedia.es/articulo/marc-21>.
- [8] **LoC.** MARC 21 Formats [28 mayo 2012]. Disponible en: <http://www.loc.gov/marc/marcdocz.html>.
- [9] **Dcmi Education Community.** Dublin Core Metadata Initiative, 2006. [20 febrero 2012]. Disponible en: <http://dublincore.org/groups/education/>.
- [10] **Bueno de la Fuente, Gema; Rodríguez Mateos, David.** Herramientas de software para OAI-PMH. 2007. Disponible en: [http://www.rojaseberhard.com.co/rojaseberhard/bibliotecologia/lib\\_oai.html](http://www.rojaseberhard.com.co/rojaseberhard/bibliotecologia/lib_oai.html).
- [11] **DSpace Documentación.** Documentation, DSpace. 2010. [6 abril 2012]. Disponible en: [http://www.dspace.org/1\\_5\\_2Documentation/DSpace-Manual.pdf](http://www.dspace.org/1_5_2Documentation/DSpace-Manual.pdf).
- [12] **Morejo Roque, Danichel y González Peraza.** Desarrollo de un bloque para Moodle 1.9.x que permita exportar cursos a un formato de libro electrónico interactivo. Habana, 2012. [6 marzo 2012].
- [13] **Luis Felipe Londono Muñoz.** Apache conceptos, instalación y configuración. 2010. [20 marzo 2012]. Disponible en: <http://es.scribd.com/doc/39143476/Presentacion-19-Apache-en-Linux>.

- [14] **Ciberaula**. Una Introducción a APACHE. Ciberaula. [28 mayo 2012]. Disponible en: [http://linux.ciberaula.com/articulo/linux\\_apache\\_intro](http://linux.ciberaula.com/articulo/linux_apache_intro).
- [15] **Pressman, Roger**. Ingeniería del Software. Un Enfoque Práctico. 2005. [26 mayo 2012].
- [16] **Andrés, María Mercedes Marques**. Herramientas CASE, 2008. [4 marzo 2012]. Disponible en: <http://www3.uji.es/~mmarques/f47/apun/node75.html>.
- [17] **UML tool**. UML tool business, process modeler and database designer for software development team. [ 24 marzo 2011]. Disponible en: <http://www.visual-paradigm.com>.
- [18] **EPIC**. Perl Editor and IDE for Eclipse. 2012. Disponible en: <http://www.epic-ide.org/>.
- [19] **E.V.A. UCI, I. D. S. Conferencia #2. Fase de Inicio, Modelo del Negocio, ISW 1.**
- [20] **Barrientos Enríquez, A. M.** El proceso Unificado de Modelado (RUP). 2005. [10 marzo 2012].
- [21] **Rumbaugh, J.** El lenguaje Unificado de modelado. Manual de Referencia.: Adison Wesley, 2000. [12 marzo 2012]. ISBN 84-7829-7037-7820.
- [22] **Eugenia Bahit**. POO y MVC en PHP:”El paradigma de la programación orientada a objetos en PHP y el patrón de arquitectura de software MVC”.
- [23] **G. M. Peñalver Romero**. “Metodología ágil para proyectos de software libre,” Master’s thesis, Universidad de las Ciencias Informáticas, 2008. [12 marzo 2012].
- [24] **Laurencio Pérez**. Sistema de Gestión de la Trayectoria Productiva de cada estudiante en la facultad 1, Universidad de las Ciencias Informáticas, 2009. [18 marzo 2012].
- [25] **2006**. Ingeniería del Software un enfoque práctico. La Habana: Félix Varela, 2006. Vol. 1.

