UNIVERSIDAD MAYOR DE SAN ANDRÉS
FACULTAD DE CIENCIAS PURAS Y NATURALES
CARRERA DE INFORMATICA

PROYECTO DE GRADO

"GENERACIÓN DE SERVIDORES DE METADATOS CON EL PROTOCOLO OAI-PMH

CASO: BIBLIOTECA CENTRAL DE LA UNIVERSIDAD MAYOR DE SAN ANDRÉS"

PARA OPTAR AL TITULO DE LICENCIATURA EN INFORMATICA
MENCION: INGENIERIA DE SISTEMAS INFORMATICOS

Postulante: Christian Jhonny Calle Jahuira

Tutor Metodológico: M.Sc. Aldo Ramiro Valdez Alvarado

Asesor: M.Sc. José María Tapia Baltazar

LA PAZ – BOLIVIA
2015

UNIVERSIDAD MAYOR DE SAN ANDRÉS
FACULTAD DE CIENCIAS PURAS Y NATURALES
CARRERA DE INFORMÁTICA

LA CARRERA DE INFORMÁTICA DE LA FACULTAD DE CIENCIAS PURAS Y NATURALES PERTENECIENTE A LA UNIVERSIDAD MAYOR DE SAN ANDRÉS AUTORIZA EL USO DE LA INFORMACIÓN CONTENIDA EN ESTE DOCUMENTO SI LOS PROPÓSITOS SON ESTRICTAMENTE ACADÉMICOS.

LICENCIA DE USO

El usuario está autorizado a:

a) visualizar el documento mediante el uso de un ordenador o dispositivo móvil.
b) copiar, almacenar o imprimir si ha de ser de uso exclusivamente personal y privado.
c) copiar textualmente parte(s) de su contenido mencionando la fuente y/o haciendo la referencia correspondiente respetando normas de redacción e investigación.

El usuario no puede publicar, distribuir o realizar emisión o exhibición alguna de este material, sin la autorización correspondiente.

TODOS LOS DERECHOS RESERVADOS. EL USO NO AUTORIZADO DE LOS CONTENIDOS PUBLICADOS EN ESTE SITIO DERIVARA EN EL INICIO DE ACCIONES LEGALES CONTEMPLADOS EN LA LEY DE DERECHOS DE AUTOR.

DEDICATORIA

A Dios.

Por haberme permitido cumplir esta meta, dándome su bendición día a día para llegar hasta este punto y haberme dado salud para lograr mis objetivos, además de su infinita bondad y amor.

A mi madre Asunta.

Por haberme apoyado en todo momento, por sus consejos, sus valores, por la motivación e insistencia constante que me ha permitido ser una persona de bien, pero más que nada, por su infinito amor.

A mi padre Zacarías.

Por los consejos y ejemplos de perseverancia y constancia que lo caracterizan y que me ha infundado siempre, por el valor mostrado para salir adelante y el apoyo económico brindado durante el desarrollo de la carrera y por su infinito amor.

AGRADECIMIENTO

A Dios por permitirme existir y por no dejarme desfallecer en los momentos más difíciles.

A mis padres Zacarías Calle y Asunta Pocoma de Calle por el amor y la confianza que me brindan día a día. Gracias por su constante dedicación por las muestras de cariño y por sobre todo por el apoyo y confianza que depositaron en mí.

A mis hermanos y hermanas por el apoyo y comprensión en cada momento y por apoyarme durante el momento de llevar adelante esta carrera.

A mi hijo Kristoffer Alejandro Calle Mendoza porque fue la inspiración y la fuerza que me impulso para llegar a mi meta.

A la Biblioteca Central de la Universidad Mayor de San Andrés por haberme abierto las puertas para la realización de este proyecto, así como por su entera colaboración durante su ejecución, mil gracias a la Licenciada Marilin Sanchez Rada, Magister José María Tapia Baltazar y al Magister Aldo Ramiro Valdez Alvarado por su colaboración en el desarrollo de este Proyecto de Grado.

A mis docentes, por haberme enseñado las herramientas necesarias en el desarrollo de la carrera.

Y a todos los que hicieron que este proyecto concluya.

Mil Gracias.....!

RESUMEN

El presente proyecto se desarrolló en la Biblioteca Central de la Universidad Mayor de San Andrés, denominado "Generación de Servidores de Metadatos con el Protocolo OAI-PMH[1] Caso: Biblioteca Central de la Universidad Mayor de San Andrés". El principal objetivo es integrar la tecnología de servidores de metadatos con el protocolo OAI-PMH a los sistemas de información de gestión de Biblioteca Virtual y Repositorio Digital de la universidad , para la publicación de todos los documentos académicos de investigación generados para la integración a redes internacionales de publicación de documentos de acceso libre (ApenAcces).

Este servidor será cosechado por una red internacional de servidores OAI-PMH e integrar a los proyectos OAISTER[2] de la OCLC[3] y WorlCat[4] que aproximadamente llegan a 30 millones de documentos con 1500 contribuidores de servidores OAI. La universidad será registrada e integrada a esta red internacional de instituciones educativas en el mundo.

En el desarrollo de este proyecto se ha aplicado metodologías agiles modernos, para el desarrollo del sistema aplicamos la metodología Proceso Unificado Ágil (AUP), y para el desarrollo en la ingeniería Web la metodología UWE UML (UML-Based Web Engineering) que desarrollamos en el capítulo II y III.

Para la medición de la calidad se aplicara el método de Calidad de Sitios Web (Web-site QEM), donde se realizara la medición de los diferentes módulos del sistema.

Finalmente para la medición de los costos y beneficios del sistema aplicaremos COCOMO (COnstructive COst MOdel), y presentamos los resultados tanto del personal como del costo económico para el desarrollo del presente proyecto.

Al concluir con el desarrollo y la implementación del proyecto se ha alcanzado a realizar en su totalidad los objetivos planteados, se logró integrar a la Universidad Mayor de San Andrés a la red OAISTER, WorldCAT de la OCLC redes internacionales de metadatos académicos y lo más importante es el resguardo de los derechos de autor, al ser publico el

---

[1] Open Archives Initiative - Protocol for Metadata Harvesting
[2] OAIster is a union catalog of millions of records www.oaister.org
[3] Online Computer Library Center (OCLC) www.oclc.org
[4] www.worldcat.org

documento generado por el autor en el estándar de Dublin Core especifica la fecha de creación del documento, compartiremos recursos con otros repositorios de materiales académicos. Por ejemplo se podría construir un contexto que exponga todos los recursos de una temática particular para que lo coseche un agregador temático o incluso de una tipología en particular para que los coseche un agregador de Tesis, Revistas o Libros.

Se definió un mecanismo interno de control y análisis del servidor-OAI, a partir del reporte en un contexto ad-hoc de los recursos que no cumplen con el nivel de calidad deseado por falta de metadatos obligatorios. Finalmente, con la metodología Website QEM se pudo lograr la formalización de la evaluación de la calidad del sistema. Tras la valoración obtenida aplicando la metodología se obtuvo un 83.13%, podemos asegurar que el producto elaborado es satisfactorio y comprensible para los usuarios del sistema.

INDICE DE CONTENIDO

DESCRIPCION **PAG.**

DEDICATORIA
AGRADECIMIENTO
RESUMEN

CAPITULO I
MARCO REFERENCIAL

1.1 INTRODUCCIÓN 1
1.2 ANTECEDENTES 2
1.2.1 ANTECEDENTES INSTITUCIONALES 2
1.2.2 ANTECEDENTES DE PROYECTOS SIMILARES 3
1.3 PLANTEAMIENTO DEL PROBLEMA 5
1.3.1 PROBLEMA CENTRAL 5
1.3.2 PROBLEMAS SECUNDARIOS 5
1.4 DEFINICIÓN DE OBJETIVOS 6
1.4.1 OBJETIVO GENERAL 6
1.4.2 OBJETIVOS ESPECÍFICOS 6
1.5 JUSTIFICACIÓN 7
1.5.1 JUSTIFICACIÓN ECONÓMICA 7
1.5.2 JUSTIFICACIÓN SOCIAL 7
1.5.3 JUSTIFICACIÓN TECNOLÓGICA 8
1.6 ALCANCES Y LÍMITES 8
1.6.1 ALCANCES 8
1.6.2 LÍMITES 9
1.7 APORTES 9
1.8 METODOLOGÍA 10
1.8.1 METODOLOGÍA UNIFICADA ÁGIL (AUP) 10
1.8.2 METODOLOGÍA UWE (UML WEB ENGINEERING) 11
1.8.3 LENGUAJE UNIFICADO DE MODELADO — UML 11
1.9 CALIDAD EN LA INGENIERÍA DE SOFTWARE 12
1.10 SEGURIDAD INFORMÁTICA 12
1.11 HERRAMIENTAS DEL SOFTWARE 14

CAPITULO II
MARCO TEORICO

2.1 INTRODUCCION 15
2.2 REFERENTE TEORICO 16
2.3 INICIATIVA DE ARCHIVOS ABIERTOS (OAI-PMH) 16

2.4 ESQUEMA DE METADATOS DC (DUBLIN CORE) 18
2.5 REPOSITORIOS ELECTRONICOS 22
2.6 DEFINICION DE BIBLIOTECA DIGITAL Y COMPONENTES 23
2.7 PROTOCOLOS DE INTERCAMBIO DE INFORMACION 24
2.7.1 QUE ES OAI-PMH 24
2.7.2 EL PROTOCOLO OAI PARA LA RECOLECCIÓN DE METADATOS (OAI-PMH) 26
2.7.3 EL PROTOCOLO AOI-PMH 27
2.7.4 PROTOCOLO Z3950 30
2.8 PROVEEDORES DE DATOS Y DE SERVICIO 31
2.9 CLASIFICACION DE LAS COLECCIONES DIGITALES DE LA BIBLIOTECA CENTRAL 32
2.10 TECNICAS PARA CREAR INDICES 33
2.11 LENGUAJE DE REPRESENTACION DE CONTENIDOS 35
2.11.1 HTML (HYPERTEXT MARKUP LANGUAGE) 35
2.11.2 XML (EXTENSIBLE MARKUP LANGUAGE) 36
2.12 METADATOS 36
2.13 TIPOS DE METADATOS 37
2.13.1 METADATOS DESCRIPTIVOS 37
2.13.2 METADATOS ESTRUCTURALES 37
2.13.3 METADATOS ADMINISTRATIVOS 38
2.14 FORMATO DE METADATOS DUBLIN CORE 38
2.15 ALMACENAMIENTO, RESPALDO Y PRESERVACION DIGITAL 41
2.16 METODOLOGIAS 42
2.16.1 PROCESO UNIFICADO AGIL AUP 42
2.16.1.1 DESCRIPCIÓN 42
2.16.1.2 PRINCIPIO BÁSICOS 43
2.16.1.3 PROCESO DE DESARROLLO 44
2.16.2 METODOOGIA UWE (UML-BASED WEB ENGINEERING) 46
2.16.2.1 LA METODOLOGÍA UWE 47
2.16.2.2 ESPECIFICACIÓN DE REQUERIMIENTOS 47
2.16.2.3 MODELO CONCEPTUAL 49
2.16.2.4 MODELO DE NAVEGACIÓN 50
2.16.2.5 MODELO DE PRESENTACIÓN 53
2.16.2.6 MODELADO DE TAREAS 54

CAPITULO III
MARCO APLICATIVO

3.1 INTRODUCCION 56
3.2 FASE DE INICIO 56
3.2.1 REQUERIMIENTOS DEL SISTEMA 56
3.2.2 REQUERIMIENTOS TECNOLÓGICOS 57
3.2.3 MODELO DE NEGOCIOS 57
3.2.4 MODELADO DE CASOS DE USO DE NEGOCIO 58

3.2.4.1 DESCRIPCIÓN DE CASOS DE USO 59
3.3 FASE DE ELABORACION 62
3.3.1 ESPECIFICACIÓN DE REQUERIMIENTOS 62
3.3.1.1 FUNCIONES DEL SISTEMA 62
3.3.2 MODELO DE CASOS DE USO DEL ANÁLISIS 63
3.3.3 CASOS DE USO 64
3.3.3.1 DETALLE DE CASOS DE USO EXPANDIDO 67
3.3.4 DIAGRAMA DE SECUENCIA 69
3.4 FASE DE CONSTRUCCION 73
3.4.1 DIAGRAMA DE CONTENIDO 73
3.4.2 DIAGRAMA DE NAVEGACIÓN 74
3.4.3 DIAGRAMA DE ESTADO 75
3.4.4 DIAGRAMA DE COMPONENTES 77
3.4.5 DIAGRAMA DE DESPLIEGUE 77
3.4.6 ARQUITECTURA DEL SISTEMA 78
3.4.7 DIAGRAMA DE CLASES 80
3.5 FASE DE IMPLEMENTACION 81
3.5.1 DISEÑO DE INTERFACES 81

CAPITULO IV
CALIDAD Y SEGURIDAD

4.1 METRICAS DE CALIDAD 84
4.2 METODOLOGÍA DE EVALUACIÓN DE CALIDAD DE SITIOS WEB 84
4.2.1 DEFINICIÓN DE LAS METAS DE EVALUACIÓN Y SELECCIÓN DEL PERFIL 85
4.2.1.1 DEFINICIÓN DEL DOMINIO Y ENTE DE EVALUACIÓN 85
4.2.1.2 DEFINICIÓN DE LAS METAS DE EVALUACIÓN 85
4.2.1.3 SELECCIÓN DEL PERFIL DE USUARIO 85
4.2.2 DEFINICIÓN Y ESPECIFICACIÓN DE LOS REQUERIMIENTOS DE CALIDAD 86
4.2.3 DEFINICIÓN E IMPLEMENTACIÓN DE LA EVALUACIÓN ELEMENTAL 89
4.2.3.1 CRITERIO DE EVALUACIÓN ELEMENTAL PARA ATRIBUTOS 89
4.2.3.1.1 ESPECIFICACIÓN PARA REPRESENTAR A LOS ATRIBUTOS DE CALIDAD 89
4.2 .3.1.2 IMPLEMENTACIÓN DE LAS MEDICIONES ELEMENTALES 94
4.2.4 DEFINICIÓN E IMPLEMENTACIÓN DE LA EVALUACIÓN GLOBAL 101
4.2.4.1 AGREGACIÓN DE LAS PREFERENCIAS ELEMENTALES 101
4.2.4.2 CÓMPUTO DE LAS PREFERENCIAS DE CALIDAD PARCIALES Y GLOBALES 104
4.2.5 ANÁLISIS DE RESULTADOS 109
4.3 SEGURIDAD 111
4.4 ANALISIS COSTO - BENEFICIO 122

CAPITULO IV
CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES 127
5.2 RECOMENDACIONES 128
BIBLIOGRAFIA
ANEXOS
Anexo A ARBOL DE PROBLEMAS
Anexo B MARCO LOGICO
Anexo C FORMATO DE METADATOS DUBLIN CORE
Anexo D CASOS DE USO
Anexo E DIAGRAMA DE SECUENCIA
Anexo F MODELO CONCEPTUAL
Anexo G MODELO DE NAVEGACIÓN
Anexo H MODELO DE ESTRUCTURA DE NAVEGACIÓN
Anexo I DIAGRAMA DE CONTENIDOS
Anexo J DIAGRAMA DE ESTADO
Anexo K DIAGRAMA DE COMPONENTES
Anexo L DIAGRAMA DE DESPLIEGUE
Anexo M CONTRATO DE OPERACIONES
Anexo N ARQUITECTURA DEL SISTEMA SERVIDOR-OAI
Anexo O DOCUMENTACION

INDICE DE TABLAS

DESCRIPCION **PAG.**

Tabla 2.1 Formato Dublin Core 21
Tabla 3.1 Casos de Uso Expandido: Solicitar publicación 59
Tabla 3.2 Casos de Uso de Expandido: Registrar Autor 59
Tabla 3.3 Casos de Uso Expandido: Registrar datos del documento 60
Tabla 3.4. Casos de Uso Expandido: Realizar la Selección de la Colección 60
Tabla 3.5. Casos de Uso Expandido: Consultar Duplicidad o Plagio del Titulo 61
Tabla 3.6 Caso de Uso Expandido: Catalogación del Documento 62
Tabla 3.7 Requerimientos Funcionales 63
Tabla 3.8 Descripción de actores principales 63
Tabla 3.9 Casos de uso: Registro de Solicitud de Publicación 76
Tabla 3.10 Casos de uso: Realizar la catalogación del documento en Dublín Core 68
Tabla 3.11 Contrato de Operaciones- Introducir login y password 69
Tabla 3.12 Contrato de Operaciones- Solicita Publicación 70
Tabla 3.13 Contrato de Operaciones- Introducir login y password 71
Tabla 3.14 Contrato de Operaciones- Crear nuevo registro de solicitud 71
Tabla 3.15 Contrato de Operaciones- Verifica datos 71
Tabla 3.16 Contrato de Operaciones- Introducir login y password 72
Tabla 3.17 Contrato de Operaciones- Registra el Documento 73
Tabla 3.18 Contrato de Operaciones- Realiza catalogación del documento 73
Tabla 4.1 Árbol de requerimientos de calidad 88
Tabla 4.2 Plantilla del atributo de Calidad – Mapa del sitio 90
Tabla 4.3 Plantilla del atributo de Calidad – Factibilidad FAQ 90
Tabla 4.4 Plantilla del atributo de Calidad – Soporte a lenguaje extranjero 91
Tabla 4.5 Plantilla del atributo de Calidad – Búsqueda de Documentos 91
Tabla 4.6 Plantilla del atributo de Calidad – Indicador del Camino 92
Tabla 4.7 Plantilla del atributo de Calidad – Promedio de enlace de página 92
Tabla 4.8 Plantilla del atributo de Calidad – Información de los autores por publicación 93
Tabla 4.9 Plantilla del atributo de Calidad – Enlaces Rotos 93
Tabla 4.10 Plantilla del atributo de Calidad – Paginas de Acceso Rápido 94
Tabla 4.11 Resultados de las preferencias de calidad elemental de la característica Usabilidad 96
Tabla 4.12 Resultados de las preferencias de calidad elemental de característica Funcionalidad 99
Tabla 4.13 Resultados de las preferencias de calidad elemental de las característica Confiabilidad 99

Tabla 4.14 Resultados de las preferencias de calidad elemental de la característica Eficiencia 100
Tabla 4.15 Valores parametrizados de "r" 105
Tabla 4.16 Resultados de la evaluación global 110
Tabla 4.17 Coeficiente Modelo COCOMO 122
Tabla 4.18 Ponderaciones para los 15 atributos de COCOMO Intermedio 123

INDICE DE FIGURAS

DESCRIPCION **PAG.**

Figura 2.1 Componentes de un repositorio 22
Figura 2.2 Componentes básicos de la biblioteca digital 24
Figura 2.3 Área de Servicios de Datos: ejemplo de sistema de agregación 31
Figura 2.4 Propuesta del Servidor OAI-PMH de la Biblioteca Central 32
Figura 2.5 software para indizar, buscar y recuperar información 33
Figura 2.6 Reporte de índice invertido 34
Figura 2.7 Diagrama de Casos de Usos de la Biblioteca 46
Figura 2.8 Modelo Conceptual para la Biblioteca 50
Figura 2.9 Modelo de Espacio de Navegación para la Biblioteca 51
Figura 2.10 Modelo de Estructura de Navegación para la Biblioteca 53
Figura 2.11 Prototipo de Presentación de Publicación para la Biblioteca 54
Figura 2.11 Modelo de tareas para la tarea Borrar una Publicación 55
Figura 3.1 Diagrama de Casos de uso del Negocio Actual 58
Figura 3.2 Diagrama de Casos de uso Procedo de Catalogación 61
Figura 3.3. Diagrama de Casos de uso del Sistema 64
Figura 3.4 Diagrama de Casos de Uso: Registrar solicitud de Publicación 65
Figura 3.5 Diagrama de Casos de Uso: Catalogación del Documento 66
Figura 3.6 Diagrama de Secuencia: Registro de Solicitud de Publicación 69
Figura 3.7 Diagrama de Secuencia: Registrar Autor 70
Figura 3.8 Diagrama de Secuencia: Registrar el Documento 72
Figura 3.9 Diagrama de Contenido 74
Figura 3.10 Diagrama de Navegación 75
Figura 3.11 Diagrama de Estado: Registrar Solicitud de Publicación 75
Figura 3.12 Diagrama de Estado: Registrar Autor 76
Figura 3.13 Diagrama de Estado: Catalogación del documento en Dublín Core 76
Figura 3.14 Diagrama de Componentes 77
Figura 3.15 Diagrama de despliegue 77

Figura 3.16 Arquitectura del Sistema Servidor OAI por capas 78
Figura 3.17 Arquitectura del Sistema Servidor OAI 79
Figura 3.18 Diagrama de Clases 80
Figura 3.19 Interfaz del menú principal 81
Figura 3.20 Interfaz de autenticación 82
Figura 3.21 Interfaz de resultados 82
Figura 3.22 Interfaz OAISTER con 30 millones de documentos y 1500 instituciones 83
Figura 3.23 Interfaz de resultados OAISTER 83
Figura 4.1 Estructura de agregación de preferencias parciales para la característica Usabilidad 102
Figura 4.2 Estructura de agregación de preferencias parciales para la característica Funcionalidad 102
Figura 4.3 Estructura de agregación de preferencias parciales para la característica Confiabilidad 103
Figura 4.4 Estructura de agregación de preferencias parciales para la característica Eficiencia 103
Figura 4.5 Estructura de agregación de características de alto nivel 103
Figura 4.6 Rango de aceptabilidad de preferencia de calidad 110
Figura 4.7 Compartir nuestros datos de usuario de varias páginas de acceso 112
Figura 4.8 Usuario accede por diferentes paginas al Cosechador OAI 113
Figura 4.9 Usuario recibe un envía al navegador del usuario una Cookie que sirve para identificarle y asociarle el HashMap 114
Figura 4.10 Usuario recibe un envía al navegador del usuario una Cookie que sirve para identificarle y asociarle el HashMap 115

CAPITULO I
MARCO REFERENCIAL

1.1 INTRODUCCIÓN

Este trabajo se sitúa en el área de interoperabilidad de repositorios digitales y bajo el protocolo definido por la Iniciativa de Archivos Abiertos (OAI, por sus siglas en inglés). El objetivo es construir y promover la construcción de servidores OAI-PMH[1] que comparten colecciones digitales en una red internacional de documentos académicos y se denomina federación. Esta federación está conformada por aquellas instituciones académicas y universidades dentro del marco OAI.

El presente trabajo permitirá la publicación y protección intelectual de la producción de documentos de la Universidad Mayor San Andrés así como documentos de investigación, proyectos de grado y tesis de pregrado y postgrado.

Este trabajo presenta el desarrollo de un servidor de metadatos con el protocolo OAI-PMH. Para esto, el bibliotecario solamente deberá conocer la estructura de las colecciones digitales y proporcionar la información necesaria al autor para la publicación de su documento en un servidor OAI. El servidor garantizara la fácil publicación de documentos académicos digitales capaces de ser validados y publicados ante la Iniciativa de Archivos Abiertos, y de esta forma ser parte de la creciente federación OAI como UMSA.

El protocolo OAI-PMH utilizado para la transmisión de metadatos en Internet. Se desarrollara y describirá su arquitectura basada en el modelo cliente-servidor, donde la Biblioteca Central, pondrá a disposición del público todas sus colecciones en formato digital con el formato de catalogación Dublin Core para que puedan ser recuperados por

[1] OAI-PMH (Open Archives Initiative - Protocol for Metadata Harvesting)
Protocolo para la Recolección de Metadatos - www.openarchives.org/pmh/

los cosechadores OAI de documentos académicos. La comunicación se realiza mediante el protocolo http y las respuestas están codificadas en XML.

1.2 ANTECEDENTES

1.2.1 ANTECEDENTES INSTITUCIONALES

La "Biblioteca Central", llamada así por ser el más grande de toda la universidad donde se depositan todos los documentos generados en la universidad por las defensas de pregrado, postgrado, tesis , libros y artículos de investigación científica.

Sus instalaciones están ubicadas en el primer subsuelo del Monoblock de la UMSA, cuenta con tres salas: el sector de atención al público, donde se encuentra el catalogo electrónico de referencia; una oficina para la administración y registro de solicitud de publicación, procesos técnicos y depósito del fondo bibliotecario; una sala de lectura silenciosa; y otra sala de lectura de hemeroteca, dividida en dos ambientes.

La nueva tecnología para la publicación de documentos en la Biblioteca Central, utiliza un sistema de publicación cliente servidor y que esta público en internet, pero no integra a redes internacionales académicos de distribución de servidores OAI y cosechadores de documentos académicos como OAISTER, WorldCAT de la OCLC[2] entre otros.

En nuestro país tampoco existe ningún antecedente que esté aplicando esta tecnología, citaremos antecedentes internacionales para tomar como referencia del presente proyecto.