### Bibliografía consultada

1. **Barrientos Enríquez, A. M.** El proceso Unificado de Modelado (RUP). 2005. [10 marzo 2012].
2. **Carsí, J. A. Ramos, I. Silva, J.** Un Generador Automático de Planes de Migración de Datos. I+D Computación. Departamento de Sistemas Informáticos y Computación, Vol. 1, No. 1, 2002. [2 marzo 2012]. Disponible en: <http://issi.dsic.upv.es/publications/archives/f-1050406290873/No1Art02.pdf>.
3. **Clifford A Lynch.** "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age" ARL, no. 226. [20 febrero 2012].
4. **Cross, David.** Data Munging with Perl. Manning Publications Co., 2001. [9 abril 2012].
5. **Dcmi Education Community.** Dublin Core Metadata Initiative, 2006. [20 febrero 2012]. Disponible en: <http://dublincore.org/groups/education/>.
6. **Duval, Erik. Hodgins, Wayne. Sutton, Stuart. Weibel, Stuartl.** Metadata Principales and Practicalities. D-Lib Magazine. 2002. [20 abril 2012]. Disponible en: <http://www.dlib.org/dlib/april02/weibel/04weibel.html>.
7. **DSpace Documentación.** Documentation, DSpace. 2010. [6 abril 2012]. Disponible en: [http://www.dspace.org/1\\_5\\_2Documentation/DSpace-Manual.pdf](http://www.dspace.org/1_5_2Documentation/DSpace-Manual.pdf).
8. **Francisca Hernández Carrascal.** Datos y metadatos, 2009. [18 abril 2012]. Disponible en: <http://www.slideshare.net/questc2d800/datos-y-metadatos-francisca-hernandez-carrascal>.
9. **García Melero, L.A y García Camarero, E.** Automatización de bibliotecas. Madrid: Arco/Libros, 1999. [15 febrero 2012].
10. **Library of Congress.** MARC to Dublin Core Crosswalk, 2008. [20 abril 2012]. Disponible en: <http://www.loc.gov/marc/marc2dc.html>.
11. **LoC.** Conociendo MARC Bibliográfico [15 abril 2012]. Disponible en: <http://www.loc.gov/marc/umbspa/um01a06.html>.
12. **LoC.** MARC 21 Formats [16 abril 2012]. Disponible en: <http://www.loc.gov/marc/marcdocz.html>.
13. **María Jesús Lamarca Lapuente.** Hipertexto: El nuevo concepto de documento en la cultura de la imagen. [20 abril 2012]. Disponible en: <http://www.hipertexto.info/documentos/metadatos.htm>.

14. **Olaran Múgica, M.** Nuevas tendencias en la catalogación a comienzos del siglo XXI, 2011. [20 marzo 2012]. Disponible en: <http://docupedia.es/articulo/marc-21>.
15. **Rodríguez Silva, Luis Racié.** El Marco Legal del Software Libre. Algunas Reflexiones. S.l.: UCI, 2006. [6 abril 2012].
16. **Tránsito Ferreras Fernández.** Repositorio Institucionales, Salamanca, 2009. [17 febrero 2012].

### Glosario de términos

**API:** (Application Programming Interface), un conjunto de rutinas, protocolos y herramientas para construir aplicaciones de interfaz.

**BSD:** (Berkeley Software Distribution), permite el uso de código fuente de software no libre. Incorpora cuatro conceptos adicionales que permiten que los usuarios puedan fácilmente extender el sistema: clase, herencia, tipos y funciones.

**Catalogación:** Tiene como objetivo principal de describir el documento físico y documentalmente para que pueda ser recuperado posteriormente con la mayor rapidez y precisión posible.

**CENIA:** Centro de Informatización Universitaria, se concentra en la informatización de la gestión universitaria y en el desarrollo del concepto de universidad digital.

**CGI:** Cell Global Identity.

**HTTP:** HTTP o HiperText Transfer Protocol (protocolo de transferencia de hipertexto). Es el grupo de reglas, o protocolos, que gobiernan la transferencia de hipertexto entre dos o más computadoras. Es muy cómodo y fácil de usar para transferir texto, imágenes, sonido.