En Santa Fe (Nuevo México, USA) se organizó una reunión, con la idea de que la interoperabilidad de estos archivos de eprints[3] era clave para aumentar su impacto entre la comunidad académica. Con ella se podrían federar varios archivos, intercambiar registros o realizar búsquedas en disciplinas relacionadas al mismo tiempo. Los participantes en esta reunión fueron especialistas en bibliotecas digitales, así como representantes de los principales archivos académicos existentes.

---

[2] OCLC (Online Computer Library Center) sin fines de lucro www.**oclc**.org

[3] **ePrint** (Pronunciado: i-Print) es la versión digital de un documento de investigación.

1.2.2 ANTECEDENTES DE PROYECTOS SIMILARES

AOISTER, WORLDCAT y OCLC han mantenido un registro de todos los repositorios académicos que han implementado el protocolo OAI-PMH. El incremento del número de repositorios en la red ha sido constante. Habría que destacar también los repositorios que se están abriendo en el área de la salud, área social, área tecnológica, bibliotecología y documentación, para el cual podemos citar los siguientes:

a) Citamos el proyecto CSIC[4] que a partir de 2012 se adopta la denominación de Red de Bibliotecas y Archivos del SCIC, con la autoría del Ministerio de Economía y Competitividad del gobierno de España para la red de Bibliotecas del Consejo Superior de Investigaciones Científicas forma parte del Sistema de Información Científica del CSIC (SIC) y está constituida por la Unidad de Recursos de Información Científica para la Investigación y por 63 bibliotecas especializadas además de 350 archivos científicos. Las bibliotecas de la Red están ubicadas en centros de investigación, propios del CSIC o que este organismo comparte con Universidades u otros organismos de investigación. La Red de Bibliotecas del CSIC también proporciona servicios bibliotecarios al personal investigador de los centros e institutos que no tienen biblioteca a través del Plan 100% Digital. La Red incluye, además, a 5 bibliotecas externas pertenecientes a fundaciones relacionadas con la investigación (Residencia de Estudiantes, Real Academia de Ciencias Exactas, Físicas y Naturales, Fundación Ortega y Gasset, Fundación García Lorca y Museu de Ciencies Naturales de la Ciudadella). La Red de Bibliotecas del CSIC es un sistema horizontal de información científica que permite la selección, organización, difusión, acceso y conservación de los recursos de información con el fin de ser un instrumento útil para la investigación que desarrolla el CSIC. La Red de Bibliotecas da soporte a la excelencia científica garantizando el acceso intelectual, físico y electrónico a la información, las ideas y los servicios para una comunidad de

[4] CSIC : Consejo Superior de Investigación Científica de España
http://bibliotecas.csic.es/servidor-oai

investigación multidisciplinar como la del CSIC mediante la utilización, en modo local o remoto, de infraestructuras y servicios de calidad.

b) También citamos el proyecto "CyberTesis"[5] de la Universidad de San Marcos, el proyecto fue creado el 25 de marzo del 2002, es un proyecto de colaboración entre varias instituciones de Perú y centros de investigación como el CNRS. En estos momentos almacena muchos documentos, la mayor parte de ellos son las tesis de la universidad. Incluye los documentos de áreas de los institutos de investigación. Difícil acceso y ponerla al alcance de cualquier usuario de la forma más sencilla posible. Es decir, trata de sacar a la luz colecciones que antes eran invisibles. Todos los recursos tienen el texto completo disponible en la red de forma que siempre se pueda llegar a los contenidos. Recoge datos de todos los archivos conocidos. En total 122 archivos con más de un millón de registros. Colaboración entre los mismos, emitir recomendaciones y crear servicios personalizados. Otros servicios también financiados por la UE son ICite y TORII ambos dentro del proyecto: Tools for Innovative Publishing in Science.

En resumen, los proveedores de servicios están proliferando y cada vez están proporcionando servicios más sofisticados. Se puede decir que existe un mercado donde los servicios pueden competir, existen hasta 10 interfaces diferentes a los datos proporcionados por arXiv.org, cada uno de ellos con unas características diferenciadoras.

[5] CyberTesis: http://cybertesis.unmsm.edu.pe/

1.3 PLANTEAMIENTO DEL PROBLEMA

1.3.1 PROBLEMA CENTRAL

En la Biblioteca Central de la Universidad Mayor de San Andrés y el resto de las instituciones educativas de nuestro país se puede observar el desconocimiento de la tecnología OAI-PMH, en es este estudio presentamos que ninguna institución boliviana está registrado como proveedor de metadatos académicos al mundo.

El 72% de nuestras instituciones educativas están migrando a bibliotecas virtuales pero sin considerar las nuevas tecnologías que ofrece el mundo globalizado, por tanto en este proyecto se plantea el problema central:

¿Cómo podemos integrar la tecnología de servidores de metadatos con el protocolo OAI-PMH a los Sistemas de Información y Gestión de Biblioteca Virtual y Digital en la Universidad Mayor de San Andrés (UMSA), para registrarla e integrarla a redes internacionales de metadatos académicos protegiendo los derechos de autor?

1.3.2 PROBLEMAS SECUNDARIOS

Otros problemas que se resolverán con este proyecto son:

- ✓ No existe una infraestructura interoperable que permita a los autores a no invertir mucho tiempo ni recursos en publicar sus documentos.
- ✓ También se pudo apreciar que algunos autores e investigadores que desean compartir con el mundo globalizado en una infraestructura que garantice y evite el plagio del documento de sus investigaciones y sus colecciones digitales están limitados de hacerlo.

- ✓ El personal bibliotecario no están familiarizado con el protocolo OAI-PMH.
- ✓ Formar parte de la Federación de la Iniciativa de Archivos Abiertos y ofrecer a la comunidad académica colecciones ordenadas de documentos.
- ✓ Generar un ambiente abierto y extensible en el cual diversas colecciones digitales de investigación coexistan y diferentes aplicaciones distribuidas las puedan acceder de manera uniforme.

1.4 DEFINICIÓN DE OBJETIVOS

1.4.1 OBJETIVO GENERAL

Como respuesta al problema central se plantea el siguiente objetivo general:

Implementar un servidor de metadatos con el protocolo OAI-PMH para generar y automatizar el proceso de construcción y publicación de un repositorio de documentos académicos en la universidad. Para esto, solamente es necesario que el bibliotecario conozca la estructura de la colección que desea compartir con el mundo, y con base en la información que proporcionará al servidor OAI, se podrá construir el código XML del servicio de publicación.

1.4.2 OBJETIVOS ESPECÍFICOS

Por otro lado los objetivos específicos dan solución a los problemas secundarios y se plantean de la siguiente manera:

- ✓ Construir una infraestructura del servidor de metadatos OAI-PMH y podrá ser validado ante la Iniciativa de Archivos Abiertos, lo que permitirá que el servidor creado por el la Biblioteca Central de la UMSA pueda formar parte de la federación OAI conformada por las colecciones digitales.
- ✓ Capacitar al personal bibliotecario el protocolo OAI-PHM.
- ✓ Registrar el servidor OAI de la Biblioteca Central de la UMSA ante la Federación de la Iniciativa de Archivos Abiertos para garantizar los derechos del autor y evitar el plagio de los documentos.

- ✓ Construir colecciones de documentos de manera ordenada por área y tipo de documento en la biblioteca central.
- ✓ Los autores solicitan el registro para la publicación de sus documentos de manera automática.
- ✓ Implementar los módulos que son parte de del servidor OAI.

1.5 JUSTIFICACIÓN

1.5.1 JUSTIFICACIÓN ECONÓMICA

La Biblioteca Central de la Universidad Mayor de San Andrés está construyendo su depósito de documentos físicos para la preservación y diseminación de los trabajos como proyectos de grado, tesis, revistas científicas, libros, entre otros, que se encuentran depositados en cantidades y que faltan espacios para el resguardo. Con la construcción de un servidor OAI se ahorra el espacio de almacenamiento y la impresión física de los documentos, de esta forma se garantiza la nueva forma de preservación y publicación usando tecnología de punta.

1.5.2 JUSTIFICACIÓN SOCIAL

El servidor de metadatos OAI-PMH de la Biblioteca Central, esta íntegramente relacionad al servicio social y de asistencia al área académica, para aquella población de estudiantes e investigadores, es por esta razón que la implementación de este sistema, será de gran utilidad para que el mismo cuente con información actualizada y oportuna a fin de brindar material de investigación certificada y garantizada en el área académica para mejorar el nivel académico y a su vez favorecer al cumplimiento de sus objetivos y metas trazadas por la Biblioteca Central.

Con el sistema serán beneficiados directamente estudiantes, docentes investigadores y población en Gral., se agilizará el acceso a información académica y científica a nivel mundial y brindar nivel internacional nuestras publicaciones y evitar el plagio y proteger los derechos del autor.

1.5.3 JUSTIFICACIÓN TECNOLÓGICA

Cuando se habla de preservación digital hay que referirse a la conservación para asegurar la accesibilidad y la recuperación de materiales digitales. También puede entenderse como preservación digital, la técnica de digitalización de documentos en soporte papel para asegurar su conservación en el tiempo y evitar el deterioro del original. En este trabajo, se justifica tecnológicamente la preservación digital de los documentos de la Biblioteca Central de acceso abierto, es decir la preservación digital dirigida a asegurar la accesibilidad y el uso a largo plazo de los objetos digitales depositados en esos repositorios y por lo tanto garantizar la accesibilidad en el futuro a estos últimos.

1.6 ALCANCES Y LÍMITES

1.6.1 ALCANCES

El servidor de metadatos OAI-PMH de la Biblioteca Central tendrá un alcance positivo para la UMSA[6] en la actualidad, por ejemplo, facilitara la generación de colecciones organizadas y compartir sus colecciones a la federación internacional de archivos abiertos. También resultara influyente para que el resto de las bibliotecas e institutos de investigación de la UMSA sean capaces de generar sus propios servidores de metadatos OAI.

La exitosa aplicación de servidores OAI para diferentes colecciones provee un excelente indicador de que en un futuro es posible que esta herramienta sea usada por muchos desarrolladores que quieran compartir sus colecciones al mundo. Entonces, el servidor-OAI contribuirá en el proceso de expansión de la comunidad OAI, y por lo tanto, será una excelente ayuda para alcanzar los objetivos trazados por la Iniciativa de Archivos Abiertos.

1.6.2 LÍMITES

El proyecto se limitara al análisis, diseño y desarrollo de los procesos que integran el manejo de documentos en función a los requerimientos y el estudio de la situación actual.

[6] UMSA: Universidad Mayor de San Andrés

También el proyecto se limitara a trabajar en la versión 2.0 del protocolo OAI-PMH, se puede plantear un serio problema sobre la aplicabilidad del Servidor OAI si se publica una nueva versión de OAI-PMH. Realmente el servidor-OAI está ligado actualmente a la versión 2.0 de OAI-PMH y precisaría adaptarse a los cambios para que pudiera soportar el proceso de actualización a las nuevas versiones del protocolo.

1.7 APORTES

Los aportes a la comunidad académica de este trabajo son:

Lo primero de un sistema de preservación digital es que la información del contenido se mantenga accesible para los usuarios por un periodo de largo tiempo. Esto significa que el sistema de preservación a instalar en el servidor de metadatos OAI-PMH de la Biblioteca Central debe responder y aportar a ser un Software Libre con las consideraciones siguientes:

- ✓ La libertad de ejecutar el programa, para cualquier propósito (libertad 0).
- ✓ La libertad de estudiar cómo trabaja el programa, y cambiarlo para que haga lo que usted quiera (libertad 1). El acceso al código fuente es una condición necesaria.
- ✓ La libertad de redistribuir copias para que pueda ayudar al prójimo (libertad 2).
- ✓ La libertad de distribuir copias de sus versiones modificadas a terceros (libertad 3). Si lo hace, puede dar a toda la comunidad una oportunidad de beneficiarse de sus cambios. El acceso al código fuente es una condición necesaria.

1.8 METODOLOGÍA

La ingeniería de software pretende utilizar los recursos computacionales de tal manera que se produzcan soluciones eficientes y eficaces a los problemas del presente trabajo., el

éxito de un proyecto de software involucra elementos como la planeación, la administración y la utilización de metodologías agiles modernos.

1.8.1 METODOLOGÍA UNIFICADA ÁGIL (AUP)

La metodología unificada ágil (AUP) es una forma simplificada del RUP desarrollada por Scott Ambler. Describe un enfoque simple del desarrollo del software usando técnicas y conceptos ágiles. Algunas técnicas usadas por AUP incluyen el desarrollo orientado a pruebas, modelado y gestión de cambios ágiles y refactorización de base de datos para mejorar la productividad.

AUP se divide en cuatro fases para el desarrollo del software:

- ✓ **Inicio**, tiene como objetivo reunir y organizar la información recogida antes de empezar a diseñar el proyecto
- ✓ **Elaboración**, permite definir la arquitectura del sistema, se realiza la especificación de los casos de uso seleccionado.
- ✓ **Construcción**, se detallan los casos de uso escenarios restantes, si es necesario modifica la descripción de la arquitectura.
- ✓ **Transición**, se detallan los casos de uso escenarios restantes, si es necesario modifica la descripción de la arquitectura.

1.8.2 METODOLOGÍA UWE (UML WEB ENGINEERING)

UWE es una metodología que permite especificar de mejor manera una aplicación Web en su proceso de creación mantiene una notación estándar basada en el uso de UML (Unified Modeling Language) para sus modelos y sus métodos, lo que facilita la transición. La metodología define claramente la construcción de cada uno de los elementos del modelo.

En su implementación se deben contemplar las siguientes etapas y modelos:

- ✓ **Análisis de requisitos**. Plasma los requisitos funcionales de la aplicación Web mediante un modelo de casos de uso.
- ✓ **Modelo de contenido**. Define, mediante un diagrama de clases, los conceptos a detalle involucrados en la aplicación.
- ✓ **Modelo de navegación**. Representa la navegación de los objetos dentro de la aplicación y un conjunto de estructuras como son índices, menús y consultas.
- ✓ **Modelo de presentación**. Representa las interfaces de usuario por medio de vistas abstractas.
- ✓ **Modelo de proceso**. Representa el aspecto que tienen las actividades que se conectan con cada clase de proceso.

1.8.3 LENGUAJE UNIFICADO DE MODELADO - UML

UML (Unified Modeling Lenguage) es una herramienta que sirve para especificar, visualizar, construir y documentar los elementos de un sistema, así como para el modelado del proceso de negocios u otros sistemas, UML reúne una colección de las mejores prácticas en la ingeniería que han sido utilizadas con éxito para modelar sistemas grandes y complejos.

UML es una herramienta que cumple la función de capturar la idea de un sistema para comunicarla posterior mente a quien forma parte del proceso de desarrollo; trabaja con un conjunta de símbolos y diagramas. Cada diagrama tiene un fin distinto en el proceso de desarrollo[7].

1.9 CALIDAD EN LA INGENIERÍA DE SOFTWARE

Para la calidad en la ingeniería del software de Sitios Web se usara el método (Web-site QEM), es un grupo de características que representa la efectividad y la eficiencia de un sistema informático. Se tiene como principales puntos que se toma en cuenta en la calidad del software[8].

---

[7] [LARMAN-1999]
[8] [150-9126]

- ✓ **Usabilidad**, si bien es cierto que la usabilidad es un factor muy importante en la calidad del software también lo es el hecho es necesarios considerar otros factores como los que se mencionan en esta sección puesto de que nada sirve un software que funcione correctamente y confiablemente.
- ✓ **Confiabilidad**, el término es necesario, será separado en varios elementos que permiten darle al software el matiz de fiable.
- ✓ **Funcionalidad**, un conjunto de atributos que se relacionan con la existencia de un conjunto de funciones y sus propiedades específicas. Las funciones son aquellas que tiene como objetivo satisfacer las necesidades implícitas o explícitas del sistema.
- ✓ **Eficiencia**, conjunto de atributos relacionados con la relación entre el nivel de desempeño del software y la cantidad de recursos necesitados bajo condiciones y establecidas.

1.10 SEGURIDAD INFORMÁTICA

En el sistema la seguridad es muy importante ya que los documentos y los datos que se manejan comprometen a la institución y al autor y que son únicamente editables y actualizables por el personal bibliotecario del área de procesos técnicos de la biblioteca.

Las estrategias que se pueden aplicar para proteger los documentos y los datos son:

a) **Replicación**. Es la estrategia más básica y es necesario que las réplicas sean más de una. Algunos archivos de información digital hacen un número fijo de réplicas en servidores distribuidos geográficamente. Además cada réplica puede ser que tenga copias de seguridad.

b) **Migración.** En sus dos aspectos de cambio de soporte y de cambio de formato representa una de las estrategias para prevenir la obsolescencia del formato y del software. A veces se hace la migración en el momento de la catalogación y entonces se habla de "normalización", decidiendo si se conservará en paralelo el formato original o únicamente el objeto normalizado.

c) **Transparencia**. Esta cualidad es esencial para que el usuario pueda confiar en un sistema ya que los resultados son siempre futuros. Es de resaltar la importancia de la utilización de código abierto para asegurar el funcionamiento del sistema a largo plazo. Es remarcable la necesidad de compartir la información sobre los casos de pérdida y las mejores prácticas, tomando como modelo y ejemplo el sistema establecido por la NASA para informar de las incidencias en la navegación aérea.

d) **Diversidad**. Los sistemas que no admiten la diversidad son mucho más vulnerables que aquellos que lo hacen. Por lo tanto un sistema de preservación debería de conllevar diversidad.

e) **Auditoría.** Es necesario establecer mecanismos de control que verifiquen en diferentes fases del proceso (Catalogación, Archivo, Difusión) y mediante técnicas diversas que la información se mantiene inalterada. Las comprobaciones regulares para evitar la acumulación de errores latentes dan confianza a los que financian el sistema de preservación digital.

f) **Contención de los costos**. Uno de los peligros más grandes para garantizar la puesta en marcha y el mantenimiento de un sistema de preservación es el costo porque pocas instituciones u organizaciones pueden dedicar financiamiento a usos futuros y por la dificultad en mostrar los resultados de las inversiones realizadas.

Se pueden entender como seguridad a una característica de cualquier sistema, las amenazas como riesgos no se pueden eliminar, se las administran para que no afecten de gran manera al sistema. El sistema cuenta con cuatro características para su mejor seguridad.

- ✓ Integridad
- ✓ Confidencialidad
- ✓ Disponibilidad
- ✓ Extensibilidad

1.11 HERRAMIENTAS DEL SOFTWARE

Para la implementación y diseño del software se utiliza las siguientes herramientas:

- Sistema operativo Linux Debian 8.1 o superior
- Lenguaje de Programación Java OpenJDK 8
- Motor de Base de Datos PostgreSQL 9.4
- Tomcat
- Spring
- XML
- Apache ANT

Estas herramientas presentan todas las funciones tanto de consulta de base de datos como la programación del sistema ya que existen versiones diferentes y las últimas versiones son más completas y estas cuentan con nuevas librerías.

CAPITULO II

MARCO TEORICO

2.1 INTRODUCCION

El trabajo que presentamos en este proyecto de grado implementaremos el protocolo OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting), en la Biblioteca Central de la Universidad Mayor de San Andres y presentaremos a nuestro país el nuevo protocolo para la transmisión de contenidos digitales academicos en internet. Paradójicamente el interés entre los investigadores y profesionales de nuestro país parece escaso cuando no nulo.

La mayoría de las bibliotecas de nuestro país no aplican esta tecnología por el desconocimiento del mismo, en un estudio de arte realizado en las diferentes universidades de nuestro país no aplican las universidades esta tecnología.

El planteamiento esencial de los servidores OAI-PMH y los archivos abiertos es permitir el acceso a los materiales en la web por medio de repositorios que, interoperando unos con otros, permitan el intercambio de metadatos, su almacenamiento y publicación. Surgió de la comunidad de e-print (ediciones preliminares electrónicas), la necesidad de una solución de interoperabilidad de bajo nivel para el acceso entre repositorios heterogéneos dio lugar a la creación de la Open Archives Initiative (OAI). Como se dice en la misión de la OAI " La Open Archives Initiative desarrolla y promueve normas de interoperabilidad que tienen por objeto facilitar la difusión eficaz de los documentos y su contenidos."

Muchas comunidades se están empezando a beneficiar, o potencialmente podrían beneficiarse, de la estrategia de archivos abiertos. Internet y la gran cantidad de documentos en formato digital han ensanchado el número de usuarios potenciales de

muchas bibliotecas y centros de información. Los documentos pueden ser más accesibles y además utilizados con distintos objetivos de los que motivaron inicialmente la creación de los repositorios. Por otra parte, la posibilidad de acceder a múltiples repositorios permite la construcción de nuevos tipos de servicios que puedan atender mejor las necesidades de los usuarios. Un estímulo adicional es la posibilidad de reducción de costes, inherente a los nuevos modelos en el proceso de comunicación académica que se apoyan en esta estrategia.

2.2 REFERENTE TEORICO

El referente teórico de esta investigación, parte de dos necesidades que se crearon en el ámbito mundial de los documentos digitales, con respecto a la forma de compartir conocimientos entre los diferentes grupos de investigación que existen en las diferentes instituciones académicas, independientes del sitio geográfico donde estas se encuentren. Es por ello que se han desarrollado metodologías, esquemas, protocolos, organizaciones y grupos de investigación alrededor de este tema, con tal fuerza y magnitud que organizaciones como la UNESCO ha sido quien más apoya este tipo de iniciativas.

El propósito fundamental con el cual se presentan los siguientes elementos es proporcionar un conocimiento profundo y actualizado del estado del conocimiento y del pensamiento sobre la manera como se comparte la información hoy por hoy en el mundo.

2.3 INICIATIVA DE ARCHIVOS ABIERTOS (OAI-PMH)

La Open Archive Initiative (OAI) se creó con la misión de desarrollar y promover estándares de interoperabilidad para facilitar la difusión eficiente de contenidos en internet. Surgió como un esfuerzo para mejorar el acceso a archivos de publicaciones electrónicas (eprints) en definitiva, para incrementar la disponibilidad de las publicaciones científicas.

Los trabajos iniciales se centraron en el desarrollo de marcos de interoperabilidad para la federación de archivos de eprints, pronto apareció evidente que dichos marcos (permitir el intercambio de múltiples formatos bibliográficos entre distintas maquinas utilizando un protocolo común) tenían aplicaciones más allá de esta comunidad. Por ello se adoptó un objetivo mucho más amplio: abrir el acceso a un rango de materiales digitales.

Por lo tanto, la OAI no es solamente un proyecto centrado en publicaciones científicas, sino en la comunicación de metadatos en publicaciones científicas, sino en la comunicación de metadatos sobre cualquier material almacenado en soporte electrónico. No hay nada en el protocolo que impida a los implantadores transmitir el contenido propiamente dicho de esos materiales. No obstante esto no es el objeto principal de OAI-PMH.

Los metadatos a transmitir vía OAI-PMH deberán codificarse en Dublin Core sin calificar con objeto de minimizar los problemas derivados de las conversaciones entre múltiples formatos. Aunque se está investigando la creación de servicios tales como una interfaz de búsqueda a través de formatos heterogéneos de metadatos, una solución menos complicada y por lo tanto más fácil de implementar es requerir a los implementadores convertir sus datos a un formato común. Los quince elementos del Dublin Core han evolucionado a lo largo de los pasados años como estándar de facto para los metadatos simples y multidisciplinarias.

La OAI no define o prescribe ningún esquema para la gestión de derechos Los temas relacionados con restricciones en el acceso y gestión de la propiedad intelectual son la responsabilidad de los proveedores de datos.

2.4 ESQUEMA DE METADATOS DC (DUBLIN CORE)

El esquema de metadatos Dublin Core es un modelo de metadatos que cuenta con quince (15) definiciones semánticas descriptivas que pretenden transmitir un significado semántico a las mismas, elaborado y auspiciado por la DCMI (Dublin Core Metadata Initiative), una organización dedicada a fomentar la adopción extensa de los estándares interoperables de los metadatos y a promover el desarrollo de los vocabularios especializados de metadatos para describir recursos y permitir sistemas más inteligentes del descubrimiento del recurso. El estándar de metadatos Dublin Core es un simple pero eficaz conjunto de elementos para describir una amplia gama de recursos de red. La norma del Dublin Core conlleva dos niveles: Simple y Cualificado. El Dublin Core Simple conlleva quince elementos; el Dublin Core Cualificado conlleva un elemento adicional, la audiencia, así como un grupo de elementos de matización (denominados por ello, cualificadores) que refinan la semántica de los elementos de tal forma que pueden ser útiles para la recuperación/ localización de recursos en Internet. La semántica del Dublin Core se ha establecido por un grupo internacional e interdisciplinario de profesionales de la biblioteconomía, la informática, la codificación textual, la comunidad museística, y otros campos teórico-prácticos relacionados[9].

Cada uno de los elementos del conjunto básico de Dublin Core es opcional y puede repetirse. La mayor parte de los elementos tienen asimismo un conjunto limitado de cualificadores o refinamientos, atributos que pueden usarse para matizar más (no para extender) el significado de un elemento. La Iniciativa de Metadatos Dublin Core (DCMI) ha establecido formas normalizadas para matizar los elementos y promover el uso de esquemas de codificación y vocabulario. El conjunto completo de elementos y

[9] (Hillmann, Diane. 2003)

matizaciones de elementos de acuerdo a las "buenas prácticas" de la DCMI está disponible, con un registro formal en proceso[10].