**Linux:** Es el nombre de un núcleo, pero se suele denominar con este nombre a un sistema operativo de libre distribución software libre (y de código abierto), donde el código fuente está disponible públicamente y cualquier persona, con los conocimientos informáticos adecuados, puede libremente estudiarlo, usarlo, modificarlo y redistribuirlo.

**MARC:** Es un estándar internacional, que permite el intercambio y la representación de información bibliográfica por medio de computadoras siguiendo las reglas de catalogación de las publicaciones.

**MARC21:** Es un registro catalográfico legible por máquina (MAchine-Readable Cataloging) "Legible por máquina" significa que un tipo particular de máquina, una computadora, puede leer e interpretar los datos contenidos en un registro catalográfico.

**Migración de datos:** Es el proceso que tiene por objeto tanto la importación como la exportación de una determinada información almacenada en un sistema de bases de datos, para llevar a cabo su traspaso.

**MySQL:** Es un sistema de gestión de bases de datos relacional que cuentan con todas las características de un motor de BD comercial: transacciones atómicas, triggers, replicación, llaves foráneas entre otras. Su ingeniosa arquitectura lo hace extremadamente rápido y fácil de personalizar.

**PDF:** Acrónimo del inglés Portable Document Format (formato de documento portátil), es un formato de almacenamiento de documentos, de tipo compuesto (imagen vectorial, mapa de bits y texto). Está especialmente ideado para documentos susceptibles de ser impresos, ya que especifica toda la información necesaria para la presentación final del documento, determinando todos los detalles de cómo va a quedar, no requiriéndose procesos anteriores de ajuste ni de maquetación.

**Registro bibliográfico:** Conjunto de datos que describen y representan un documento concreto, con vistas a su recuperación dentro de un catálogo de biblioteca o en una base de datos bibliográfica. Contiene todos los datos necesarios para describir (de forma breve o completa, según el nivel de la descripción) el documento, presentando dichos datos en un formato bibliográfico específico normalizado. En la actualidad dicho formato es legible por ordenador (MARC). Un registro bibliográfico suele incluir datos sobre el autor, el título, la edición, el editor, serie, notas y números normalizados del documento.

**Requerimientos:** Condición o necesidad de un usuario para resolver un problema o alcanzar un objetivo.

**Repositorio Institucional:** Aplicación informática destinada para la gestión, intercambio y conservación de la producción científica de una institución.

**RUP:** Son las siglas de Rational Unified Process. Se trata de una metodología para describir el proceso de desarrollo de software.

**Script:** En informática, un script es un guión o conjunto de instrucciones. Permiten la automatización de tareas creando pequeñas utilidades. Es muy utilizado para la administración de sistemas UNIX. Son ejecutados por un intérprete de línea de órdenes y usualmente son archivos de texto.

**SIGB:** Aplicación informática destinada a automatizar los sistemas y entornos bibliotecarios.

**SQL:** (Structured Query Language) es un lenguaje de consulta.

**XML:** Es un metalenguaje extensible de etiquetas desarrollado por el World Wide Web Consortium (W3C).

## Anexos

## Anexo 1: Prototipo de interfaz “Seleccionar registros”.

USUARIOS BÚSQUEDA AVANZADA LISTAS ACERCA DEL SISTEMA ? Ayuda | admin | Salir

Inicio > Herramientas > Migrar registros bibliográficos > Búsqueda en el catálogo Biblioteca UCI

Número de carnet o apellido

128 resultado(s) encontrados. Ordenar por: Relevancia

Seleccionar todo |  Desmarcar todo No resaltar

Prestar  Devolver  
 Buscar en catálogo

**Menú**

- HERRAMIENTAS
- NOTICIAS
- ETIQUETAS
- CALENDARIO
- AVISOS
- DISPARADORES DE AVISO
- COMENTARIOS DEL OPAC
- PESTAÑAS DE USUARIO
- VISOR DEL LOG
- INVENTARIO/EXISTENCIAS
- PREPARAR REGISTROS MARC
- ADMINISTRAR REGISTROS MARC
- EXPORTAR REGISTROS
- MIGRAR REGISTROS
- IMPORTAR USUARIOS
- USUARIOS
- SUBIR IMÁGENES DE USUARIOS
- PLANIFICADOR DE TAREAS

<input type="checkbox"/>	1.	Propuesta de tareas para el aprendizaje desarrollador en la asignatura Problemas Sociales de la <b>Ciencia</b> y la Tecnología para la formación de Ingenieros en <b>Ciencias</b> Informáticas por Vázquez González, Jorge Luis <b>Publicación:</b> 2009 , Tesis presentada para optar el grado académico de Maestro en <b>Ciencias</b> en Docencia Superior Agraria. <b>Fecha:</b> 2009 <b>Disponibilidad:</b> No hay ejemplares disponibles:
<input type="checkbox"/>	2.	Sistema de Información de Gobierno: mercado de datos <b>ciencia</b> e innovación tecnológica por Cespedes Maceo, Dainelis <b>Publicación:</b> 2011 , Trabajo de Diploma para optar por el título de Ingeniero en <b>Ciencias</b> Informáticas <b>Fecha:</b> 2011 <b>Disponibilidad:</b> Copias disponibles: Biblioteca de la UCI (1),
<input type="checkbox"/>	3.	Crónicas del mañana : 50 años de cuentos cubanos de <b>ciencia</b> ficción por Sánchez, José Miguel <b>Publicación:</b> La Habana Editorial Letras Cubanas 2008 . 342 p. <b>Fecha:</b> 2008 <b>Disponibilidad:</b> Copias disponibles: Biblioteca de la UCI (1),
<input type="checkbox"/>	4.	Implementación del sistema de gestión y control para la evaluación del sistema de <b>ciencia</b> e innovación tecnológica de la Dirección de deportes en la UCI por Pozo Hernández, Mario Luis del <b>Publicación:</b> 2010 , Trabajo de Diploma para optar por el título de Ingeniero en <b>Ciencias</b> Informáticas. <b>Fecha:</b> 2010 <b>Disponibilidad:</b> Copias disponibles: Biblioteca de la UCI (1),
<input type="checkbox"/>	5.	Análisis, diseño e implementación del mercado de datos indicadores relacionados con la <b>ciencia</b> e innovación tecnológica para la Oficina Nacional de Estadísticas por Hechavarría, Yunió Ricardo <b>Publicación:</b> 2010 , Trabajo de Diploma para optar por el título de Ingeniero en <b>Ciencias</b> Informáticas. <b>Fecha:</b> 2010 <b>Disponibilidad:</b> Copias disponibles: Biblioteca de la UCI (1),
<input type="checkbox"/>	6.	<b>Ciencia</b> por doquier por Pérez Sanfel, Francisco H. <b>Publicación:</b> La Habana Gente Nueva 2008 . 239 p. <b>Fecha:</b> 2008 <b>Disponibilidad:</b> Copias disponibles: Biblioteca de la UCI (1),
<input type="checkbox"/>	7.	Estudios de Historia de la <b>Ciencia</b> y la Tecnología <b>Publicación:</b> La Habana ACADEMIA 1994 . 89 p. <b>Fecha:</b> 1994 <b>Disponibilidad:</b> Copias disponibles: Biblioteca de la UCI (2),
<input type="checkbox"/>	8.	Portal de noticias del Programa Nacional de <b>Ciencia</b> de Materiales. por Rodríguez Vera, Abel <b>Publicación:</b> 2008 , Trabajo de Diploma para optar por el título de Ingeniero en <b>Ciencias</b> Informáticas. <b>Fecha:</b> 2008 <b>Disponibilidad:</b> Copias

TODOS LOS DERECHOS RESERVADOS - DISEÑO PARA UNA RESOLUCIÓN DE 1024 PX DE ANCHO - 2011

## Anexo 2: Prototipo de interfaz “Exportar registro seleccionado”.

The screenshot displays a web interface with a list of records. A modal dialog is open over the list, titled "Abriendo registro-11104.dublin-core.xml". The dialog contains the following text:

Usted ha seleccionado abrir  
**registro-11104.dublin-core.xml**  
el cual es: documento XML  
de: http://10.53.9.10:3387

¿Qué debiera hacer Firefox con este archivo?