El conjunto de elementos de metadatos Dublin Core es un estándar para la descripción de recursos entre distintos dominios informativos. Aquí, se entiende un recurso de información como "algo que tiene identidad". Esta es una definición utilizada en la RFC 2396, "Identificadores Uniformes de Recursos (URI): Sintaxis Genérica", realizada por Tim Berners-Lee y otros. No existen restricciones fundamentales para los tipos de recursos a los que se les pueden asignar metadatos Dublin Core. (Dublin Core Metadata Initiative. 2003).

Dublin Core presenta dos niveles para la clasificación de sus elementos; el nivel simple y el avanzado. El nivel sencillo de Dublin Core abarca quince (15) elementos: Título, Creador, Materia, Descripción, Editor, Colaborador, Fecha, Tipo, Formato, Identificador, Fuente, Lengua, Relación, Cobertura, Derechos. Para ver el detalle de los 15 elementos ver el anexo A. “Dublin Core Metadata Element Set”.

Mientras que el nivel avanzado aporta otros tres elementos (Audiencia, Origen y Titular de derechos) además de un grupo de elementos refinados que se ocupan de la semántica de los elementos de manera útil para el descubrimiento de recursos. Para ver con más detalle cada uno de los 15 elementos y sus cualificadores, ver la tabla 2.1 en el cual se presenta el esquema de Metadatos Dublin Core para Sistemas Servidor OAI, que están recogidos en DCMI Metadata Terms (http://dublincore.org/documents/dcmi-terms/):

[10] (Hillmann, Diane. 2003)

ETIQUETA DEL ELEMENTO DUBLIN CORE http://dublincore.org/documents/dcmi-terms/	DESCRIPCIÓN
DC. Title	**Título: El nombre dado a un recurso.** Típicamente, un título es el nombre formal por el que es conocido el recurso.
DC. Creator	**Autor: La entidad primariamente responsable de la creación del contenido intelectual del recurso.** Entre los ejemplos de un creador se incluyen una persona, una organización o un servicio. Típicamente, el nombre del creador podría usarse para indicar la entidad.
DC. Subject	**Materias y palabras clave: El tema del contenido del recurso.** Un tema será expresado como palabras clave, frases clave o códigos de clasificación que describan el tema de un recurso. Se recomienda seleccionar un valor de un vocabulario controlado o un esquema de clasificación formal.
DC. Description	**Descripción: La descripción del contenido del recurso.** La descripción puede incluir, pero no se limita a: un resumen, tabla de contenidos, referencia a una representación gráfica de contenido o una descripción de texto libre del contenido.
DC. Publisher	**Editor: La entidad responsable de hacer que el recurso se encuentre disponible.** Ejemplos de editores son una persona, una organización o un servicio. Típicamente, el nombre de un editor podría usarse para indicar la entidad.
DC. Contributor	**Colaborador. La entidad responsable de hacer colaboraciones al contenido del recurso.** Ejemplos de colaboradores son una persona, una organización o un servicio. Típicamente, el nombre del colaborador podría usarse para indicar la entidad.
DC. Date	**Fecha: Una fecha asociada con un evento en el ciclo de vida del recurso.** Típicamente, la fecha será asociada con la creación o disponibilidad del recurso. Se recomienda utilizar un valor de datos codificado definido en el documento "Date and Time Formats",http://www.w3.org/TR/NOTE-datetime que sigue la norma ISO 8601 que sigue el formato YYYY-MM-DD.
DC. Type	**Tipo: la naturaleza o categoría del contenido del recurso.** El tipo incluye términos que describen las categorías generales, funciones, géneros o niveles de agregación del contenido. Se recomienda seleccionar un valor de un vocabulario controlado (por ejemplo, el *DCMI Vocabulary* -DCMITYPE- http://dublincore.org/documents/dcmi-type-vocabulary/). Para describir la manifestación física o digital del recurso, se usa el elemento Formato.
DC. Format	**Formato: la manifestación física o digital del recurso.** El formato puede incluir el tipo de media o dimensiones del recurso. Podría usarse para determinar el *software*, *hardware* u otro equipamiento necesario para ejecutar u operar con el recurso. Ejemplos de las dimensiones son el tamaño y la duración. Se recomienda seleccionar un valor de un vocabulario controlado (por ejemplo, la lista de Internet Media Types (MIME) que define los formatos de medios de ordenador).

DC. Identifier	**Identificación: Una referencia no ambigua para el recurso dentro de un contexto dado.** Se recomienda identificar el recurso por medio de una cadena de números de conformidad con un sistema de identificación formal, tal como un URI (que incluye el Uniform Resource Locator -URL, el Digital Object Identifier (DOI) y el International Standard Book Number (ISBN).
DC. Source	**Fuente**: **Una referencia a un recurso del cual se deriva el recurso actual.** El recurso actual puede derivarse, en todo o en parte, de un recurso fuente. Se recomienda referenciar el recurso por medio de una cadena o número de conformidad con un sistema formal de identificación.
DC. Language	**Lengua**: **La lengua del contenido intelectual del recurso.** Se recomienda usar RFC 3066 http://www.ietf.org/rfc/rfc3066.txt en conjunción con la ISO 639 [ISO639] http://www.loc.gov/standards/iso639-2/, que define las etiquetas de dos y tres letras primarias para lenguaje, con subetiquetas opcionales. Ejemplo: "en" u "eng" para Inglés, "akk" para Acadio, y "en-GB" para inglés usado en Reino Unido.
DC. Relation	**Relación**: **Una referencia a un recurso relacionado.** Se recomienda referenciar el recurso por medio de una cadena de números de acuerdo con un sistemas de identificación formal.
DC. Coverage	**Cobertura: La extensión o ámbito del contenido del recurso.** La cobertura incluiría la localización espacial (un nombre de lugar o coordenadas geográficas), el período temporal (una etiqueta del período, fecha o rango de datos) o jurisdicción (tal como el nombre de una entidad administrativa). Se recomienda seleccionar un valor de un vocabulario controlado (por ejemplo, del Thesaurus of Geographic Names (TGN) y que, donde sea apropiado, se usen preferentemente los nombres de lugares o períodos de tiempo antes que los identificadores numéricos tales como un conjunto de coordenadas o rangos de datos.
DC. Rights	**Derechos**: **La información sobre los derechos de propiedad y sobre el recurso.** Este elemento podrá contener un estamento de gestión de derechos para el recurso, o referencia a un servicio que provea tal información. La información sobre derechos a menudo corresponde a los derechos de propiedad intelectual, copyright y otros derechos de propiedad.

Tabla 2.1 Formato Dublin Core
Fuente: NCSA (National Center for Supercomputing Applications)
OCLC (On Line Library Computer Center)

2.5 REPOSITORIOS ELECTRONICOS

En el sentido más amplio, los repositorios electrónicos son sitios activos y colaborativos, en los que se depositan, por diferentes personas, diversos tipos de objetos electrónicos, usualmente relacionadas por una temática o por una comunidad, además, tecnológicamente deben facilitar la interoperabilidad para el intercambio de información con otros repositorios o con otras aplicaciones usando tecnologías como el Z3959, OAI-PMH, etc.

En la figura 2.1 presentamos la arquitectura de un repositorio digital y la tecnología requerida con los estándares y formatos de catalogación.

Figura 2.1 Componentes de un repositorio

Fuente: Elaboración propia

2.6 DEFINICION DE BIBLIOTECA DIGITAL Y COMPONENTES

El término biblioteca digital posee otros sinónimos como: biblioteca virtual, biblioteca electrónica. Existen varias definiciones de biblioteca digital, pero la definición está fuertemente relacionada con la forma en que está representado el contenido, es decir una forma digital.

De acuerdo con la Digital Library Federation, las bibliotecas digitales son organizaciones que proveen recursos, incluyendo personal especializado, para seleccionar, estructurar, ofrecer acceso intelectual, interpretar distribuir, preservar la integridad y asegurar la persistencia en el tiempo de las colecciones digitales en la medida en que estas estén de forma rápida y económica, disponible para el uso de una o varias comunidades de usuarios.

La definición de biblioteca digital para el presente proyecto de grado es:

La biblioteca digital (BD) plenamente desarrollada es una colección organizada, con servicios asociados, en la que la información está almacenada en formato digital y es accesible en un sistema de información virtual.

La BDU (Biblioteca Digital Universitaria) consta de tres componentes básicos (Figura 2.2):

1. Tecnologías de la Información y la Comunicación necesarias para la BDU.
2. Contenidos digitales organizados.
3. Servicios digitales de información.

Tecnologías de la Información y la Comunicación necesarias para la BDU.

Figura 2.2 Componentes básicos de la biblioteca digital

Fuente: Elaboración propia

Cada una de estas partes tendrá características propias del momento tecnológico actual, pues no hay que olvidar que la biblioteca en el medio digital experimenta cambios provocados por los vertiginosos avances en las Tecnologías de la Información y la Comunicación. Esto provoca que nos movamos constantemente de una imagen de biblioteca a otra, o que el modelo prevaleciente sea modificado debido a las características de las nuevas herramientas. Así, a esta biblioteca se le adjudican tantos nombres como usos tecnológicos que van apareciendo.

2.7 PROTOCOLOS DE INTERCAMBIO DE INFORMACION

2.7.1 QUE ES OAI-PMH

En primer lugar hemos de matizar varios puntos respecto a su nombre. El término archivo refleja los orígenes de la iniciativa en el seno de las comunidades de eprints donde es sinónimo de depósito de documentos científicos a texto completo. No tiene nada que ver aquí con el concepto tradicional de archivo con connotaciones de preservación y

conservación. Se utiliza por lo tanto con un sentido mucho más amplio, como un depósito para almacenar cualquier tipo de información. El término abierto se refiere al punto de vista de la arquitectura del sistema. Se tratan de definir interfaces que faciliten la disponibilidad de contenidos procedentes de una variedad de proveedores.

La Open Archives Initiative (OAI) se creó con la misión de desarrollar y promover estándares de interoperabilidad para facilitar la difusión eficiente de contenidos en Internet. Surgió como un esfuerzo para mejorar el acceso a archivos de publicaciones electrónicas (eprints), en definitiva, para incrementar la disponibilidad de las publicaciones científicas. Los trabajos iniciales se centraron en el desarrollo de marcos de interoperabilidad para la federación de archivos de eprints, pronto apareció evidente que dichos marcos (permitir el intercambio de múltiples formatos bibliográficos entre distintas máquinas utilizando un protocolo común) tenían aplicaciones más allá de esta comunidad. Por ello se adoptó un objetivo mucho más amplio: abrir el acceso a un rango de materiales digitales.

Los metadatos a transmitir vía OAI-PMH deberán codificarse en Dublin Core sin calificar con objeto de minimizar los problemas derivados de las conversiones entre múltiples formatos. Aunque se está investigando la creación de servicios tales como una interfaz de búsqueda a través de formatos heterogéneos de metadatos, una solución menos complicada y por lo tanto más fácil de implementar es requerir a los implementadores convertir sus datos a un formato común. Los quince elementos del Dublin Core han evolucionado a lo largo de los pasados años como el estándar de facto para los metadatos simples y multidisciplinares.

La OAI no define o prescribe ningún esquema para la gestión de derechos. Los temas relacionados con restricciones en el acceso y gestión de la propiedad intelectual son la responsabilidad de los proveedores de datos.

2.7.2 EL PROTOCOLO OAI PARA LA RECOLECCIÓN DE METADATOS (OAI-PMH)

El Protocolo OAI para la Recolección de Metadatos (OAI-PMH), define un mecanismo para la recolección de registros que contienen los metadatos de los repositorios. El OAI-PMH ofrece a los Proveedores de Datos una opción técnica sencilla para poner sus metadatos a disposición de servicios basados en los estándares abiertos HTTP (Hypertext Transport Protocol) y XML (Extensible Markup Language). Los metadatos que son recolectados pueden estar en cualquier formato establecido por una comunidad (o por cualquier conjunto específico de Proveedores de Datos y Proveedores de Servicios), con independencia de que hayan establecido el Dublin Core no cualificado para proporcionar un nivel básico de interoperabilidad. Por lo tanto, los metadatos de distintas fuentes pueden ser reunidos en una base de datos, y se pueden ofrecer servicios sobre la base de esta recolección centralizada, o "agregación" de datos. El vínculo entre estos metadatos y el contenido no está definido por el protocolo OAI. Es importante darse cuenta de que OAI-PMH no proporciona una búsqueda a través de estos datos, simplemente permite reunir los datos en un sitio. Para ofrecer un servicio, el método de la recolección debe combinarse con otros mecanismos.

El uso del protocolo es muy prometedor dentro de la estrategia de los archivos abiertos. El beneficio más conocido es servir de sustento de un nuevo modelo de comunicación académica. Tal vez los objetivos más sencillos de alcanzar son: hacer aflorar a la superficie "los recursos ocultos" y la interoperabilidad a bajo coste. Aunque el OAI-PMH es técnicamente muy simple, la construcción coherente de servicios que satisfagan las necesidades de los usuarios sigue siendo compleja. El protocolo OAI-PMH podría convertirse en parte de la infraestructura de la Web, tanto como el protocolo HTTP lo es ahora, si gracias a la combinación de su relativa sencillez y el éxito demostrado en los

primeros servicios implementados, condujese a una aceptación amplia en las instituciones de investigación, editores, e "instituciones de la memoria".

2.7.3 EL PROTOCOLO AOI-PMH

Los que participaron en el desarrollo de OAI.PMH tomaron una decisión clave en cuanto a la arquitectura del protocolo. Adoptaron un modelo que rechazaba la búsqueda distribuida (como hace el Z39.50) a favor de simplemente tener servidores que proporcionen metadatos, sujetos sólo a criterios de alcance bastante simples, tales como por ejemplo proporcionar todos los registros añadidos o cambiados desde una fecha específica.

Vamos a entrar aquí en una descripción técnica del protocolo, pero básicamente OAI-PMH utiliza transacciones HTTP para emitir preguntas y obtener respuestas entre un servidor o archivo y un cliente o servicio recolector de metadatos. El segundo puede pedir al primero que le envíe metadatos según determinados criterios como por ejemplo la fecha de creación de los datos. En respuesta el primero devuelve un conjunto de registros en formato XML, incluyendo identificadores (URLs por ejemplo) de los objetos descritos en cada registro.

Las peticiones se emiten utilizando los métodos GET o POST del protocolo HTTP y constan de una lista de opciones con la forma de pares del tipo: clave=valor. Existen seis peticiones que un cliente puede realizar a un servidor:

a) **GetRecord**. Utilizado para recuperar un registro concreto. Necesita dos argumentos: identificador del registro pedido y especificación del formato bibliográfico en que se debe devolver.

b) **Identify**. Utilizado para recuperar información sobre el servidor: nombre, versión del protocolo que utiliza, dirección del administrador, etc.

c) **ListIdentifiers**. Recupera los encabezamientos de los registros, en lugar de los registros completos. Permite argumentos como el rango de fechas entre los que queremos recuperar los datos.

d) **ListRecords**. Igual que el anterior pero recupera los registros completos.

e) **ListSets**. Recupera un conjunto de registros. Estos conjuntos son creados opcionalmente por el servidor para facilitar una recuperación selectiva de los registros. Sería una clasificación de los contenidos según diferentes entradas. Un cliente puede pedir que se recuperen solo los registros pertenecientes a una determinada clase. Los conjuntos pueden ser simples listas o estructuras jerárquicas.

f) **ListMetadataFormats.** Devuelve la lista de formatos bibliográficos que utiliza el servidor.

El protocolo soporta múltiples formatos para expresar los metadatos, no obstante requiere que todos los servidores ofrezcan los registros utilizando Dublin Core no calificado, codificado en XML. Además de éste formato cada servidor es libre de ofrecer los registros en otro/s formatos adicionales (MARC por ejemplo). Un cliente puede pedir que los registros se le sirvan en cualquiera de los formatos soportados por el servidor. La idea subyacente aquí es que en el futuro las diferentes comunidades que utilicen el protocolo definan sus propios formatos que sean más ricos y más precisos que el Dublin Core. Por ejemplo la comunidad de archivos de eprints está trabajando en un formato denominado AMF [11]que sea capaz de describir todos los elementos que intervienen en el proceso de comunicación científica: documentos, autores, instituciones y canales de distribución de documentos.

[11]AMF (Academic Metadata Format)

Las respuestas del servidor estarán formateadas según el protocolo HTTP con los adecuados encabezamientos. Serán documentos XML correctos que se podrán validar contra el esquema definido por el protocolo OAI-PMH versión 2.0.

Petición y respuesta seria:

Petición:
http://dc.umsa.bo/OAI-script?
verb=GetRecord&identifier=oai:arXiv:hep-th/9901001&metadataPrefix=oai_dc
Respuesta:


```
<?xml version="1.0" encoding="UTF-8" ?>
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
        xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
        http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
 <responseDate>2015-05-01T19:20:30Z</responseDate>
 <request verb="GetRecord" identifier="oai:arXiv:hep-th/9901001"
         metadataPrefix="oai_dc">http://dc.umsa.bo/OAI-script</request>
 <GetRecord>
   <record>
     <header>
            <identifier>oai:arXiv:cs/0112017</identifier>
            <datestamp>2015-12-14</datestamp>
            <setSpec>cs</setSpec>
            <setSpec>matematicas</setSpec>
    </header>
<metadata>
     <oai_dc:dc
        xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
        xmlns:dc="http://purl.org/dc/elements/1.1/"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
        xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
        http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
       <dc:title> Uso estructural de metadatos para localizar experiencia de contenidos
digitales
       </dc:title>
       <dc:creator>Calle, Christian</dc:creator>
       <dc:subject>Biblioteca Digital Universitaria</dc:subject>
       <dc:description> Con la creciente sofisticación técnica de
tanto los consumidores y los proveedores de información, no es
```

creciente demanda de las experiencias más significativas de la tecnología digital
información. Se presenta un marco que separa digitales
objetar la experiencia, o de la prestación, desde el almacenamiento objeto digital
y la manipulación, por lo que la representación se puede adaptar a
comunidades particulares de los usuarios. </dc:description>
<dc:description>Comentario: 23 paginas, incluye 2 apendices,
8 figuras</dc:description>
<dc:date>2015-11-14</dc:date>
</oai_dc:dc>
</metadata>
</record>
</GetRecord>
</OAI-PMH>

Aspectos que no trata el protocolo son por ejemplo cuestiones de gestión o autorización para el acceso de los clientes. El servidor deberá recurrir a métodos externos si desea limitar los clientes a los que sirva información. En relación con este punto está la utilización que los clientes hagan de los datos. También queda fuera del protocolo. Finalmente, tampoco trata el tema de cómo los clientes pueden localizar aquellos servidores que contengan los datos que necesitan.

2.7.4 PROTOCOLO Z3950

Z39.50 es una norma establecida para consultar catálogos de bibliotecas a través de internet. Se basa en el modelo cliente-servidor. El cliente se conecta, busca y extrae información de un servidor.

Hoy en día es quizá la norma más importante en el mundo de las bibliotecas, la información en el servidor es agrupada en bases de datos. Este protocolo es mantenido por la Biblioteca del Congreso de los Estados Unidos ANSI/NISO Z39.50, ISO23950. Los catálogos se presentan generalmente en formato MARC21, En ocasiones estos se pueden presentar en otros formatos como XML.

2.8 PROVEEDORES DE DATOS Y DE SERVICIO

De la sección anterior se desprende que la arquitectura de OAI-PMH se basa en clientes y servidores. Los primeros son los archivos que proporcionan la información y los segundos son los cosechadores (recolectores) o servicios que toman los datos, con el objetivo de incorporarlos algún valor añadido y presentarlos a los usuarios finales.

OAI-PMH separa claramente los proveedores de datos (en sentido general bibliotecas o archivos) de los proveedores de servicios (p.ej. Agregadores), aunque no hay impedimento conceptual en la coexistencia de roles en un único agente. En este sentido, el protocolo no trata de la interoperabilidad igual-a-igual entre Archivos, sino que considera la interrelación entre archivos digitales y agentes con capacidades de agregación y recopilación, éstos previsiblemente ofreciendo servicios de valor añadido a los usuarios finales.

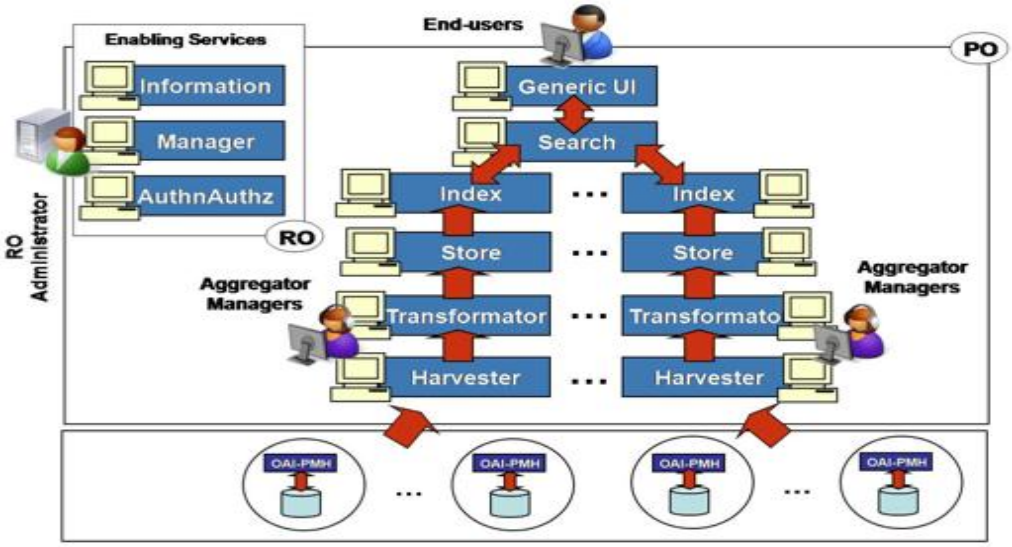


Figura 2.3 Área de Servicios de Datos: ejemplo de sistema de agregación

Fuente: Paolo Manghi

OAI-PMH no está limitado a metadatos en formato DC, aunque esta malinterpretación está algo extendida. El estándar exige al menos el soporte de Dublin Core, pero las comunidades de conocimiento específico podrían (y deberían) usar recolecciones basadas en conjunto de metadatos de mayor precisión semántica.

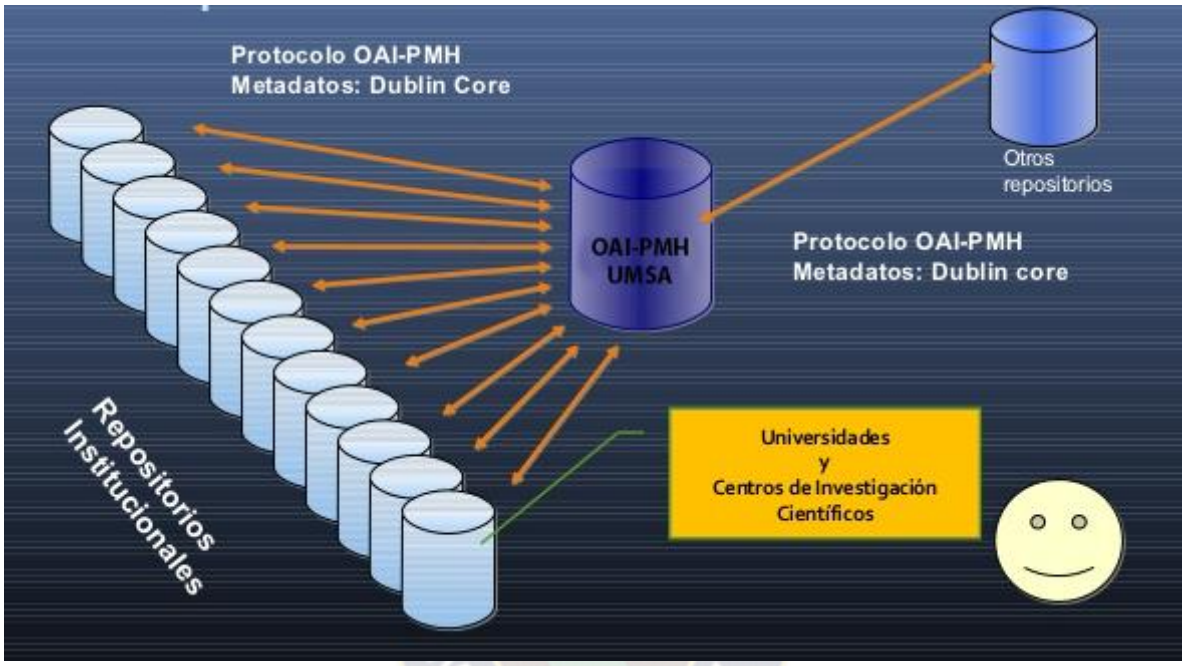


Figura 2.4 Propuesta del Servidor OAI-PMH de la Biblioteca Central

Fuente: Elaboración propia

2.9 CLASIFICACION DE LAS COLECCIONES DIGITALES DE LA BIBLIOTECA CENTRAL

En la práctica, los servidores OAI han agrupado las colecciones digitales de diferente forma y esto ha dado origen a una división de los repositorios, principalmente como:

- ✓ Repositorio de eprints.
- ✓ Repositorios temáticos.
- ✓ Repositorios de materiales académicos.
- ✓ Repositorios institucionales.
- ✓ Repositorios de objetos de aprendizaje.