Abrir con Navegador web Firefox (predeterminado)

Guardar Archivo

Hacer esto para los archivos de este tipo de ahora en adelante.

Buttons:

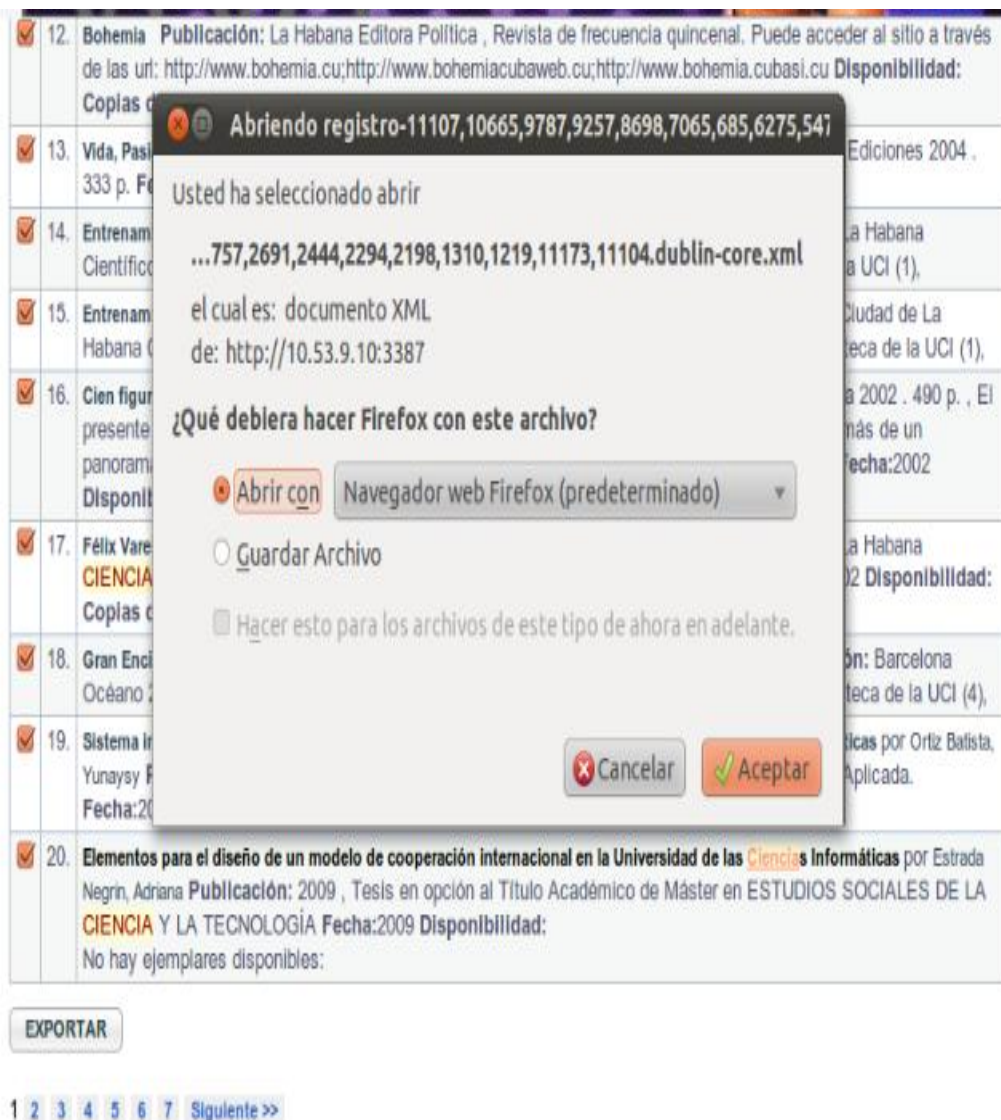
The background interface shows a table of records. Record 20 is selected (checkbox checked):