2.10 TECNICAS PARA CREAR INDICES

Actualmente existen 3 técnicas para realizar el índice en un sistema:

a) Índice Invertido.

b) Índice Semántico Latente.

c) Índice Conceptual Efectivo.

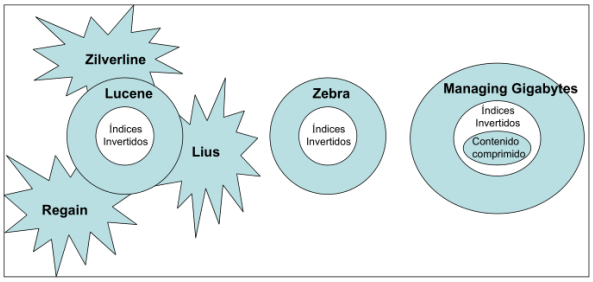


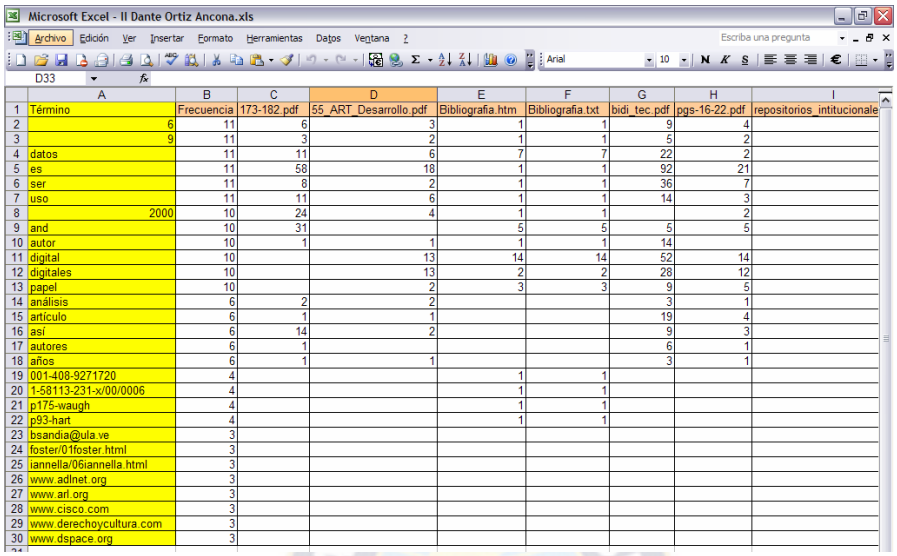
Figura 2.5 software para indizar, buscar y recuperar información

Fuente: GNU

a) INDICE INVERTIDO (II)

La representación de Índice Invertido (II) es el método dominante para indizar texto, pero no es conveniente en búsqueda de similitud entre documentos. El desempeño de la representación II empeora cuando se incrementa el número de palabras en un documento o en los casos en que una palabra tiene una lista invertida demasiado grande.

Sistemas que usan el índice invertido por ejemplo tenemos el WinISIS, ISISMARC, Microisis, etc.



	A	B	C	D	E	F	G	H	I
1	Término	Frecuencia	173-182.pdf	55_ART_Desarrollo.pdf	Bibliografia.htm	Bibliografia.txt	bidi_tec.pdf	pgs-16-22.pdf	repositorios_intitucionale
2	6	11	6	3	1	1	9	4	
3	9	11	3	2	1	1	5	2	
4	datos	11	11	6	7	7	22	2	
5	es	11	58	18	1	1	92	21	
6	ser	11	8	2	1	1	36	7	
7	uso	11	11	6	1	1	14	3	
8	2000	10	24	4	1	1		2	
9	and	10	31		5	5	5	5	
10	autor	10	1	1	1	1	14		
11	digital	10		13	14	14	52	14	
12	digitales	10		13	2	2	28	12	
13	papel	10		2	3	3	9	5	
14	análisis	6	2	2			3	1	
15	artículo	6	1	1			19	4	
16	así	6	14	2			9	3	
17	autores	6	1				6	1	
18	años	6	1	1			3	1	
19	001-408-9271720	4			1	1			
20	1-58113-231-x/00/0006	4			1	1			
21	p175-waugh	4			1	1			
22	p93-hart	4			1	1			
23	bsandia@ula.ve	3							
24	foster/01foster.html	3							
25	iannella/06iannella.html	3							
26	www.adlnet.org	3							
27	www.arl.org	3							
28	www.cisco.com	3							
29	www.derechoycultura.com	3							
30	www.dspace.org	3							

Figura 2.6 Reporte de índice invertido

Fuente: Elaboración propia

b) INDICE SEMANTICO LATENTE (ISL)

Es una técnica para mejorar la calidad en la búsqueda por similitud transformando documentos del conjunto de palabras original a un espacio de conceptos.

La idea principal de este método es proyectar los datos en un espacio pequeño, de los datos originales eliminando los efectos nocivos de sinonimia y polisemia. Trata de minimizar ambigüedad, redundancia y vocabulario sin comprimir la representación. ISL transforma los datos de una representación indizada dispersa (como en el índice invertido) con dimencionalidad alta a una representación en un espacio real mucho menos disperso.

Desafortunadamente, ISL transforma los datos en un dominio que no es posible brindar técnicas de indizado efectivas.

c) INDICE CONCEPTUAL EFECTIVO (ICE)

Un documento se representa como conjunto de atributos que corresponden a conceptos con significado. Cada uno de estos conceptos es definido por una palabra con un peso asociado (frecuencia). La palabra con el peso representa a un conjunto de palabras relacionadas semánticamente. La representación ICE es una representación comprimida que reduce ambigüedad, redundancia y vocabulario no relacionado en un documento se utiliza el método de índices invertidos para indizar los documentos.

El método ICE es mucho mejor que II en búsquedas de similitud y preserva la misma calidad de los resultados, tiene una gran eficiencia computacional y de almacenamiento por ejemplo, en una muestra de 167000 documentos se requirió 87.7 Mb usando el método II 8.3 Mb usando ICE.

2.11 LENGUAJE DE REPRESENTACION DE CONTENIDOS

2.11.1 HTML (HYPERTEXT MARKUP LANGUAGE)

Es un lenguaje de marcado diseñado para estructurar textos y presentarlos en forma de hipertexto, que es el formato estándar para las páginas web.

- ✓ Es el formato más popular que existe para la edición de documentos en la Word Wide Web.
- ✓ Se creó con objetivos de divulgación, por lo que no tuvo una planificación adecuada.
- ✓ El marcado HTML está básicamente orientado hacia la presentación.

2.11.2 XML (EXTENSIBLE MARKUP LANGUAGE)

Es un metalenguaje extensible de etiquetas desarrollado por el World Wide

Web Consortium (W3C).

- ✓ Es una simplificación del SGML y permite definir la gramática de lenguajes específicos. Por lo tanto XML no es un lenguaje en particular, sino una manera de definir lenguajes para diferentes necesidades.
- ✓ XML es una tecnología sencilla que tiene a su alrededor otras que lo complementan.
- ✓ El marcado XML se puede orientar hacia la presentación o hacia el contenido. Sin embargo, su uso correcto debe estar orientado hacia la descripción de la estructura lógica y semántica del contenido.

2.12 METADATOS

Cuando se busca el concepto de metadato en medio físico o digital, en un alto porcentaje se acuña el término con la definición que es un "dato sobre los datos", que en términos generales no deja muy claro que la palabra "datos" es la cosa a la que se está haciendo referencia. En este trabajo se entiende por metadatos, siguiendo a Eva Méndez y José A. Senso, como (toda aquella información descriptiva sobre el contexto, calidad, condición o características de un recurso, dato u objeto que tiene la finalidad de facilitar su recuperación, autentificación, evaluación, preservación y/o interoperabilidad)[12].

El objetivo final de los metadatos es ayudar a publicitar y dar soporte a los datos que una persona u organización ha producido. En la Biblioteca Central de la Universidad de Antioquia, existe dentro de sus sistema de gestión de calidad un proceso técnico en el cual participan profesionales de la información, con la responsabilidad de asignar a cada

[12] Méndez, Eva & Senso, José A. 2004

recurso los metadatos asociados, ya sea utilizando un tesauro dispuesto para tal fin o agregando términos en caso de ser necesario.

Para el caso particular del Servidor OAI de la Biblioteca Central, no se tiene que inventar ningún estándar de metadatos, a pesar de que en el medio existen varios, el servidor OAI trae por defecto el estándar Dublin Core, logrando con esto el uso de herramientas estándares con las cuales se pueden intercambiar y consultar información desde el exterior. Aunque el sistema da la posibilidad de hacer cambios en el Dublin Core, lo recomendable es no agregar elementos, con seguridad, si se hace una revisión del estándar existente, se encontrará un lugar para esa información. Lo que si puede ocurrir es que el nombre o la posición del elemento del estándar pueden ser diferentes de la que se espera, lo cual es fácil de resolver ayudándose de los archivos de configuración que están escrito en lenguaje xml.

2.13 TIPOS DE METADATOS.

2.13.1 METADATOS DESCRIPTIVOS

Su objetivo es la descripción e identificación de los recursos de información bien en un sistema local para permitir la búsqueda y recuperación de una determinada categoría de documentos o imágenes (p.e: fotografías de automóviles), bien en aplicación web para facilitar a los usuarios la localización de recursos (colecciones digitalizadas sobre prensa).

2.13.2 METADATOS ESTRUCTURALES

Facilitan la navegación y presentación de los recursos electrónicos proporcionando información sobre la estructura interna de los mismos (página, sección, capítulo, numeración, índices, tablas de contenidos); la relación entre los materiales (fotografía B incluida en el manuscrito A) y unen los archivos de imagen y textos relacionados (archivo Aes el formato JPEG de la imagen del archivo B)

2.13.3 METADATOS ADMINISTRATIVOS

Facilitan la gestión y procesamiento de las colecciones digitales tanto a corto como a largo plazo incluyendo datos técnicos sobre la creación y el control de calidad, gestión de derechos, control de acceso y utilización y condiciones de preservación (ciclos de actualización, migración, etc.).

2.14 FORMATO DE METADATOS DUBLIN CORE

La Iniciativa Dublin Core (DCMI)[13] comenzó su andadura en 1995 en un encuentro en Dublin, Ohio (USA) en el que participaron el NCSA (National Center for Supercomputing Applications) y OCLC (On Line Library Computer Center), junto con representantes de la IETF (Internet Engineering Task Force) y en el que bibliotecarios, proveedores de contenido y expertos en lenguajes de marcado logo DCMIpretendieron desarrollar estándares para describir los recursos de información y facilitar su recuperación. Así nació un pequeño conjunto de descriptores, en principio pensados para que fuera el propio autor el que los incluyera en el documento o recurso, pero que rápidamente adquirieron alcance global porque también se interesaron en ellos numerosos y variados proveedores de información pertenecientes a diferentes sectores como el de las artes, las ciencias, la educación, los negocios y la administración. Hoy, los metadatos Dublin Core se han convertido en uno de los estándares más extendidos para la recuperación de información en la World Wide Web y el DC se ha convertido en un vocabulario muy utilizado no sólo en el ámbito bibliotecario y documental, sino en otros muchos sectores. Además, este conjunto de metadatos se puede utilizar no sólo con HTML, sino sobre otros lenguajes estructurados como XML y conjuntamente con otros lenguajes de descripción como RDF. The Dublin Core metadata element set se convirtió en norma ISO 15836/2003 en febrero de 2003.

El conjunto de elementos Dublin Core se centró en 13 elementos, pero concluyó con 15 descriptores como resultado de un consenso y un esfuerzo interdisciplinar e internacional.

[13] http://dublincore.org

Ya existen transcripciones a 20 idiomas y ha sido adoptado por el CEN/ISS (European Committee for Standardization / Information Society Standardization System) y posee dos RFCs de Internet (Requests for Comments) (RFC2413) y (RFC2731). Es también el estándar oficial del WWW Consortium y el estándar del Z39.50. Los metadatos Dublin Core han sido aprobados por el organismo nacional de estandarización norteamericano (ANSI/NISO Z39.85) y los utilizan como base tanto gobiernos como agencias supranacionales y muchas otras iniciativas de metadatos pertenecientes a comunidades específicas como bibliotecas, archivos, en educación, negocios, etc.

Los metadatos Dublin Core tratan de ubicar, dentro de Internet, los datos necesarios para describir, identificar, procesar, encontrar y recuperar un documento introducido en la red. Si este conjunto de elementos Dublin Core se lograra aceptar internacionalmente supondría que todos los procesos que indizan documentos en Internet encontrarían, en la cabecera de los mismos, todos los datos necesarios para su indización y además estos datos serían uniformes. Si el Dublin Core lograra estandarizar los metadatos de la cabecera de los documentos se facilitaría su indización automática y mejoraría la efectividad de los motores de búsqueda.

El conjunto de elementos de metadatos Dublin Core es un conjunto de metadatos previsto para describir documentos. Los elementos poseen etiquetas descriptivas que pretenden transmitir un significado semántico a los mismos.

Cada elemento es opcional y puede repetirse. Además, los elementos pueden aparecer en cualquier orden. Aunque algunos entornos, como HTML, no diferencian entre mayúsculas y minúsculas, es recomendable escribir correctamente cada metadato, según su definición, para evitar conflictos con otros entornos, como SGML y XML

La descripción de referencia de los 15 elementos principales de los metadatos Dublin Core la podemos encontrar en: http://dublincore.org/documents/dces/ (en español en:

http://www.rediris.es/metadata/), aunque con más detalle aparecen estos elementos, junto con el resto de términos utilizados por la Iniciativa de Dublin Core en DCMI Metadata Terms (http://dublincore.org/documents/dcmi-terms/).

Podemos clasificar el conjunto de elementos Dublin Core en 3 grupos que indican la clase o el ámbito de la información que contienen:

Elementos relacionados principalmente con el contenido del recurso:

- Title (titulo)
- Subject (tema)
- Description (descripción)
- Source (fuente)
- Languaje (lenguaje)
- Relation (relación)
- Coverage (cobertura).

Elementos relacionados principalmente con el recurso cuando es visto como una propiedad intelectual:

- Creator (autor)
- Publisher (editor) y, otras colaboraciones
- Contributor (otros autores/colaboradores)
- Rights (derechos).

Elementos relacionados principalmente con la instanciación del recurso:

- Date (fecha)
- Type (tipo de recurso)
- Format (formato)
- Identifier (identificador)

Veamos con más detenimiento cada uno de estos 15 elementos que están recogidos en la tabla 2.1 de esta sección.

2.15 ALMACENAMIENTO, RESPALDO Y PRESERVACION DIGITAL

Las estrategias que se pueden aplicar para proteger los documentos y los datos son:

g) **Replicación**. Es la estrategia más básica y es necesario que las réplicas sean más de una. Algunos archivos de información digital hacen un número fijo de réplicas en

servidores distribuidos geográficamente. Además cada réplica puede ser que tenga copias de seguridad.

h) **Migración.** En sus dos aspectos de cambio de soporte y de cambio de formato representa una de las estrategias para prevenir la obsolescencia del formato y del software. A veces se hace la migración en el momento de la catalogacion y entonces se habla de "normalización", decidiendo si se conservará en paralelo el formato original o únicamente el objeto normalizado.

i) **Transparencia**. Esta cualidad es esencial para que el usuario pueda confiar en un sistema ya que los resultados son siempre futuros. Es de resaltar la importancia de la utilización de código abierto para asegurar el funcionamiento del sistema a largo plazo. Es remarcable la necesidad de compartir la información sobre los casos de pérdida y las mejores prácticas, tomando como modelo y ejemplo el sistema establecido por la NASA para informar de las incidencias en la navegación aérea.

j) **Diversidad**. Los sistemas que no admiten la diversidad son mucho más vulnerables que aquellos que lo hacen. Por lo tanto un sistema de preservación debería de conllevar diversidad.

k) **Auditoría.** Es necesario establecer mecanismos de control que verifiquen en diferentes fases del proceso (Catalogación, Archivo, Difusión) y mediante técnicas diversas que la información se mantiene inalterada. Las comprobaciones regulares para evitar la acumulación de errores latentes dan confianza a los que financian el sistema de preservación digital.

l) **Contención de los costos**. Uno de los peligros más grandes para garantizar la puesta en marcha y el mantenimiento de un sistema de preservación es el costo porque pocas instituciones u organizaciones pueden dedicar financiamiento a usos futuros y por la dificultad en mostrar los resultados de las inversiones realizadas.

2.16 METODOLOGIAS

2.16.1 PROCESO UNIFICADO AGIL AUP

La definición de esta metodología, se ha realizado tomando como base las siguientes referencias: [14]

2.16.1.1 DESCRIPCIÓN

El Proceso Unificado Ágil (en adelante, AUP), se describe como una metodología fácil de entender para el desarrollo de aplicaciones software de negocio, utilizando técnicas ágiles y conceptos aun fieles a las de RUP, por lo tanto, es una versión simplificada del Rational Unified Process (RUP).

AUP, incluye o hace uso de las siguientes técnicas ágiles:

- Test driven development (TDD).
- Agile model driven development (AMDD).
- Agile change management.
- Data base refactoring.

AUP, es una metodología que tiene la adopción de muchas de las técnicas ágiles de la metodología XP (Extreme Programming) y de las formalidades de RUP, teniendo como filosofía adaptarse a las necesidades del proyecto y no al contrario como lo planteado en las metodologías tradicionales. Esta metodología, plantea un ciclo de vida iterativo, que se basa en la ampliación y refinamiento sucesivo del sistema, mediante múltiples iteraciones con retroalimentación cíclica y adaptación como elementos principales que dirigen para converger hacia un sistema adecuado.

[14] [CRLA02], [SABL06]. [KINT07], [WAUP09].

2.16.1.2 PRINCIPIO BÁSICOS

Como parte de esta metodología, en la presente sección se detalla, tanto las fases y disciplinas de su planteamiento:

a) Fases

Al igual que RUP, AUP tiene las 4 fases clásicas consecutivas y que acaban cada uno de estas, con hitos claros alcanzados:

- ✓ **Inception (Concepción):** El objetivo de esta fase es obtener una comprensión común cliente equipo de desarrollo, del alcance del nuevo sistema y definir una o varias arquitecturas candidatas para el mismo.
- ✓ **Elaboración:** El objetivo es que el equipo de desarrollo profundice en la comprensión de los requisitos del sistema y en validar la arquitectura.
- ✓ **Construcción:** Durante la fase de construcción el sistema es desarrollado y probado al completo en el ambiente de desarrollo.
- ✓ **Transición**: el sistema se lleva a los entornos de preproducción donde se somete a pruebas de validación y aceptación y finalmente se despliega en los sistemas de producción.

b) Disciplinas

El proceso AUP, establece un modelo más simple que el planteado en RUP, por lo que reúne en una única disciplina: el modelado de negocio, requisitos y análisis y diseño. El resto de disciplinas coinciden con las restantes de RUP, a continuación se describe las disciplinas consideradas por AUP:

- ✓ **Modelado:** Se busca entender el negocio de la organización, el problema de dominio que se abordan en el proyecto, y determinar una solución viable.
- ✓ **Implementación**: Consiste en transformar los modelos o modelo en código ejecutable y realizar un nivel básico de las pruebas, en particular *Unit testing*.

- ✓ **Pruebas**: Se busca realizar una evaluación objetiva para garantizar la calidad. Esto incluye la búsqueda de defectos, validar que el sistema funciona tal como está establecido, y verificando que se cumplan los requisitos.
- ✓ **Despliegue:** Consiste en la elaboración de un plan para la entrega del sistema y ejecutar el plan para que el sistema esté a disposición de los usuarios finales.
- ✓ **Gestión de Configuración**: Consiste en administrar el acceso a los artefactos del proyecto. Esto incluye no sólo el seguimiento de las versiones de los artefactos en el tiempo, sino también el control y la gestión de los cambios de estos.
- ✓ **Administración del Proyecto**: Consiste en dirigir las actividades que se llevan a cabo dentro del proyecto. Esto incluye la gestión de riesgos, la dirección de personas (la asignación de tareas, el seguimiento de los progresos, etc), y de coordinación con el personal y los sistemas fuera del alcance del proyecto para asegurarse de que el software final sea entregado a tiempo y dentro del presupuesto.
- ✓ **Entorno**: Es un soporte para el resto de los esfuerzos para garantizar un proceso adecuado, orientación (normas y directrices), y herramientas (hardware, software, etc) estén disponibles para el equipo según sea necesario.

2.16.1.3 PROCESO DE DESARROLLO

En esta metodología las disciplinas se llevan a cabo de manera iterativa, con la definición de las actividades de los miembros del equipo de desarrollo, con el fin de desarrollar, validar y entregar el software que responda a las necesidades de los stakeholders.

En cada disciplina la metodología plantea las diferentes actividades y artefactos a producir, lo cual no implica que se realicen o se produzcan todo lo planteado sino más bien lo que se necesita en el proyecto.

Las fases que platea la metodología no constituye el antiguo ciclo de vida secuencial o en cascada, sino más bien, es planteada de la siguiente manera:

- ✓ La fase de Inicio (Inception), no es una fase de requisitos, sino más bien una especia de fase de viabilidad, donde se lleva acabo el estudio suficiente, para decidir si continuar o no el proyecto.
- ✓ La fase de Elaboración no es una fase de requisitos o diseño, sino que es una fase donde se implementa de manera iterativa la arquitectura que constituye el núcleo central del sistema, y es donde se mitiga las cuestiones de alto riesgo.
- ✓ En la fase de construcción, se implementa de manera iterativa el resto de requisitos (de menor riesgo), se realiza pruebas y se prepara para el despliegue.

Por cada una de las fases e iteraciones planteadas en las mismas, se puede hacer uso de la totalidad de las disciplinas o solo de algunas, esto dependerá de la iteración en la que se encuentre, debido a que el esfuerzo relativo en las disciplinas disminuye de iteración en iteración.

AUP, plantea que en lugar del "*big bang*" en la entrega del software, una liberación del producto en partes, por ejemplo versión 1, luego versión 2 del software en producción y así sucesivamente hasta tener el producto completado. De acuerdo a este planteamiento AUP distingue dos tipos de iteraciones:

- ✓ Versión de desarrollo, cuyo resultado está un despliegue en entorno QA (Aseguramiento de la calidad) o Demo área, estas versiones deben ser rápidas, de ser posible de una a tres semanas de desarrollo.

- ✓ Versión producto, cuyo resultado es un despliegue en producción, donde la primera entrega estable tardará en promedio 12 meses, la siguiente 9 meses y las siguientes cada 6 meses.

AUP, se preocupa principalmente de la gestión de riesgos. Propone que aquellos elementos con alto riesgo obtengan prioridad en el proceso de desarrollo y sean abordados en etapas tempranas del mismo. Especialmente relevante en este sentido es el

desarrollo de prototipos ejecutables durante la fase de elaboración del producto, donde se demuestre la validez de la arquitectura para los requisitos clave del producto y que determinan los riesgos técnicos.

2.16.2 METODOOGIA UWE (UML-BASED WEB ENGINEERING)

UWE[15] es un proceso de desarrollo para aplicaciones Web que se centra en el diseño sistemático, la personalización y la generación semi-automática del sistema. Describe una metodología de diseño basado exclusivamente en técnicas, notaciones y mecanismos de extensión UML (Unified Modeling Lenguage). De manera a cubrir algunas características especiales que respectan al diseño de una aplicación Web, UWE define algunas vistas especiales, representados por diagramas UML, el modelo de navegación y el modelo de presentación. De manera adicional, el diseñador Web podría hacer uso de diferentes técnicas de modelado agregando más vistas si fuese necesario, no existen limitaciones para el número de vistas. UWE hace uso de los mecanismos de extensión de UML para definir estereotipos de manera a generar las vistas especiales el modelado de las aplicaciones Web.

La ventaja de utilizar diagramas UML es el conocimiento semántico de estos diagramas. Resulta fácil de entender, para cualquier diseñador de software, un modelo basado en UML. UWE recomienda la separación estricta de intereses en las primeras fases del desarrollo y también la implementación de un proceso de desarrollo dirigido por modelos. El mayor desafío es realizar un proceso de desarrollo que permita la generación automática total de sistemas web.

[15] Based[OMG, 2003]

2.16.2.1 LA METODOLOGÍA UWE

La metodología UWE provee algunas normas para la construcción sistemática de modelos propuestos por Hennicker y Koch[16]. Los principales modelos utilizados son: análisis de requerimientos, modelo conceptual, modelo de navegación y modelo de presentación, complementados con tareas y visualización de escenarios Web. A continuación se explican brevemente cada uno de estos modelos.