<input checked="" type="checkbox"/>	20.	Elementos para el diseño de un modelo de cooperación internacional en la Universidad de las Ciencias Informáticas por Estrada Negrin, Adriana	Publicación: 2009 , Tesis en opción al Título Académico de Máster en ESTUDIOS SOCIALES DE LA CIENCIA Y LA TECNOLOGÍA Fecha:2009 Disponibilidad: No hay ejemplares disponibles:
-------------------------------------	-----	---	--

Below the table is an "EXPORTAR" button and a pagination control showing "1 2 3 4 5 6 7 Siguinte >>".



## Anexo 3: Prototipo de interfaz “Exportar todos los registros seleccionados”.



The screenshot displays a web interface with a list of records. A modal dialog box is open, titled "Abriendo registro-11107,10665,9787,9257,8698,7065,685,6275,547". The dialog contains the following text:

Usted ha seleccionado abrir  
...757,2691,2444,2294,2198,1310,1219,11173,11104.dublin-core.xml  
el cual es: documento XML  
de: http://10.53.9.10:3387

¿Qué debiera hacer Firefox con este archivo?

Abrir con Navegador web Firefox (predeterminado)

Guardar Archivo

Hacer esto para los archivos de este tipo de ahora en adelante.

Buttons: Cancelar, Aceptar

Below the list, there is an "EXPORTAR" button and a pagination control showing "1 2 3 4 5 6 7 Siguinte >>".

**Anexo 4:** Prototipo de interfaz “Seleccionar archivos y documentos”.

The screenshot shows a web application interface for a library management system. At the top, there is a header with the logo "Sistema de Gestión Bibliotecaria" and the date "Martes, 29 de Mayo de 2012". Below the header is a navigation bar with links for "USUARIOS", "BÚSQUEDA AVANZADA", "LISTAS", and "ACERCA DEL SISTEMA", along with "Ayuda", "admin", and "Salir".

The main content area is titled "Subir archivo para importación". It features a search bar with the placeholder "Número de carnet o apellido" and a search icon. Below the search bar are radio buttons for "Prestar" (selected), "Devolver", and "Buscar en catálogo".

On the left side, there is a "Menú" section with a list of items, each with a checked checkbox: "HERRAMIENTAS", "NOTICIAS", "ETIQUETAS", "CALENDARIO", "AVISOS", "DISPARADORES DE AVISO", "COMENTARIOS DEL OPAC", "PESTAÑAS DE USUARIO", "VISOR DEL LOG", "INVENTARIO/EXISTENCIAS", "PREPARAR REGISTROS MARC", "ADMINISTRAR REGISTROS MARC", and "EXPORTAR REGISTROS".

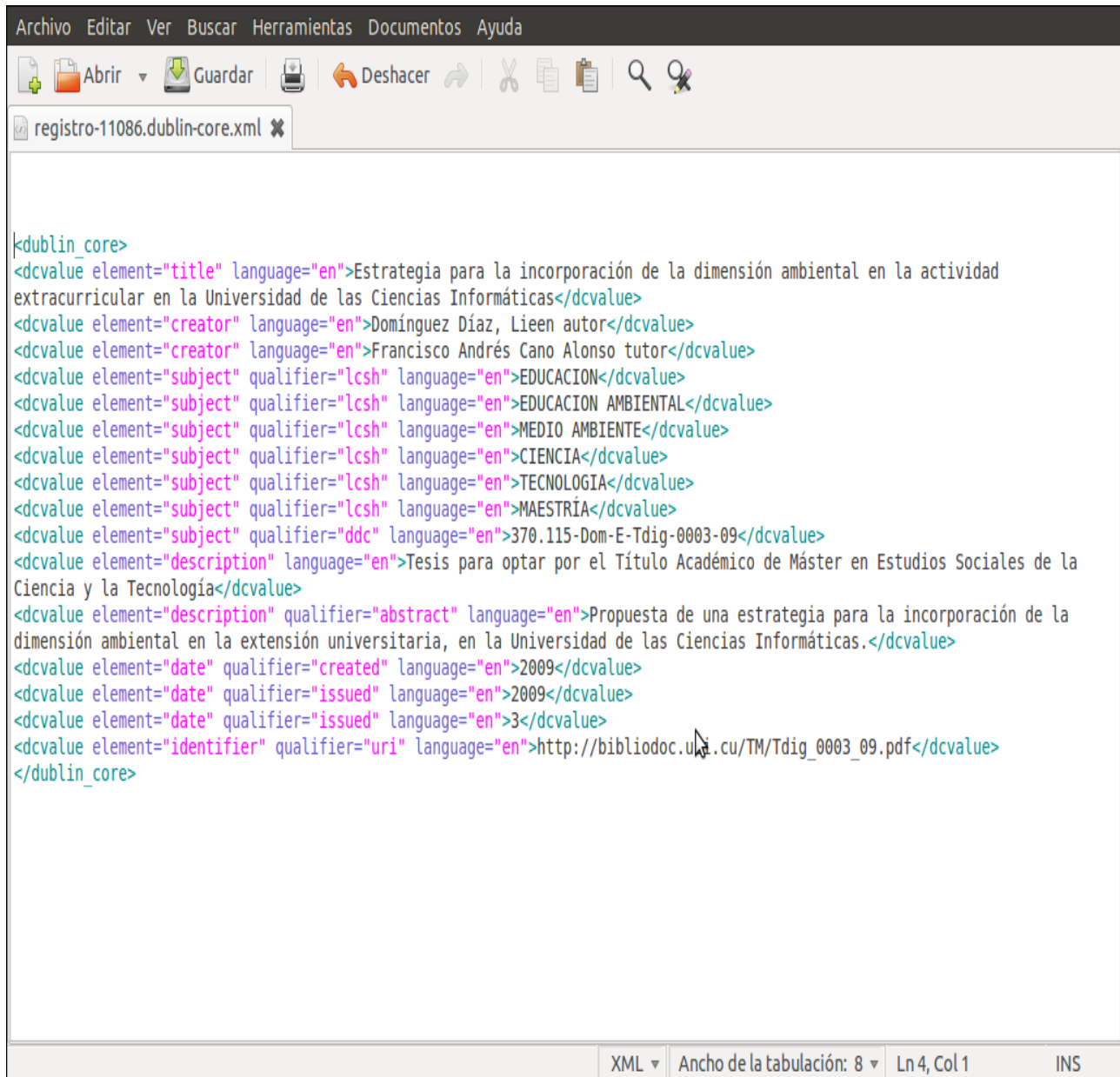
The main content area contains two sections for file selection:

- Seleccione el archivo a subir:** A text input field containing "/home/proyecto/Escritorio" and an "Examinar..." button. Below it is a note: "NOTA: Sólo se admiten archivos XML".
- Selecciones:** Two radio buttons: "archivo zip" (selected) and "archivo pdf".
- Selecciones:** A second text input field containing "/home/proyecto/Escritorio" and an "Examinar..." button. Below it is a note: "NOTA: Sólo se admiten archivos PDF y ZIP".
- Buttons:** "SUBIR" and "CANCELAR" buttons.

At the bottom of the page, there is a footer with the text: "TODOS LOS DERECHOS RESERVADOS - DISEÑADO PARA UNA RESOLUCIÓN DE 1024 PX DE ANCHO - 2011".