2.16.2.2 ESPECIFICACIÓN DE REQUERIMIENTOS

La especificación de requerimientos se realiza a través del diagrama de casos de usos UML.

Aplicando las técnicas UML utilizadas en UWE en el servidor de metadatos de publicación de documentos académicos, este servidor ofrece información acerca de publicaciones a usuarios anónimos y a usuarios que solicitan registrar su publicacion. Una publicación contiene información relativa a proyectos de pregrado, postgrado y libros artículos de investigación, etc. Esta publicación se describe mediante un título, un número, una editorial, una fecha de publicación, un conjunto de artículos y autores para cada uno de los artículos. Un libro consiste en exactamente un artículo, en el cual el título del artículo es el mismo que el título del libro. De manera adicional, cada artículo y publicación poseen un conjunto de palabras claves asociadas. Se distinguen los siguientes tipos de usuarios: lectores registrados o no registrados y un administrador. El lector registrado es modelado de manera tal a que en cada visita pueda encontrar los artículos que son propios de su interés en particular. El lector registrado también puede marcar un artículo como de interés especial. La aplicación administra una lista de claves para cada lector. El sistema administra la actualización del modelo de usuario de acuerdo a las observaciones obtenidas en las conductas de los lectores (para este caso se limita sólo a los artículos marcados o visitados frecuentemente). La lista puede incluir tanto claves positivas como

[16] Año de publicación [2000].

negativas. Las claves que son marcadas como negativas son utilizadas de manera tal a ocultar publicaciones y artículos que son irrelevantes para el lector.

La figura 2.7 muestra el diagrama de casos de usos para aplicación Web de la biblioteca central.

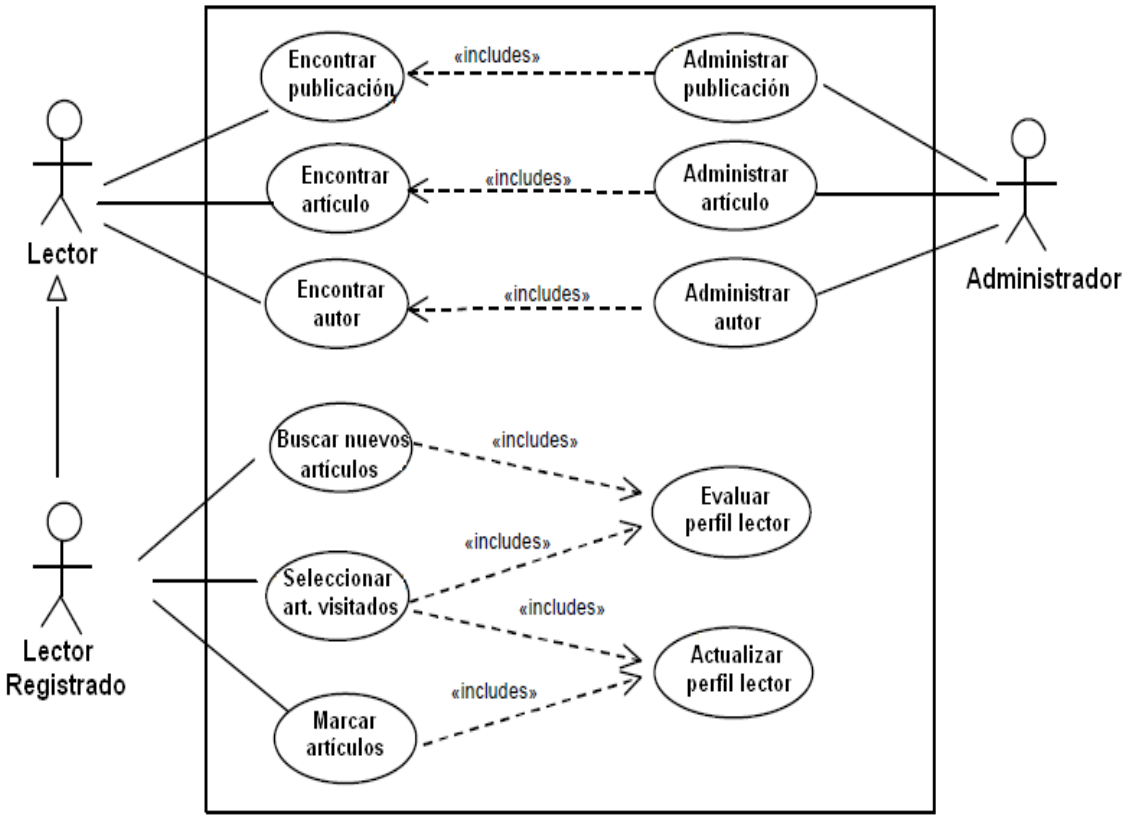


Figura 2.7 Diagrama de Casos de Usos de la Biblioteca

Fuente: Elaboración Propia

2.16.2.3 MODELO CONCEPTUAL

El modelo conceptual se realiza a través del diagrama de clases UML. Siguiendo el modelo OOHDM[17], UWE construye el modelo conceptual de una aplicación Web de manera tal a ignorar lo más que se pueda los aspectos relacionados a la navegación, presentación e interacción, para luego enfocarse en estos aspectos con el modelo navegacional y el modelo de presentación respectivamente.

El diagrama de clases posee una variedad de características adicionales que pueden ser utilizados semánticamente para mejorar estos diagramas. Algunos ejemplos de estas características son la asociación (agregación, composición), nombre de roles, multiplicidad. Si el modelo conceptual presenta una gran variedad de clases, se recomienda que los mismos sean agrupados utilizando el elemento de modelado UML denominado package.

Basado en los casos de la sección anterior y la descripción detallada de los mismos se construye el modelo conceptual. La figura 2.8 muestra el modelo conceptual para el servidor OAI. En el diagrama se limita a los datos y funcionalidades principales, se podrían incluir otros aspectos a través de un proceso iterativo e incremental del proceso.

[17] (Rossi, 2000)

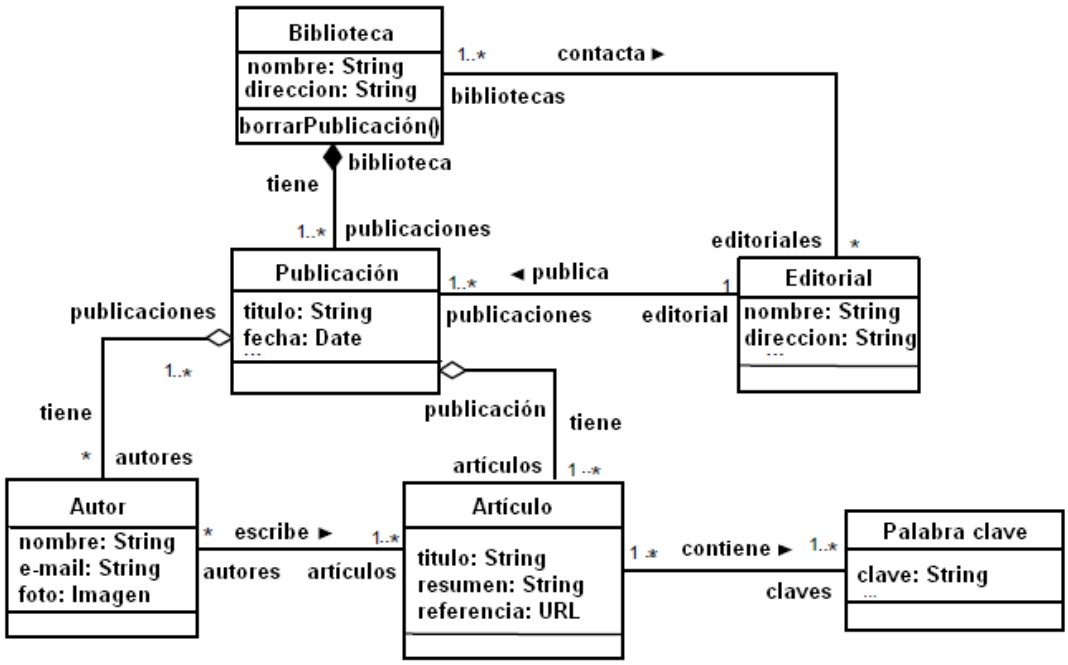


Figura 2.8 Modelo Conceptual para la Biblioteca

Fuente: Elaboración propia

2.16.2.4 MODELO DE NAVEGACIÓN

El modelo de navegación se realiza básicamente dos etapas. En la primera se construye el modelo de espacio de navegación y en la segunda el modelo de estructura de navegación. El primero especifica que objetos pueden ser visitados en la aplicación a través de la navegación. Mientras que el segundo define como se puede acceder a dichos objetos. Los modelos de navegación son representados a través de estereotipos de los diagramas de clases UML.

La figura 2.9 muestra el modelo de espacio de navegación para la Biblioteca Web del del servidor OAI. Los elementos principales modelados son las clases estereotipos <<navegation class>> y la asociación estereotipada <<direct navigability>>. Es importante notar que sólo las clases que son relevantes en el modelo conceptual son incluidas en el

modelo de navegación. Algunas informaciones de clases omitidas podrían ser incluidas como atributos de alguna otra clase de navegación, el nuevo atributo editorial introducido dentro de la clase de navegación Publicación. Se utiliza OCL (Object Constraint Language) para expresar las relaciones existentes entre clases conceptuales y clases de navegación o atributos de clases de navegación.

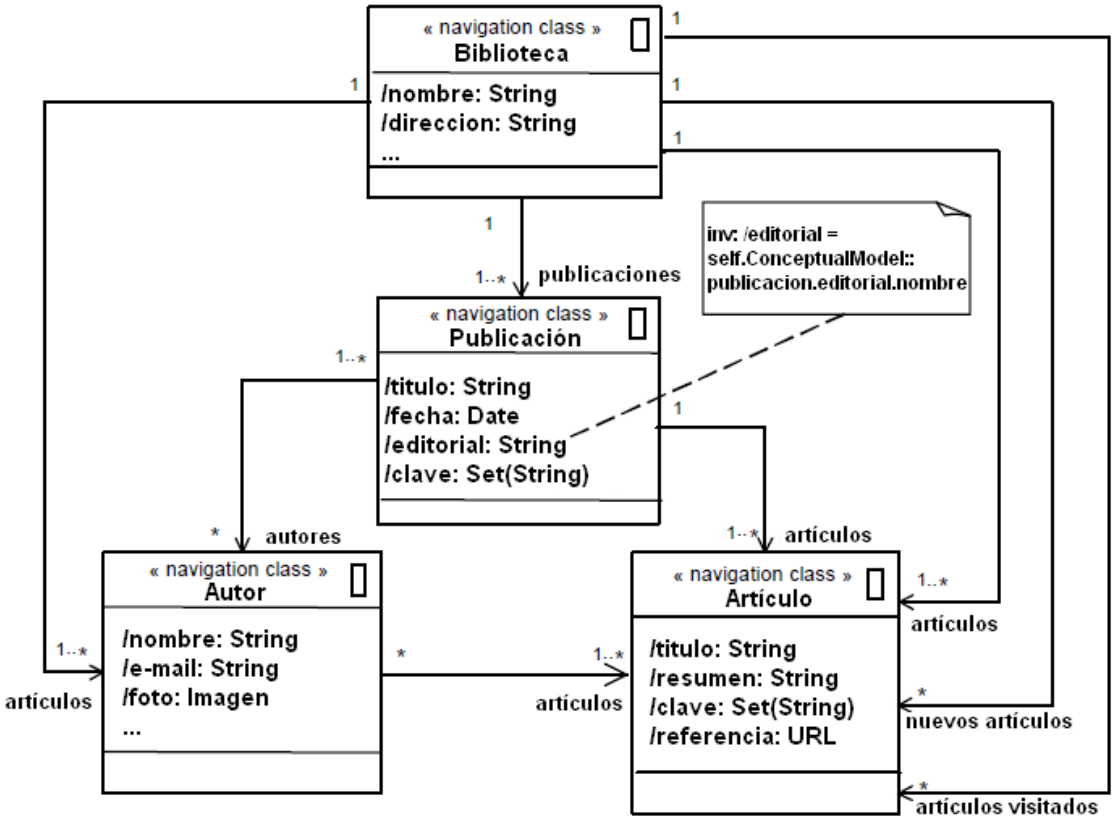


Figura 2.9 Modelo de Espacio de Navegación para la Biblioteca

Fuente: Elaboración propia

El modelo de estructura de navegación se construye utilizando como base el modelo de espacio de navegación. Puede ser considerado como un paso de refinamiento en el proceso de diseño UWE.

UWE define un conjunto de pasos y mecanismos semiautomáticos para modelar la navegación de una aplicación Web[18], los cuales consisten reforzar el modelo de espacio de navegación con índices, guías, consultas y menús.

Las clases estereotipadas para el acceso elementos son <<index>>, <<guided tour>>, <<query>> y <<menú>>. La semántica e iconos para estos estereotipos se describen en Baumeister (1999). La figura 2.10 muestra el modelo de estructura de navegación. Se puede observar las etiquetas {sorted} que indican que los índices correspondientes se encuentran ordenados de acuerdo a las preferencias del usuario. El sistema obtiene las preferencias de los valores actuales del modelo de usuario. Una vez realizado este diagrama el mismo resulta bastante útil para representar la estructura estática de la navegación.

[18] Hennicker y Koch, 2000; Hennicker y Koch, 2001

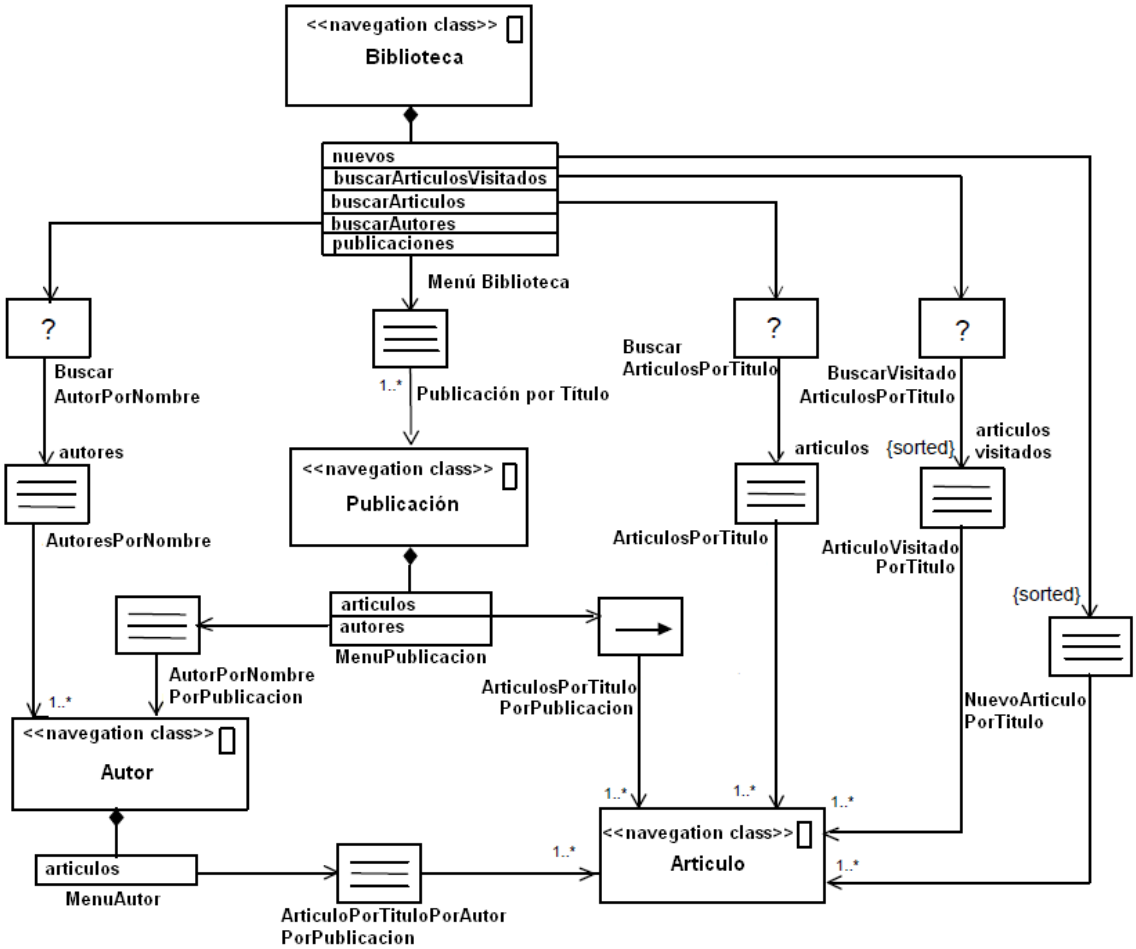


Figura 2.10 Modelo de Estructura de Navegación para la Biblioteca

Fuente: Elaboración propia

2.16.2.5 MODELO DE PRESENTACIÓN

Se utiliza un formulario particular de un diagrama de clases para representar el modelo de presentación. Esto se obtiene utilizando la composición UML para las clases. Este tipo de representación es apropiado para modelar interfaces de usuario que permitan ordenar y dimensionar en forma relativa la información disponible en dicha interfaz, aunque el mismo no sea manejado directamente por las herramientas case estándares.

El modelo de presentación permite definir dónde y cómo serán presentados los objetos navegacionales y accesos primitivos al usuario. La figura 2.11 muestra un prototipo para la clase de navegación Publicación. La elaboración de este tipo de prototipos es muy útil a la hora de realizar la captura de requerimientos con el cliente.

El conjunto de elementos estereotipados propuestos por UWE para la definición del modelo de presentación son: <<text>>, <<form>>, <<buttom>>, <<image>>, <<audio>>, <<anchor>>, <<collection>> y <<anchored collection>>.

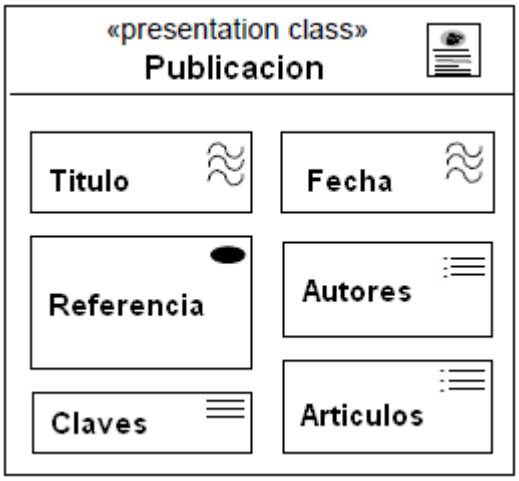


Figura 2.11 Prototipo de Presentación de Publicación para la Biblioteca

Fuente: Elaboración propia

2.16.2.6 MODELADO DE TAREAS

Los modelos de tareas son definidos como descripciones lógicas de las actividades a llevar a cabo para alcanzar los objetivos o metas del usuario, se centran por tanto en la percepción que el usuario posee del sistema.

Atendiendo a esta definición y teniendo en cuenta la necesidad de describir y caracterizar de forma clara las tareas en las organizaciones de carácter colaborativo los modelos de

tareas se convierten en el medio ideal para especificar qué datos intervienen en la gestión de esas tareas y cómo se relacionan entre sí.

Se utilizan diagrama de actividades para el modelado de tareas o flujo de procesos. La figura 2.11 muestra un modelo de tareas para la tarea Borrar una Publicación.

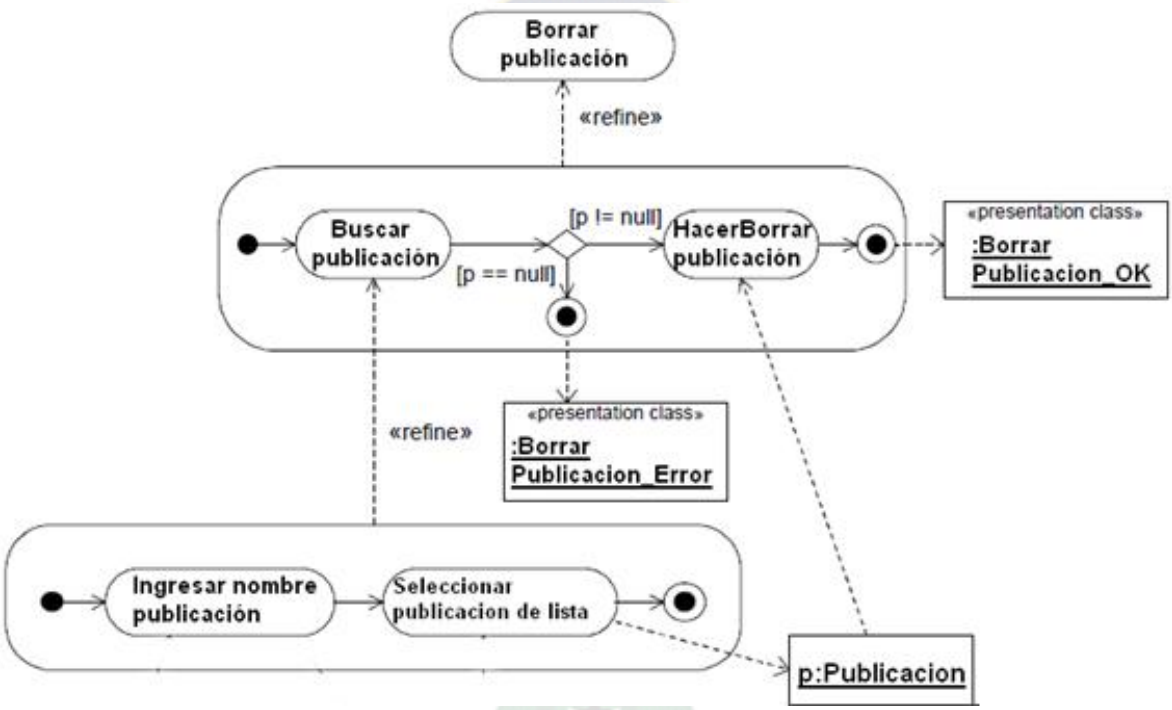


Figura 2.11 Modelo de tareas para la tarea Borrar una Publicación

Fuente: Elaboración propia

CAPITULO III

MARCO APLICATIVO

3.1 INTRODUCCION

En el capítulo presente se describe el análisis y diseño del sistema donde se aplica la metodología AUP en sus cuatro fases y sus siete disciplinas iterativas y UWE como herramienta de modelado.

3.2 FASE DE INICIO

En la fase de inicio se realiza las actividades del AUP y la planificación inicial del sistema. La Biblioteca Central de la UMSA se encarga de recepcionar las tesis, proyectos de grado y otros documentos de investigación para la preservación y publicación de los mismos para las diferentes áreas del conocimiento dentro de la universidad.

3.2.1 REQUERIMIENTOS DEL SISTEMA

La lista de requerimientos aumenta a medida que se añade nuevos elementos en la lista de requisitos para el buen funcionamiento del área de recepción y secretaria.

- ✓ Registrar datos del documento, autores, revisores y tutores.
- ✓ Entrega del formulario de recepciones.
- ✓ Registrar nuevos documentos y actualizados.
- ✓ Registro de documentos de acuerdo a la solicitud de necesidades.
- ✓ Resguardar los documentos.
- ✓ Restringir el ingreso al personal no autorizado.

- ✓ Para ingresar al sistema debe pasar por una autenticación donde debe ingresar datos del usuario y contraseña.
- ✓ Selección de documentos por colecciones para su publicación.
- ✓ Cálculo de cantidades de documentos publicados en el servidor OAI para generar informes mensuales a nivel general para la biblioteca central.
- ✓ Generación de reportes.

3.2.2 REQUERIMIENTOS TECNOLÓGICOS

Requerimientos de Software:
En el requerimiento del software para la implementación del sistema, se utiliza:

- Sistema operativo Linux Debian 8.1 o superior
- Lenguaje de Programación Java OpenJDK 8
- Motor de Base de Datos PostgreSQL 9.4
- Tomcat
- Spring
- XML
- APACHE ANT

Requerimientos de Hardware:

- Procesador I3 o superior
- Memoria 8Gb como mínimo
- Disco Duro de almacenamiento 1 Tera
- Tarjeta de red de norma IEEE 802.11n a velocidad de 300Mbps

3.2.3 MODELO DE NEGOCIOS

El modelo del negocio se basa en el estudio preliminar de la Biblioteca Central se procede a identificar y describir cada uno de los procesos de negocio, determinando el tipo de información, actividades, roles y reglas del negocio implicadas. Con el modelado de negocio lo que se pretende es comprender toda la actividad de la Biblioteca Central relacionada con el sistema.

Entender el modelo de negocio de la biblioteca central, el problema de dominio que se aborda en este proyecto, y determinar una solución viable para resolver el problema de dominio.

3.2.4 MODELADO DE CASOS DE USO DE NEGOCIO

Para comprender el contexto del sistema, se opta por realizar los modelos de casos de uso del negocio y actores que corresponden con los procesos del negocio y los clientes.