**Anexo 5:** Prototipo de interfaz “Generar estructura de carpeta”.

The screenshot displays the 'Sistema de Gestión Bibliotecaria' interface. At the top, the title 'Sistema de Gestión Bibliotecaria' is shown in a dark blue header, with the date 'Martes, 29 de Mayo de 2012' on the right. Below the header is a navigation bar with links for 'USUARIOS', 'BÚSQUEDA AVANZADA', 'LISTAS', and 'ACERCA DEL SISTEMA', along with 'Ayuda | admin | Salir'. A breadcrumb trail indicates the current path: 'Inicio > Herramientas > Migrar registros bibliográficos > Registros dublin core preparados para migración'. The main content area is titled 'Registros dublin core preparados para importación' and features a yellow success message: '• Los archivos han sido importados exitosamente.' Below this message are two buttons: 'GENERAR' and 'CANCELAR'. On the left side, there is a search bar labeled 'Número de carnet o apellido' and a sidebar menu with various options, including 'HERRAMIENTAS', 'NOTICIAS', 'ETIQUETAS', 'CALENDARIO', 'AVISOS', 'DISPARADORES DE AVISO', 'COMENTARIOS DEL OPAC', 'PESTAÑAS DE USUARIO', 'VISOR DEL LOG', 'INVENTARIO/EXISTENCIAS', 'PREPARAR REGISTROS MARC', 'ADMINISTRAR REGISTROS MARC', and 'EXPORTAR REGISTROS'. The footer contains the text 'TODOS LOS DERECHOS RESERVADOS - DISEÑADO PARA UNA RESOLUCIÓN DE 1024 PX DE ANCHO - 2011'.

**Anexo 6:** Prototipo de interfaz “Mostrar registros migrados”.

The image shows a screenshot of a text editor window with the title "registro-11086.dublin-core.xml". The editor displays XML code for a Dublin Core record. The code is as follows:

```
<dublin_core>
<dcvalue element="title" language="en">Estrategia para la incorporación de la dimensión ambiental en la actividad
extracurricular en la Universidad de las Ciencias Informáticas</dcvalue>
<dcvalue element="creator" language="en">Domínguez Díaz, Lien autor</dcvalue>
<dcvalue element="creator" language="en">Francisco Andrés Cano Alonso tutor</dcvalue>
<dcvalue element="subject" qualifier="lcs" language="en">EDUCACION</dcvalue>
<dcvalue element="subject" qualifier="lcs" language="en">EDUCACION AMBIENTAL</dcvalue>
<dcvalue element="subject" qualifier="lcs" language="en">MEDIO AMBIENTE</dcvalue>
<dcvalue element="subject" qualifier="lcs" language="en">CIENCIA</dcvalue>
<dcvalue element="subject" qualifier="lcs" language="en">TECNOLOGIA</dcvalue>
<dcvalue element="subject" qualifier="lcs" language="en">MAESTRÍA</dcvalue>
<dcvalue element="subject" qualifier="ddc" language="en">370.115-Dom-E-Tdig-0003-09</dcvalue>
<dcvalue element="description" language="en">Tesis para optar por el Título Académico de Máster en Estudios Sociales de la
Ciencia y la Tecnología</dcvalue>
<dcvalue element="description" qualifier="abstract" language="en">Propuesta de una estrategia para la incorporación de la
dimensión ambiental en la extensión universitaria, en la Universidad de las Ciencias Informáticas.</dcvalue>
<dcvalue element="date" qualifier="created" language="en">2009</dcvalue>
<dcvalue element="date" qualifier="issued" language="en">2009</dcvalue>
<dcvalue element="date" qualifier="issued" language="en">3</dcvalue>
<dcvalue element="identifier" qualifier="uri" language="en">http://bibliodoc.ula.ve/TM/Tdig_0003_09.pdf</dcvalue>
</dublin_core>
```

The status bar at the bottom of the editor shows "XML", "Ancho de la tabulación: 8", "Ln 4, Col 1", and "INS".