Casos de uso. Para de solicitar el resguardo del documento

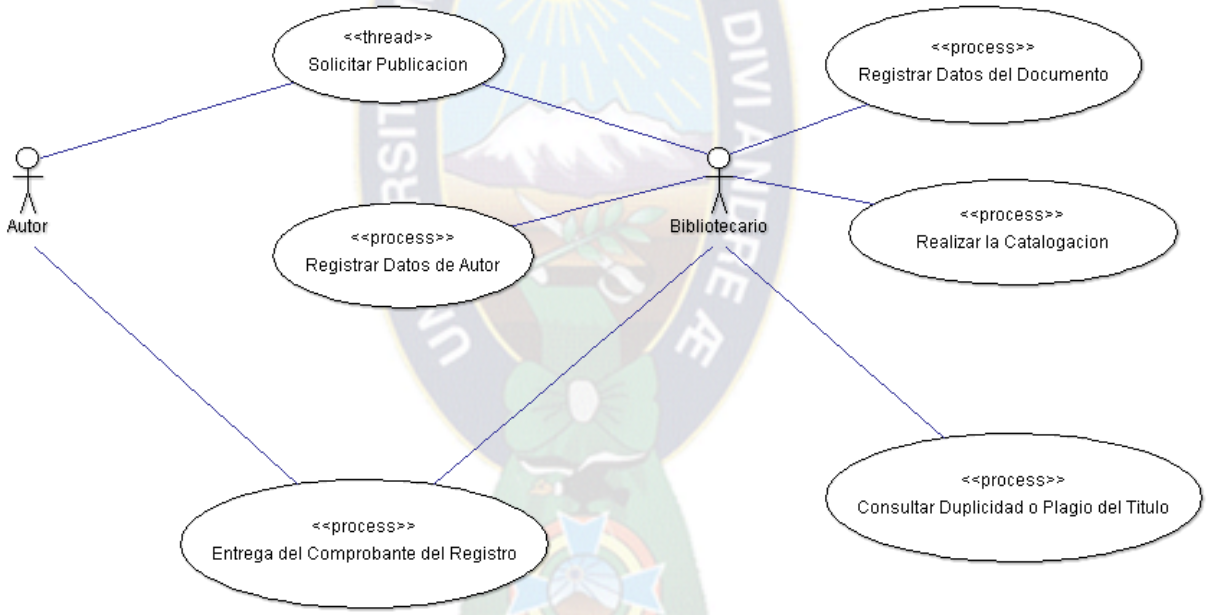


Figura 3.1 Diagrama de Casos de uso del Negocio Actual

Fuente: Elaboración Propia

3.2.4.1 DESCRIPCIÓN DE CASOS DE USO

Solicitar Publicación

Casos de uso:	Solicitar publicación
Actor	Autor y Bibliotecario
Propósito:	Realizar la solicitud de la publicación del documento.
Tipo:	Primario
Descripción:	El autor solicita la Publicación del documento en el servidor OAI, luego el bibliotecario elabora el registro para su posterior publicación.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- El caso de uso tiene su inicio cuando el autor solicita la publicación del documento. 3.- luego llena el formulario de registro con sus respectivos datos verificando que este correctamente llenado con los datos completos.	2.- el bibliotecario registra lo documento y datos del autor, selecciona la colección correcta para su publicación.

Tabla 3.1 Casos de Uso Expandido: Solicitar publicación
Fuente: Elaboración propia

Registra Datos del Autor

Casos de uso:	Registra datos del autor
Actor	Bibliotecario
Propósito:	Registra al autor que quiere la publicación del documento.
Tipo:	Primario
Descripción:	Se realiza el registro del autor para una mejor información y referencia del autor.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- El caso de uso tiene su inicio cuando el bibliotecario realiza el registro del autor con sus datos completos y referencias autor solicita la publicación del documento. 3.- El bibliotecario guarda la copia como registro y al autor se extiende un comprobante de registro original.	2.- los datos son registrados manualmente en libros de registro que son hojas con copia que sirven para tener un registro del título y el autor y sirve para el control de inventarios.

Tabla 3.2 Casos de Uso de Expandido: Registrar Autor
Fuente: Elaboración propia

Registrar datos del documento

Casos de uso:	Registra datos del documento
Actor	Bibliotecario
Propósito:	Se realiza el registro del documento.
Tipo:	Primario
Descripción:	Se realiza el registro del documento para un mejor control de datos.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- El caso de uso tiene su inicio cuando el bibliotecario realiza el registro del documento: título del documento autor al que corresponde y otros datos. 3.- El bibliotecario se encarga de guardar los datos del documento y su autor en copias para un inventario adecuado.	2.- los datos del documento son registrados manualmente con sus respectivos datos, en formularios que contienen copias.

Tabla 3.3 Casos de Uso Expandido: Registrar datos del documento
Fuente: Elaboración propia

Realizar la Catalogación

Casos de uso:	Realizar la Catalogación
Actor	Bibliotecario
Propósito:	Realiza por el titulo la catalogación en la colección al que corresponde.
Tipo:	Primario
Descripción:	Se realiza la catalogación del documento en la colección seleccionada.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- En este caso de uso el bibliotecario elabora la catalogación del documento. 3.- Luego de realizar la catalogación se guarda una copia de respaldo por el bibliotecario.	2.- Se elabora la catalogación del documento, se busca la colección al que pertenece el documento y se llena con los respectivos datos del documento.

Tabla 3.4. Casos de Uso Expandido: Realizar la Selección de la Colección
Fuente: Elaboración propia

Consultar Duplicidad o Plagio del Titulo

Casos de uso:	Consultar Duplicidad o Plagio del Titulo
Actor	Bibliotecario
Propósito:	Realiza por el titulo la verificación del plagio o duplicidad del contenido.
Tipo:	Primario
Descripción:	Se realiza la consulta a la base de datos de la existencia de otro título similar y contenido plagiado.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- En este caso de uso el bibliotecario verifica la existencia de otro título Similar. 3.- Luego de realizar la verificación se procede a la catalogación.	2.- Se elabora la verificación del documento, se busca en la colección al que pertenece el documento.

Tabla 3.5. Casos de Uso Expandido: Consultar Duplicidad o Plagio del Titulo
Fuente: Elaboración propia

Caso de uso: Para la Catalogación del Documento

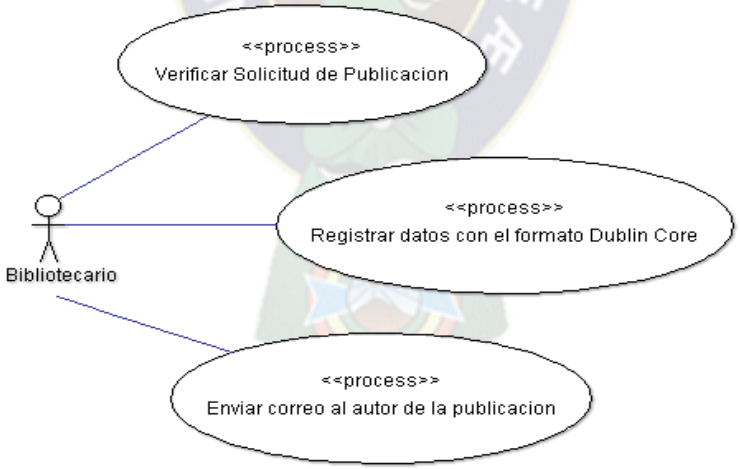


Figura 3.2 Diagrama de Casos de uso Procedo de Catalogación
Fuente: Elaboración Propia

Catalogación del Documento

Casos de uso:	Catalogación del Documento
Actor	Bibliotecario
Propósito:	Realiza por el titulo la catalogación del documento en Dublín Core.
Tipo:	Primario
Descripción:	Se realiza la verificación de la solicitud de publicación y se describen los datos del documento en el formato Dublín Core.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- En este caso de uso el bibliotecario verifica la solicitud de la publicación. 3.- Luego se envía un correo al autor informando que su documento ha sido dado de alta en el servidor OAI.	2.- Se elabora el registro de los datos del documento en formato Dublín Core. Esto se registra en el sistema.

Tabla 3.6 Caso de Uso Expandido: Catalogación del Documento
Fuente: Elaboración propia

3.3 FASE DE ELABORACION

3.3.1 ESPECIFICACIÓN DE REQUERIMIENTOS

La especificación de requerimientos es de acuerdo a la funcionalidad del sistema, se tomara en cuenta los casos de uso del negocio y los casos de uso del diseño del sistema.

3.3.1.1 FUNCIONES DEL SISTEMA

Las funciones del sistema son todas las cualidades a la que el sistema debe responder, estas tareas se detallan en la siguiente tabla

Funciones Básicas del Sistema

S: opcional **O: oculta** **E: evidente**

Referencia	Función	Categoría
R1	Realiza autenticación de login y password para ingresar al sistema	**E**
R2	El sistema deberá restringir el manejo de opciones del menú principal de acuerdo al tipo de usuario.	**O**
R3	El sistema deberá presentar ventanas independientes de	**E**

	operaciones de registro y consulta	
R4	El sistema deberá controlar el ingreso de datos en la catalogación con el formato establecido para su posterior almacenamiento en la base de datos	**E**
R5	El sistema deberá generar datos estadísticos o consultas para su posterior impresión.	**O**
R6	El sistema deberá registrar nuevos datos de bibliotecarios, autores.	**E**
R7	El sistema deberá contar con la debida seguridad respecto a los datos almacenados en la base de datos	**O**
R8	En el sistema se deberá realizar copias de seguridad de la base de datos por dispositivos de salida.	**O**

Tabla 3.7 Requerimientos Funcionales
Fuente: Elaboración propia

3.3.2 MODELO DE CASOS DE USO DEL ANÁLISIS

Se describen los principales actores del negocio encontrados los cuales serán utilizados como entrada en el desarrollo del modelo de casos de uso del sistema.

LISTA DE ACTORES

BIBLIOTECARIO	ADMINISTRADOR DEL SISTEMA	AUTOR
El encargado de la biblioteca se encarga de realizar las siguientes tareas: I. Registrar la solicitud de la publicación del documento. II. Realizar el registro del autor. III. Realizar el registro del documento. IV. Realizar la catalogación con Dublín Core. V. Realizar los reportes de los documentos publicados en el servidor OAI.	El administrador se encarga de realizar las siguientes tareas: I. Adiciona II. Modifica III. Elimina IV. Actualiza	El autor del documento y sus requerimientos. I. Solicita la publicación II. Analiza la Publicación III. Aprueba la Publicación.

Tabla 3.8 Descripción de actores principales
Fuente: Elaboración propia

3.3.3 CASOS DE USO

Los casos de uso son iteraciones del sistema donde el usuario interactuar con el sistema. Una vez identificado los actores y los casos de uso entonces procedemos a modelar.

Figura 3.3. Diagrama de Casos de uso del Sistema
Fuente: Elaboración Propia

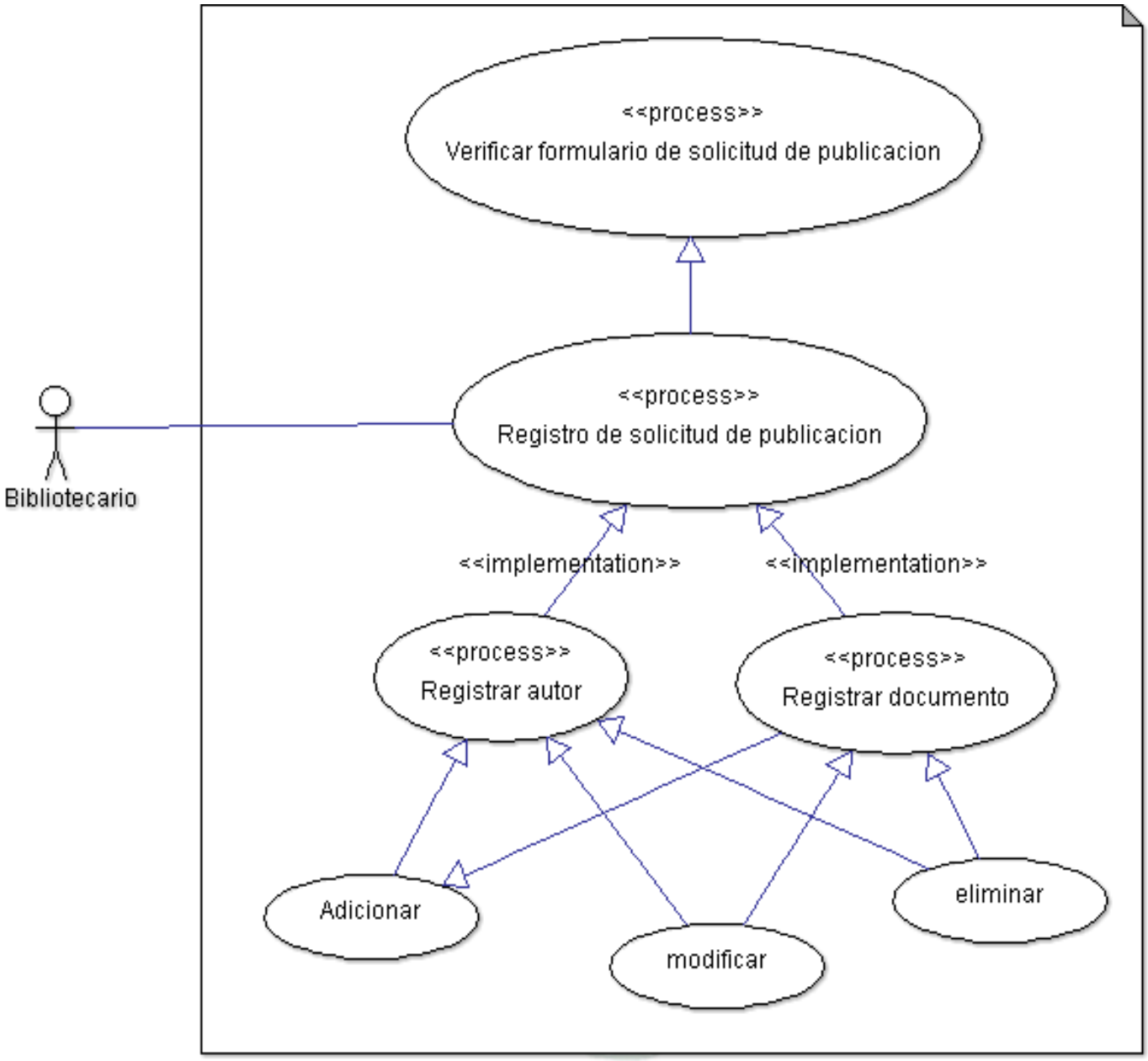


Figura 3.4 Diagrama de Casos de Uso: Registrar solicitud de Publicación
Fuente: Elaboración Propia

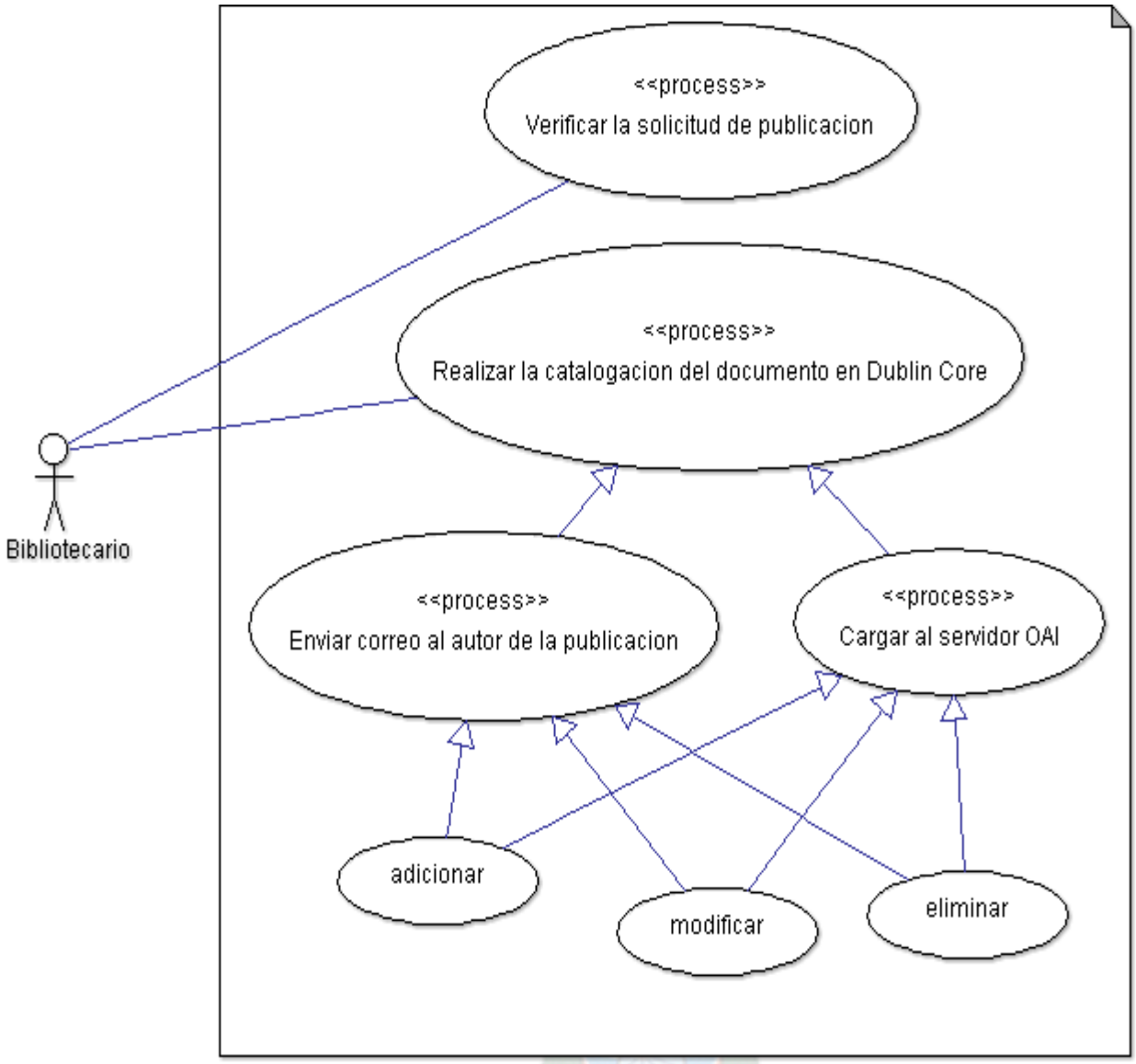


Figura 3.5 Diagrama de Casos de Uso: Catalogación del Documento
Fuente: Elaboración Propia

3.3.3.1 DETALLE DE CASOS DE USO EXPANDIDO

Los casos de uso expandido describen de forma detallada los procesos de cada caso de uso en forma teórica, cada acción que realiza el personal de la biblioteca central y la respuesta del sistema.

Registro de Solicitud de Publicación

Casos de uso:	Registro de solicitud de publicación
Actor	Bibliotecario y Autor
Propósito:	Verifica la solicitud de la publicación
Tipo:	Primario
Descripción:	El bibliotecario verifica el documento de publicación y los datos de la solicitud para luego proceder con el llenado del formulario de solicitudes.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- El bibliotecario introduce su login y password. 3.- Elige la opción crear solicitud de publicación. 5.- El bibliotecario escoge por orden las solicitudes de publicación y los documentos a registrar.	2.-El sistema verifica si los datos introducidos son válidos y son los correctos. a) Si los datos son correctos se muestra la interfaz del menú principal de usuario, b) Si los datos no son correctos el sistema solicita nuevamente los datos y muestras los errores. 4.- Se muestra el menú de registro de solicitudes.6.- Si los datos son correctamente llenados se acepta la solicitud de publicación. 7.- Luego de realizar la solicitud completa se guarda y actualiza la base de datos.

Tabla 3.9 Casos de uso: Registro de Solicitud de Publicación
Fuente: Elaboración propia

Casos de uso:	Realizar la catalogación del documento en Dublín Core
Actor	Bibliotecario
Propósito:	Catalogar todos los documentos.
Tipo:	Primario
Descripción:	Se realiza la catalogación de todos los documentos en Dublín Core de todas las solicitudes realizadas.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- El bibliotecario introduce su login y password.	2.-El sistema verifica si los datos introducidos son válidos y son los correctos. a) Si los datos son correctos se muestra la interfaz del menú principal de usuario, b) Si los datos no son correctos el sistema solicita nuevamente los datos y muestras los errores.
3.- Elige la opción catalogar documento.	4.- Muestra el formulario de catalogación de documentos.
5.- El bibliotecario llena todos los datos del nuevo documento con reglas de catalogación e indización en el formato Dublín Core.	6.- Si los datos de la catalogación son correctamente llenados se acepta la publicación. 7.- Luego de realizar la catalogación completa se guarda y publica el documento.

Tabla 3.10 Casos de uso: Realizar la catalogación del documento en Dublín Core
Fuente: Elaboración propia

3.3.4 DIAGRAMA DE SECUENCIA

En los diagramas de secuencias se muestran un conjunto de objetos que a través de mensajes permite que el sistema y el actor interactúen para dar cumplimiento a ciertas tareas que ordenadamente interactúan según la secuencia temporal de eventos por tanto se realiza lo más importante.

Registrar Solicitud de Publicación.

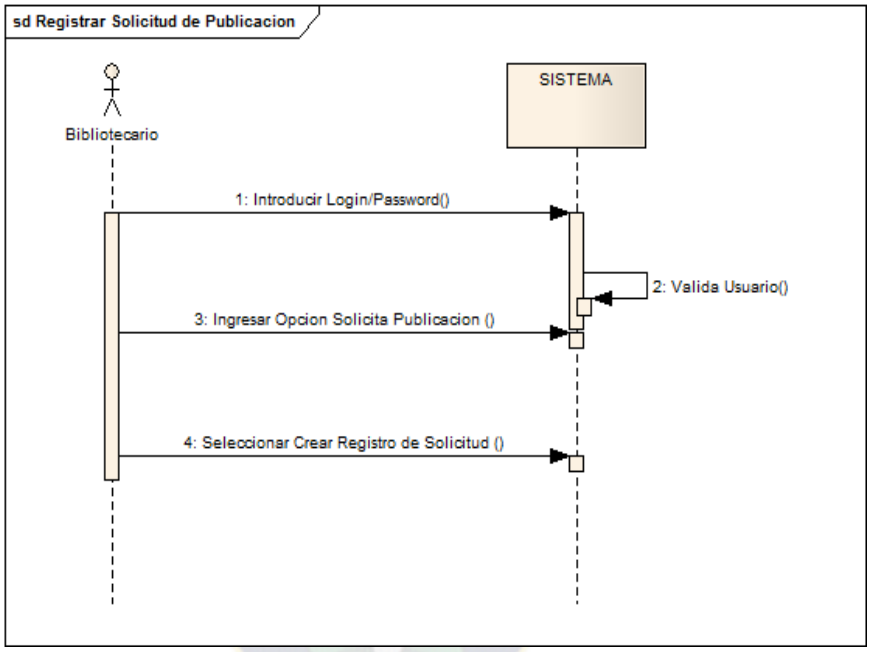


Figura 3.6 Diagrama de Secuencia: Registro de Solicitud de Publicación
Fuente: Elaboración Propia

Contrato de Operaciones

Nombre:	Introducir login y password
Responsabilidades:	Despliega y confirma autenticación
Tipo:	Manual
Referencias:	Entrega la solicitud
Precondición:	el bibliotecario debe estar registrado en el sistema
Postcondición:	Proporciona la comprobante de registro de solicitud.

Tabla 3.11 Contrato de Operaciones- Introducir login y password
Fuente: Elaboración propia

Nombre:	Solicita Publicación
Responsabilidades:	Despliega el formulario de solicitud de publicación
Tipo:	Sistema
Referencias:	Detalles del formulario de solicitud de publicación
Precondición:	El bibliotecario para registrar nueva solicitud de publicación
Postcondición:	Proporciona el formulario

Tabla 3.12 Contrato de Operaciones- Solicita Publicación
Fuente: Elaboración propia

Registrar Autor

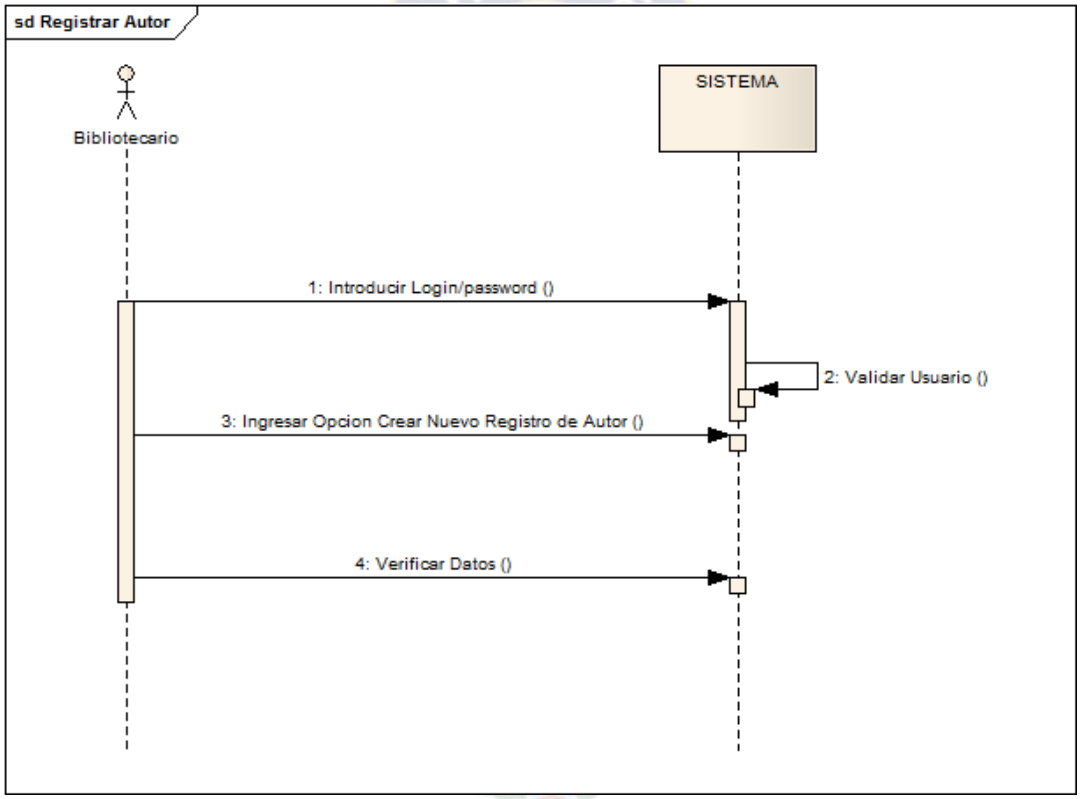


Figura 3.7 Diagrama de Secuencia: Registrar Autor
Fuente: Elaboración Propia

Contrato de Operaciones

Nombre:	Introducir login y password
Responsabilidades:	Despliega y confirma autenticación
Tipo:	Manual
Referencias:	Entrega de registro del autor
Precondición:	El usuario debe estar registrado en el sistema
Postcondición:	Proporciona el formulario de registro

Tabla 3.13 Contrato de Operaciones- Introducir login y password
Fuente: Elaboración propia

Nombre:	Crear nuevo registro de solicitud
Responsabilidades:	Despliega nuevo registro de solicitud
Tipo:	Manual
Referencias:	Registra los datos para almacenar en la base de datos
Precondición:	El autor debe tener un comprobante de la solicitud
Postcondición:	Proporciona una información correcta del autor

Tabla 3.14 Contrato de Operaciones- Crear nuevo registro de solicitud
Fuente: Elaboración propia

Nombre:	Verifica datos
Responsabilidades:	Se encarga de verificar los datos completos
Tipo:	Manual
Referencias:	Se verifica mediante el sistema los datos
Precondición:	EL sistema proporciona una base de datos actualizados
Postcondición:	Proporciona la actualización de los datos del autor

Tabla 3.15 Contrato de Operaciones- Verifica datos
Fuente: Elaboración propia

Registrar Documento

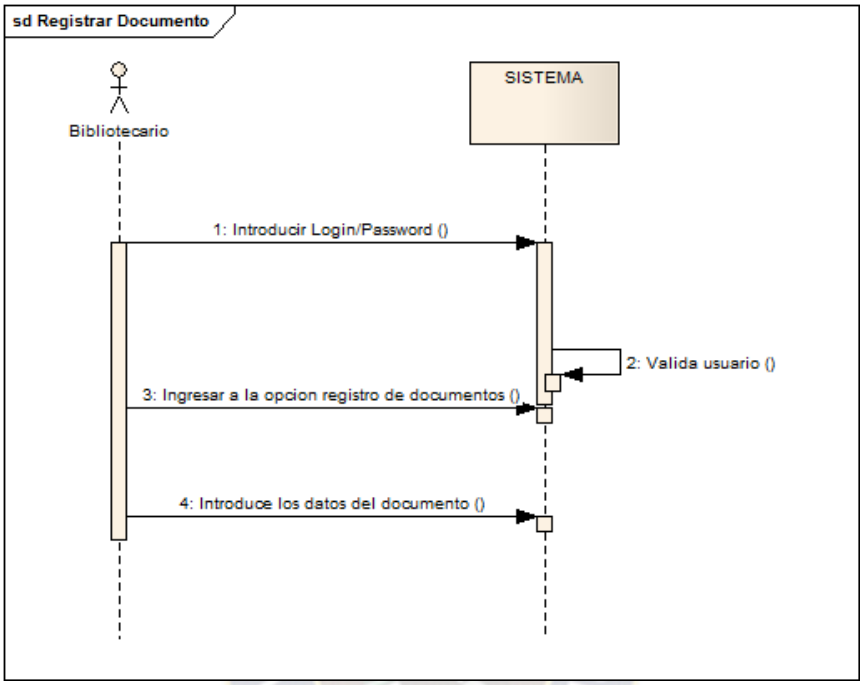


Figura 3.8 Diagrama de Secuencia: Registrar el Documento
Fuente: Elaboración Propia

Contrato de Operaciones

Nombre:	Introducir login y password
Responsabilidades:	Despliega y confirma autenticación
Tipo:	Manual
Referencias:	Entrega registro del autor
Precondición:	el bibliotecario debe estar registrado en el sistema
Postcondición:	Proporciona la proforma de registro

Tabla 3.16 Contrato de Operaciones- Introducir login y password
Fuente: Elaboración propia

Nombre:	Registra el Documento
Responsabilidades:	Despliega el formulario del nuevo registro de documento
Tipo:	Sistema
Referencias:	Registra los datos del documento en la base de datos
Precondición:	El autor del documento debe tener un comprobante
Postcondición:	Proporciona una información correcta del documento

Tabla 3.17 Contrato de Operaciones- Registra el Documento
Fuente: Elaboración propia

Nombre:	Realiza catalogación del documento en Dublin Core
Responsabilidades:	Se encarga de catalogar los datos completos del documento
Tipo:	Manual
Referencias:	Se verifica mediante el sistema datos completos y en norma
Precondición:	El sistema proporciona una base de datos actualizados
Postcondición:	Proporciona la actualización de los datos de la catalogación

Tabla 3.18 Contrato de Operaciones- Realiza catalogación del documento en Dublin Core
Fuente: Elaboración propia

3.4 FASE DE CONSTRUCCION

3.4.1 DIAGRAMA DE CONTENIDO

Para comprender el contexto del sistema el modelo especifica cómo se encuentra relacionados los contenidos del sistema, es decir, define la estructura de los datos que se encuentran alojados en el sitio web. A continuación se el contenido en la página web del sistema aplicando la metodología UWE.

Diagrama de Contenido

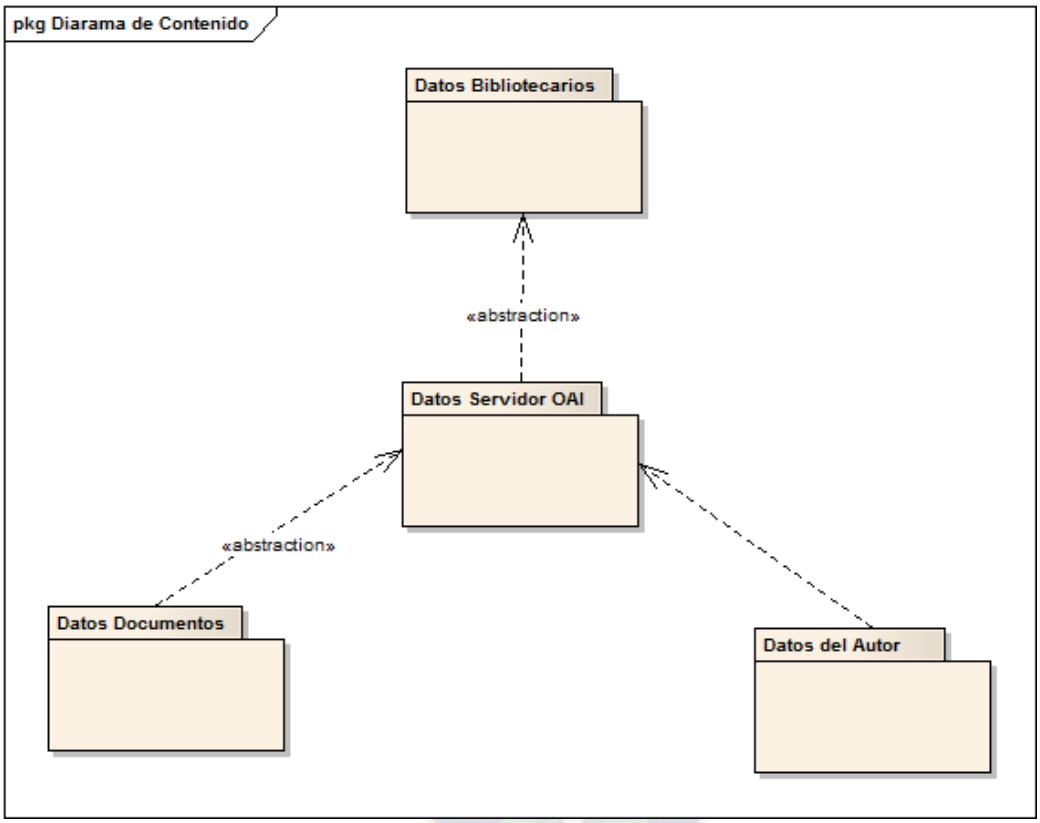


Figura 3.9 Diagrama de Contenido
Fuente: Elaboración Propia

3.4.2 DIAGRAMA DE NAVEGACIÓN

Este diagrama indica como el sistema de páginas web del sitio está relacionado internamente. Es decir cómo se enlazan los elementos de navegación.

Diagrama de navegación

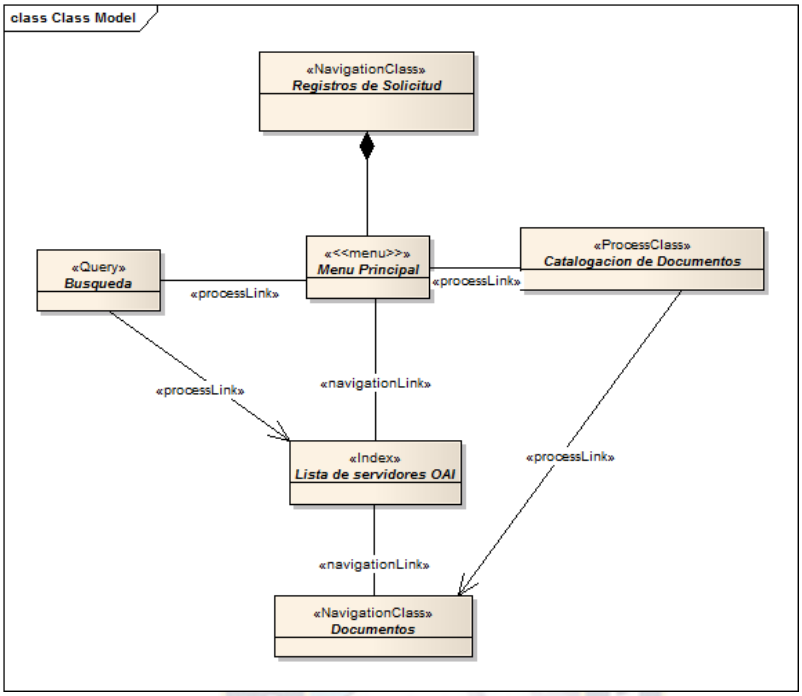


Figura 3.10 Diagrama de Navegación
Fuente: Elaboración Propia

3.4.3 DIAGRAMA DE ESTADO

Registrar Solicitud de Publicación

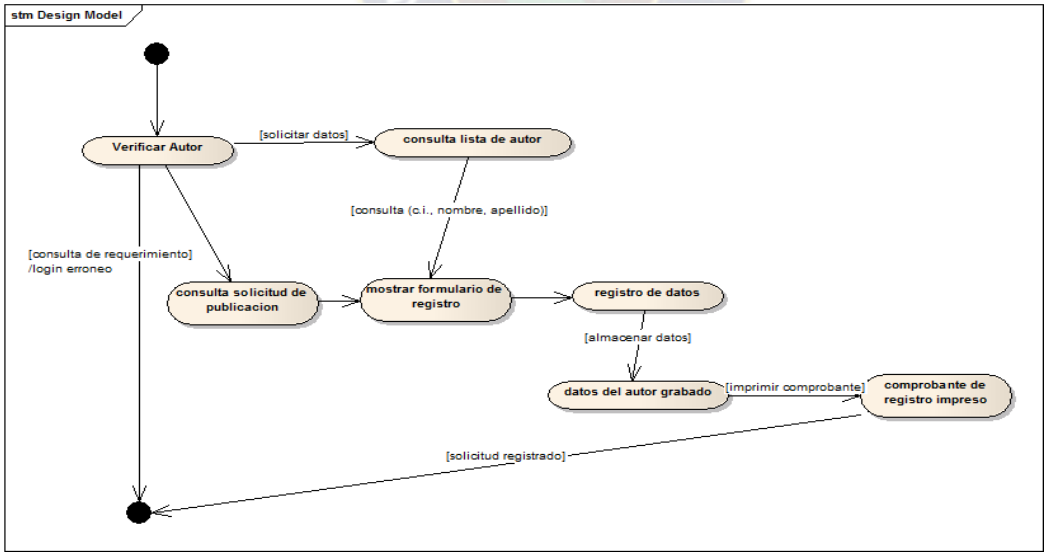


Figura 3.11 Diagrama de Estado: Registrar Solicitud de Publicación
Fuente: Elaboración Propia

Registrar Autor

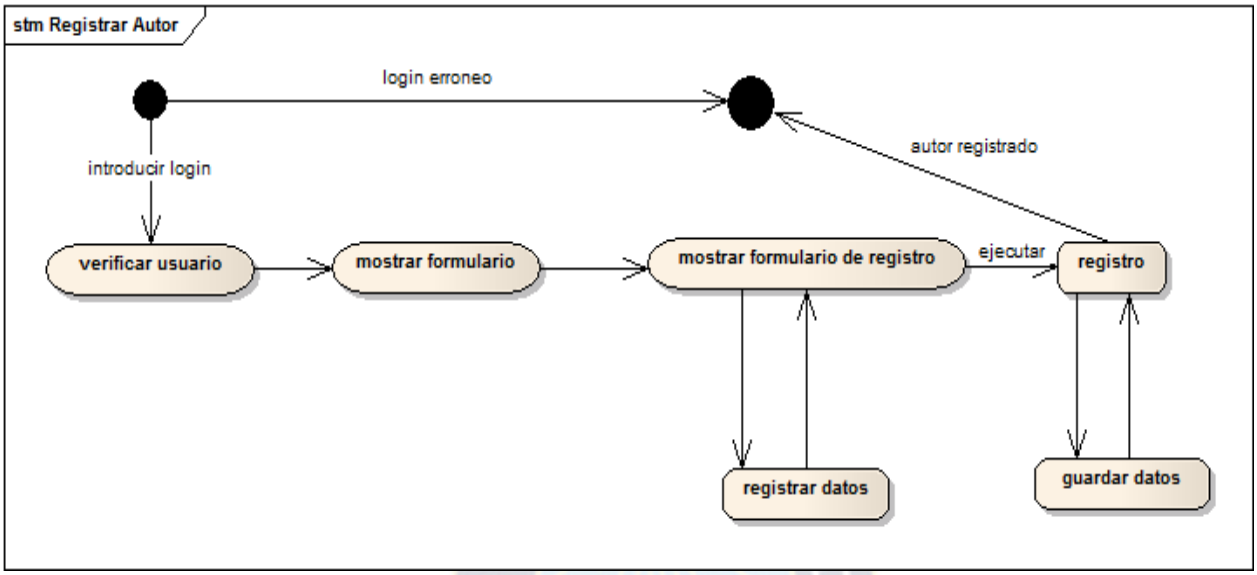


Figura 3.12 Diagrama de Estado: Registrar Autor
Fuente: Elaboración Propia

Catalogación en Dublin Core

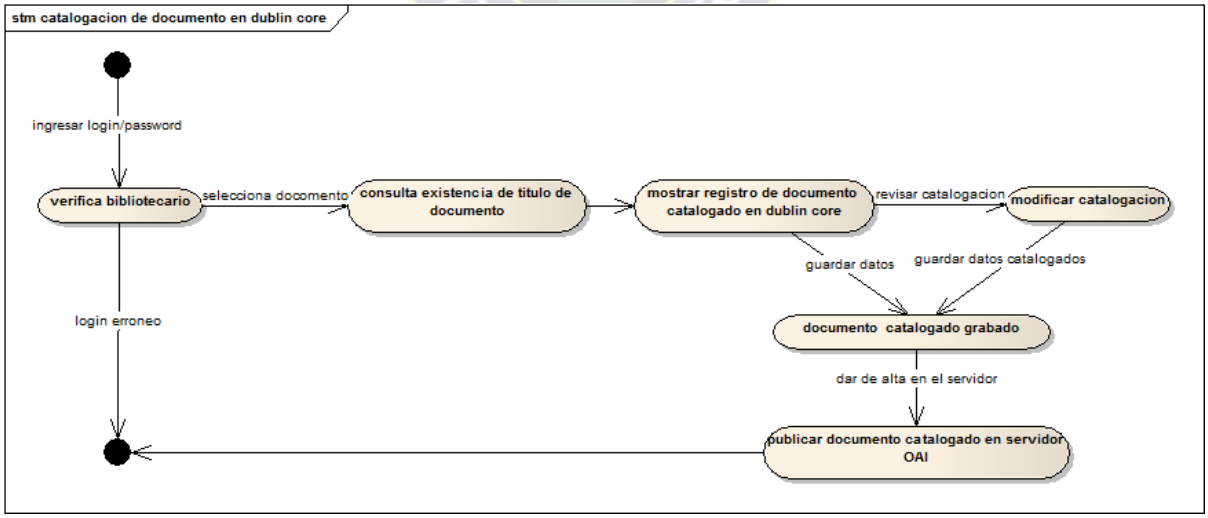


Figura 3.13 Diagrama de Estado: Catalogación del documento en Dublín Core
Fuente: Elaboración Propia

3.4.4 DIAGRAMA DE COMPONENTES

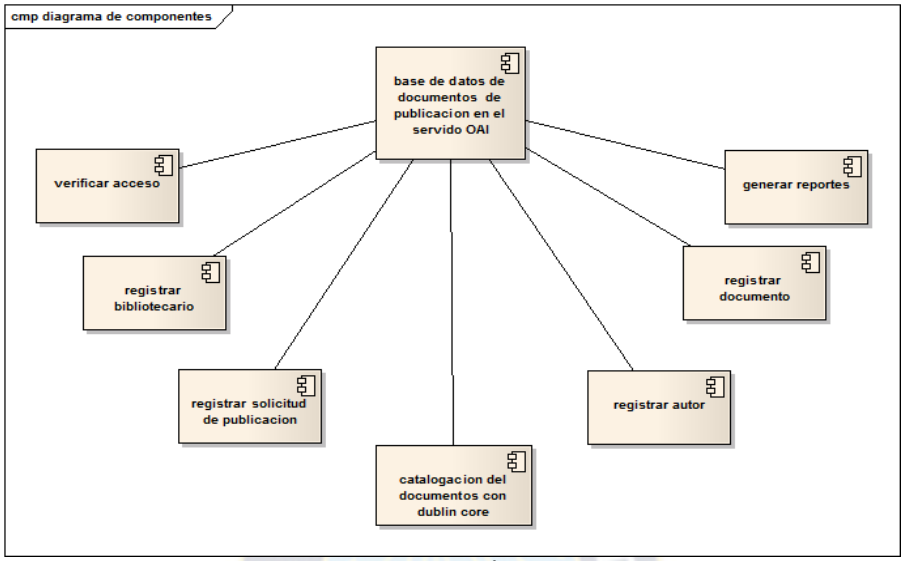


Figura 3.14 Diagrama de Componentes
Fuente: Elaboración Propia

3.4.5 DIAGRAMA DE DESPLIEGUE

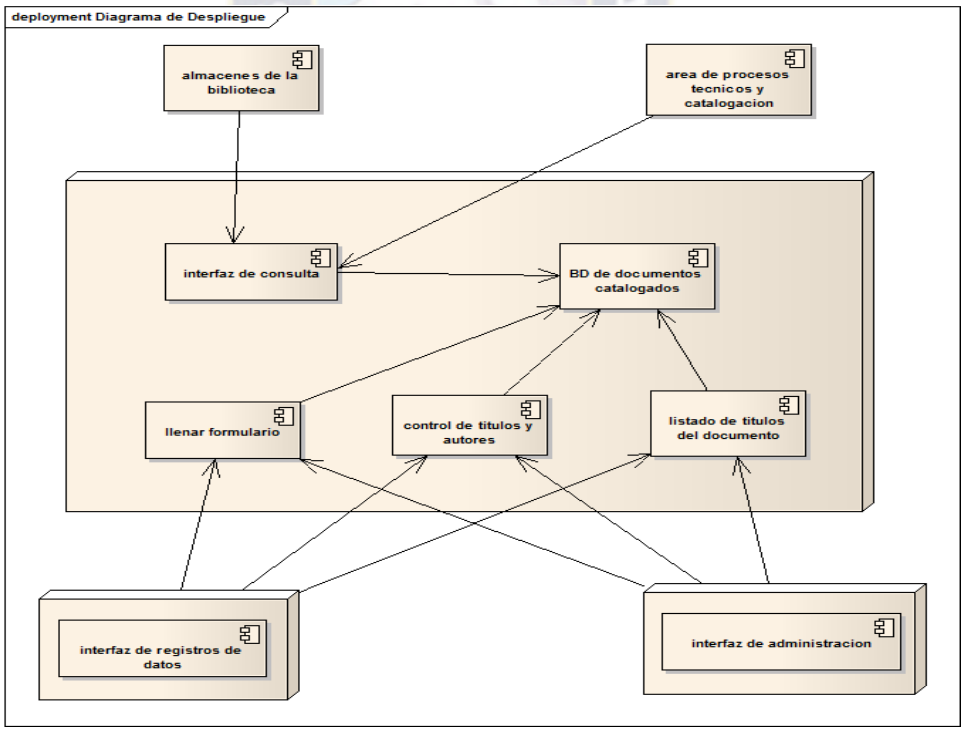


Figura 3.15 Diagrama de despliegue
Fuente: Elaboración Propia

3.4.6 ARQUITECTURA DEL SISTEMA

La arquitectura del sistema implica tres elementos básicos: plataformas de proceso programables en Java, separación entre la función y el proceso de implicaciones y la comunicación entre procesos. El diseño de la arquitectura diseñada para el Sistema es el de Cliente/Servidor, el cual consta de un Servidor de Base de datos PostgreSQL al cual se conectan todas las terminales mediante internet, todos los cosechadores de documentos de los servidores OAI-PMH de la Biblioteca Central tendrán acceso al sistema como se muestra en la siguiente figura.

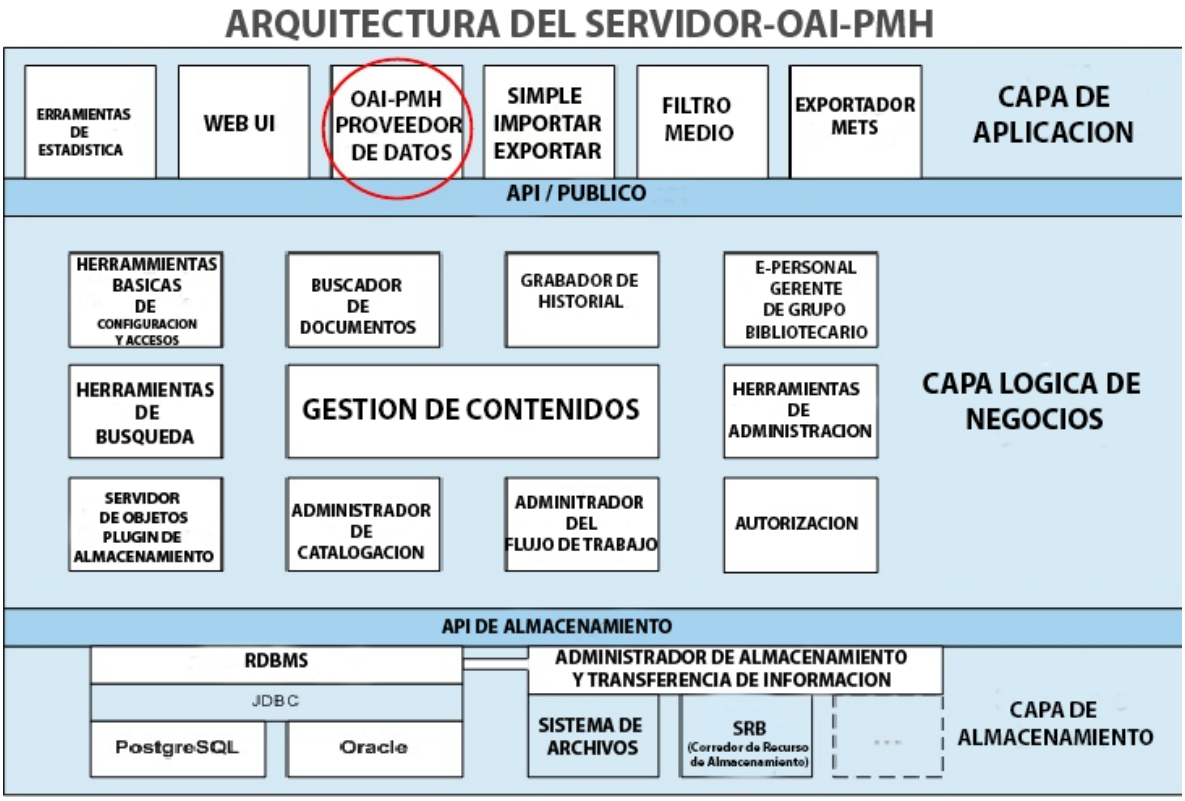


Figura 3.16 Arquitectura del Sistema Servidor OAI por capas
Fuente: Elaboración Propia

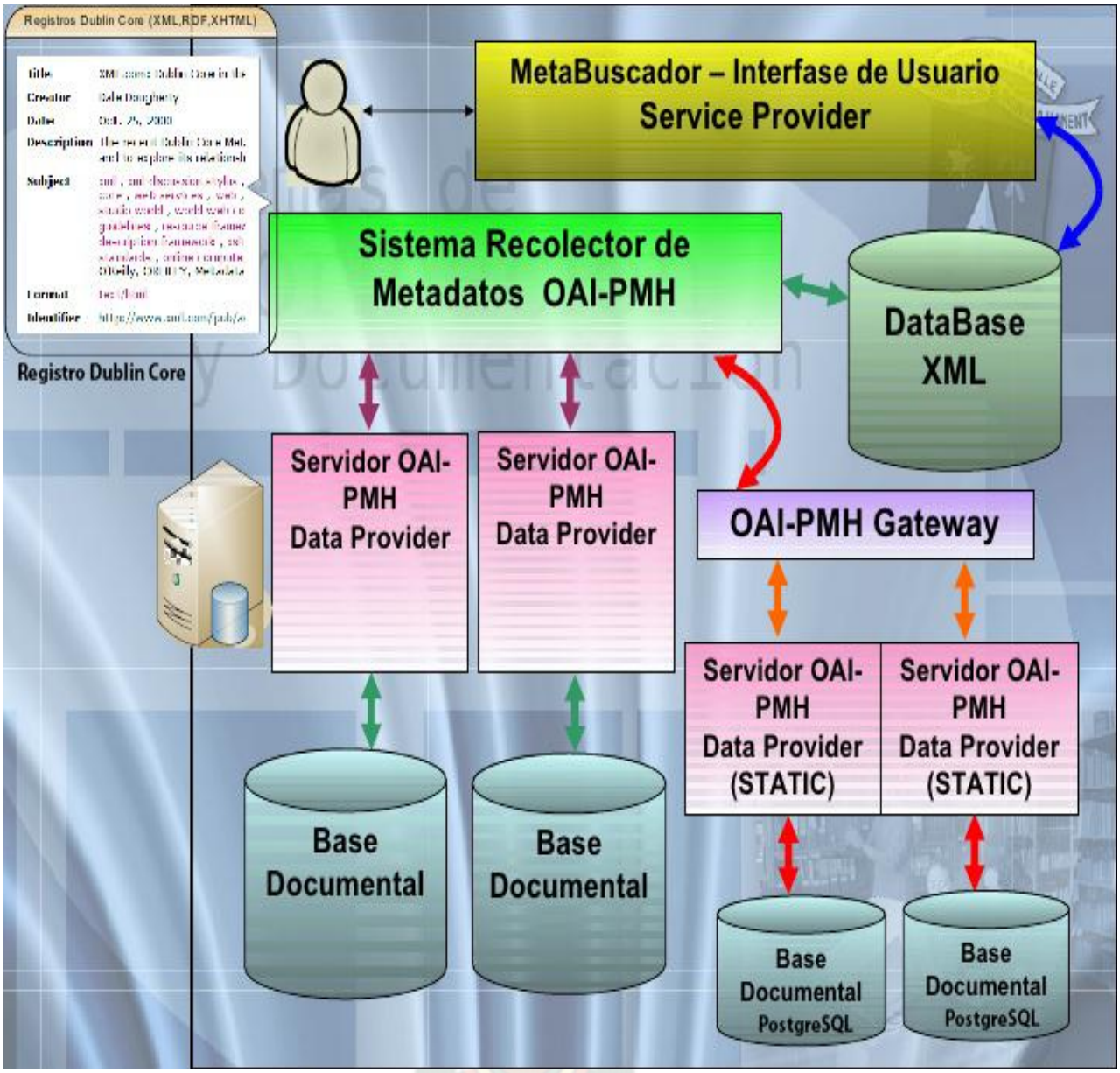


Figura 3.17 Arquitectura del Sistema Servidor OAI
Fuente: Elaboración ISTEC

3.4.7 DIAGRAMA DE CLASES

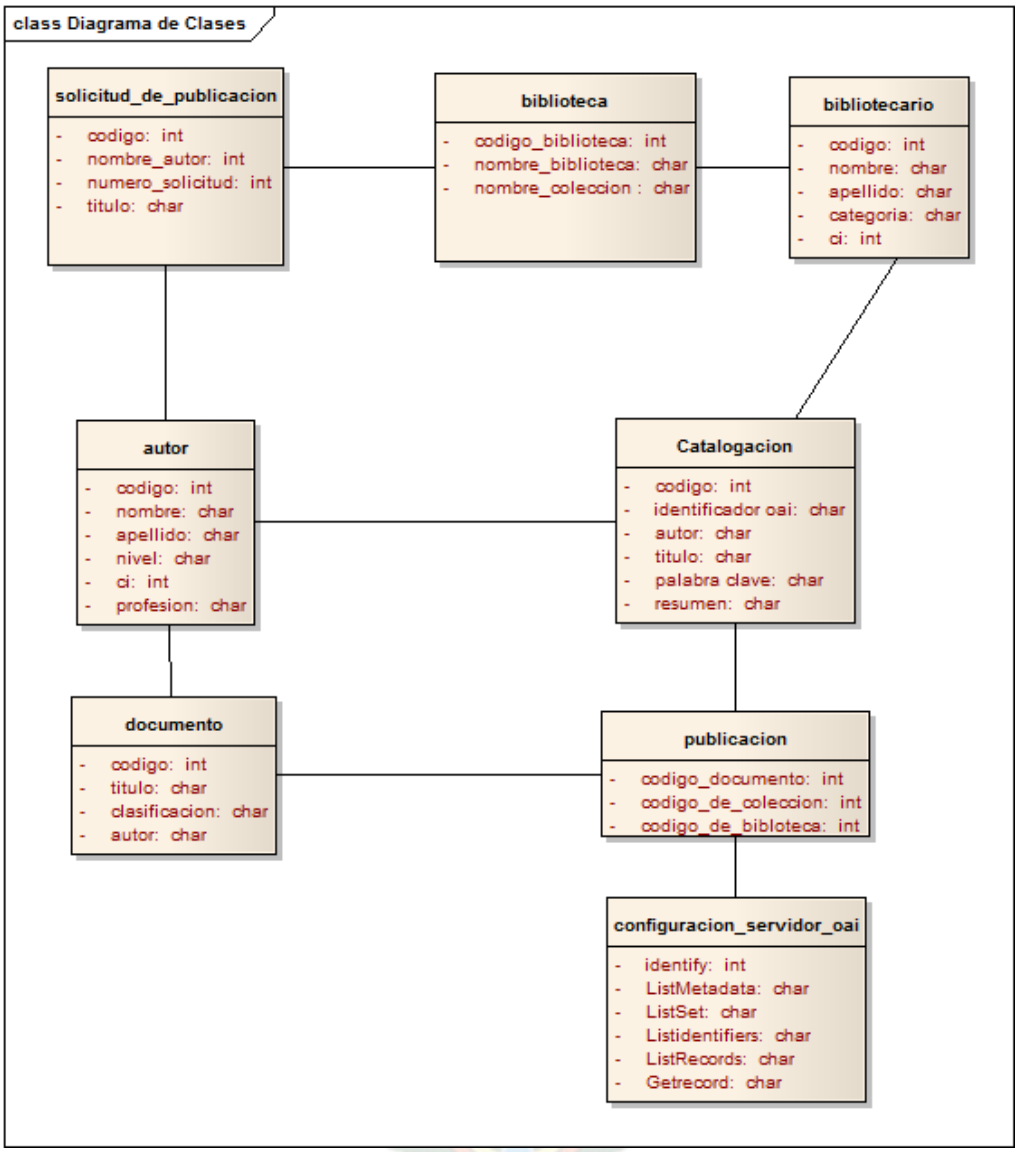


Figura 3.18 Diagrama de Clases
Fuente: Elaboración Propia

3.5 FASE DE IMPLEMENTACION

3.5.1 DISEÑO DE INTERFACES

En el diseño de interfaz de sistema el objetivo es implementar los requisitos desarrollados en el análisis, esto permite realizar las interfaces que son los prototipos o pantallas, la pantalla principal se encarga de autenticar al usuario.

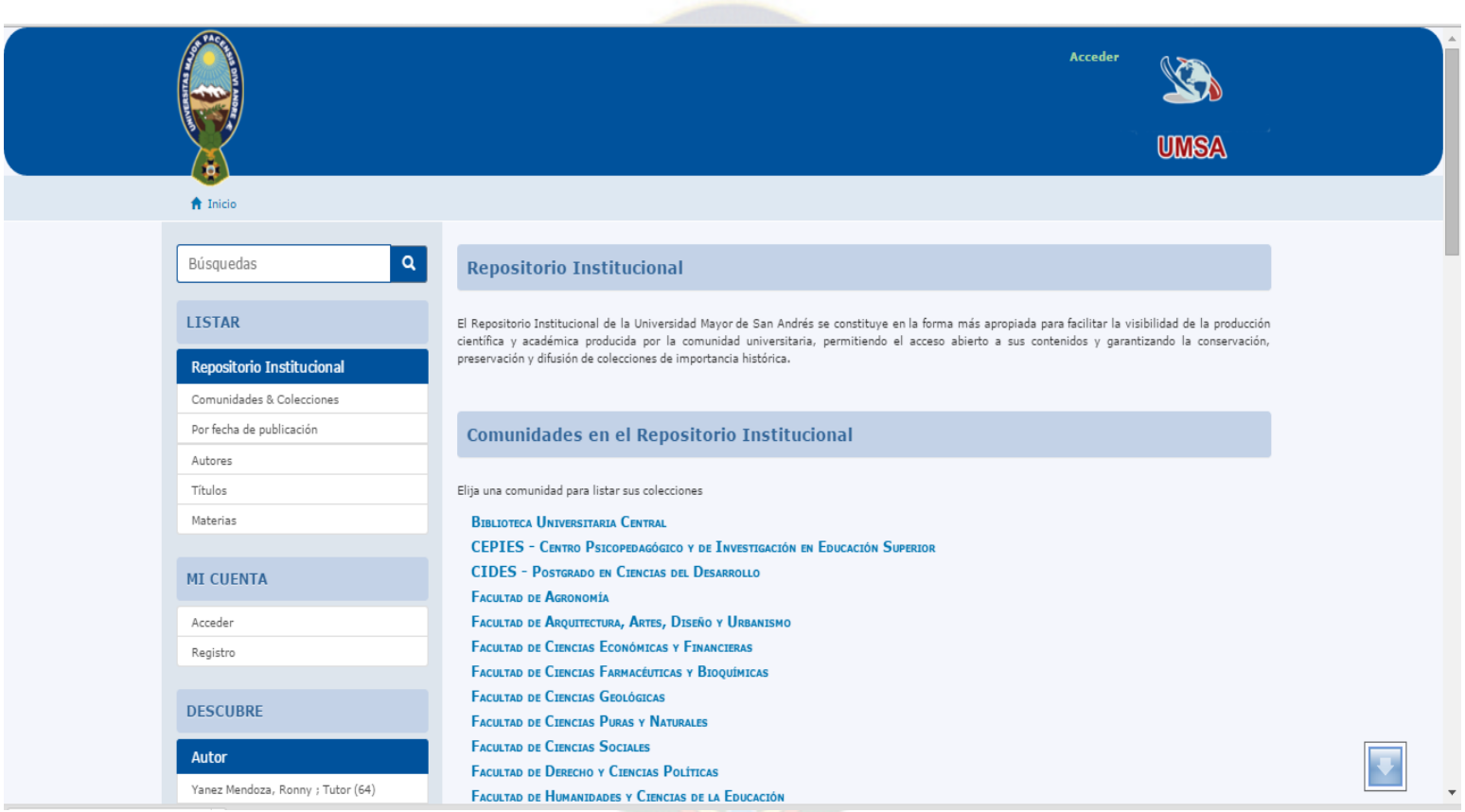


Figura 3.19 Interfaz del menú principal
Fuente: Elaboración Propia

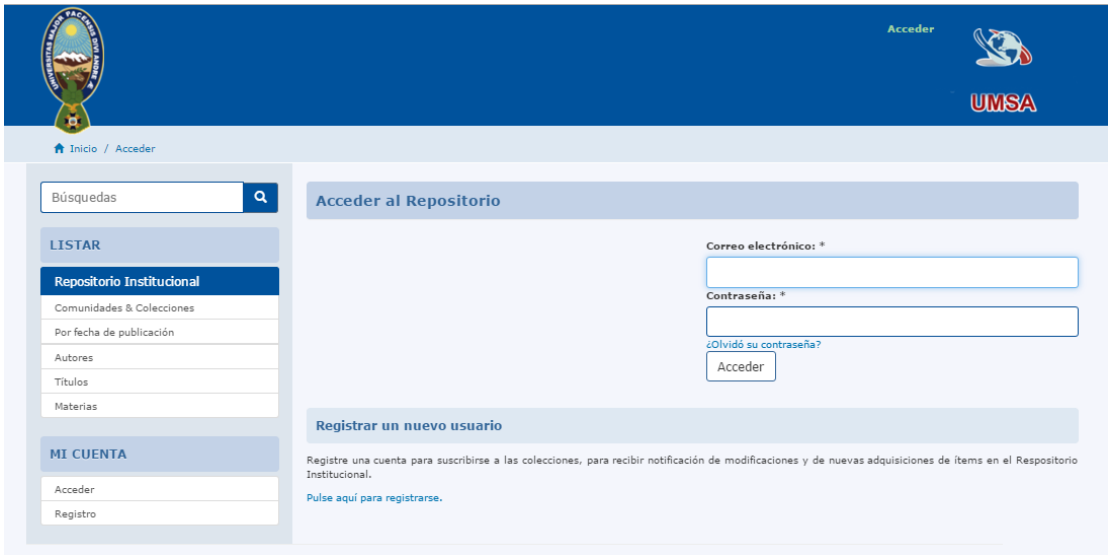


Figura 3.20 Interfaz de autenticación
Fuente: Elaboración Propia

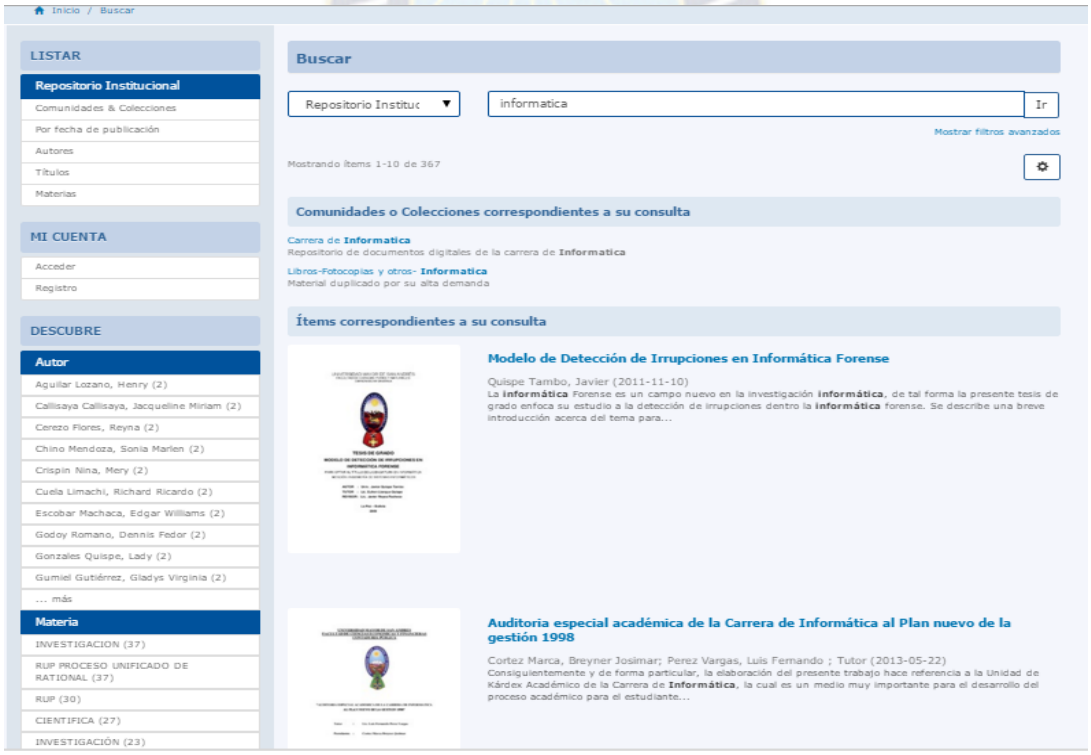


Figura 3.21 Interfaz de resultados
Fuente: Elaboración Propia

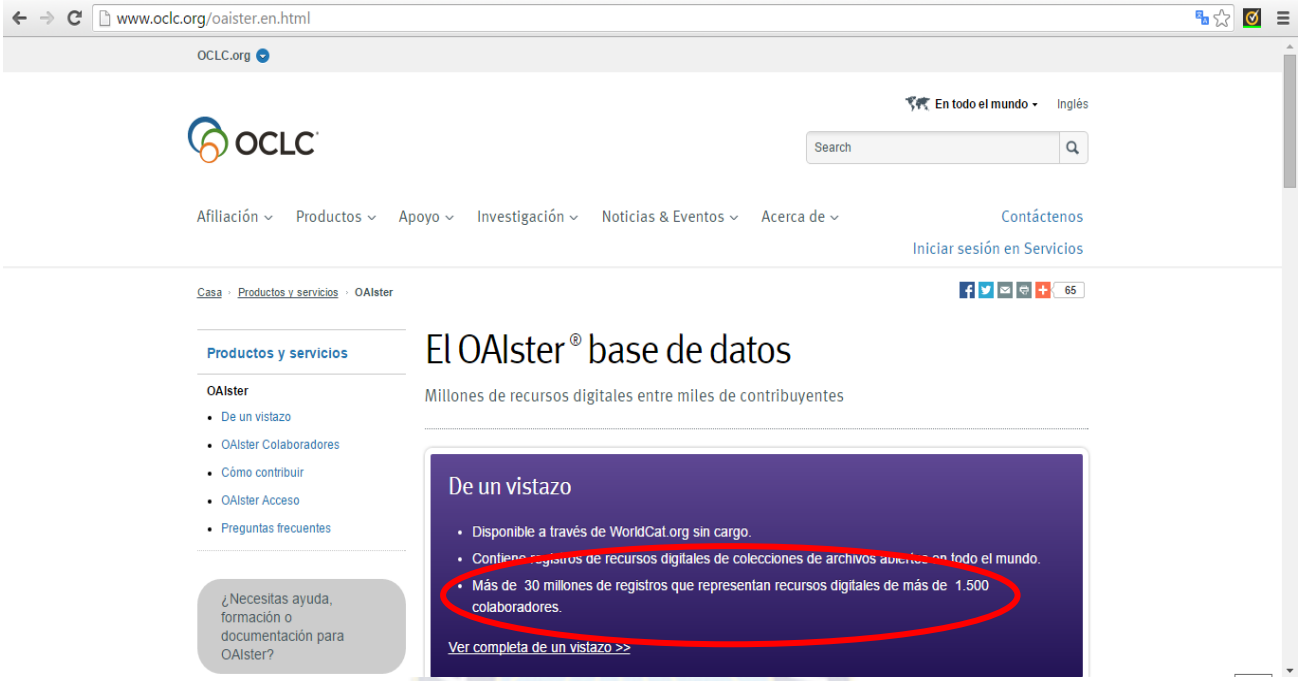


Figura 3.22 Interfaz OAISTER con 30 millones de documentos y 1500 instituciones
Fuente: Elaboración Propia

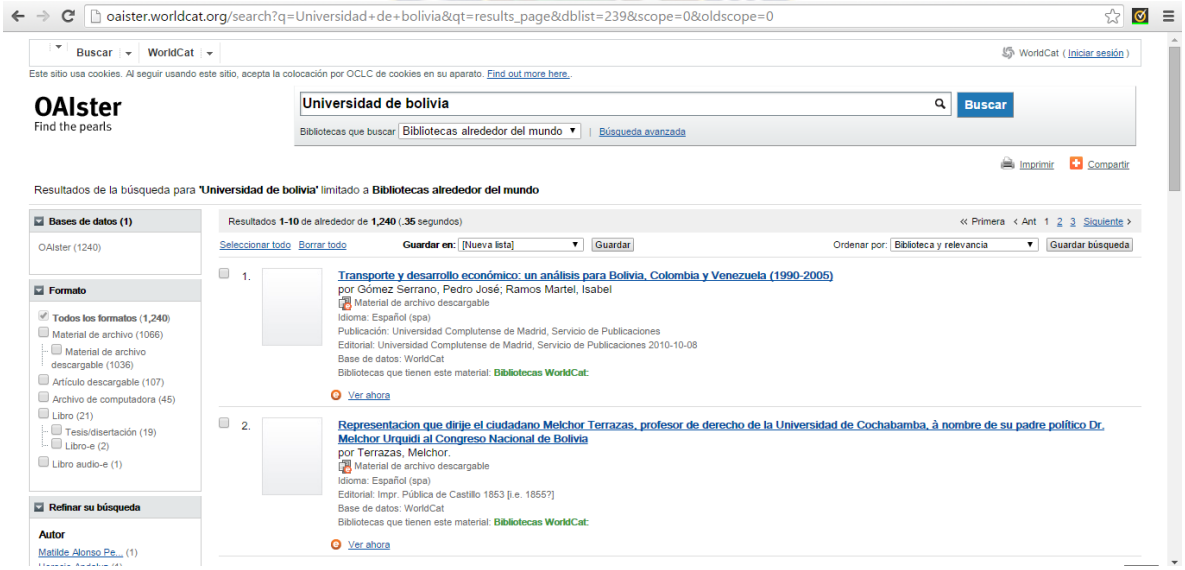


Figura 3.23 Interfaz de resultados OAISTER
Fuente: Elaboración Propia

CAPITULO IV

CALIDAD Y SEGURIDAD

4.1 METRICAS DE CALIDAD

A la conclusión del desarrollo del sistema es menester conocer la calidad del sistema siendo este un factor importante, para el funcionamiento de nuestro servidor de metadatos OAI-PMH y la publicación de documentos académicos.

Las aplicaciones de software centradas en la Web son cada vez más complejas y están creciendo rápidamente. Por lo tanto, este rápido crecimiento genera nuevos desafíos, como el diseñar y producir sistemas Web para diferentes dominios, necesidades y comportamientos de usuarios, teniendo a la calidad como objetivo esencial o, desde otro punto de vista, cómo medir, evaluar y potencialmente mejorar la calidad de los sitios Web[19].

4.2 METODOLOGÍA DE EVALUACIÓN DE CALIDAD DE SITIOS WEB

Para el tratamiento de Metodología de Evaluación de Calidad de Sitios Web (Web-site QEM), que incluye un conjunto de fases, actividades, productos, modelos y constructores de proceso que en el presente capitulo se irá desarrollando.

Si bien la metodología puede ser empleada en cualquier fase del ciclo de vida del producto, en este proyecto se empleará la metodología en la fase operativa, es decir cuando el sistema este implantado, culminado y en funcionamiento.

[19] [Olsina, 99]

4.2.1 DEFINICIÓN DE LAS METAS DE EVALUACIÓN Y SELECCIÓN DEL PERFIL

4.2.1.1 DEFINICIÓN DEL DOMINIO Y ENTE DE EVALUACIÓN

Podemos definir al dominio de la aplicación, desde el punto de vista de la evaluación, en este caso el dominio de la evaluación se centraría en el : "Sistema de Generación de Servidores de Metadatos con el Protocolo OAI-PMH para la Biblioteca Central de la Universidad Mayor de San Andrés", que consiste de un conjunto de entes a los que se le atribuyen propiedades (atributos, características) y manifiestan un comportamiento y se relacionan, este modelo de clases se puede observar en el modelo conceptual.

4.2.1.2 DEFINICIÓN DE LAS METAS DE EVALUACIÓN

En esta parte se debe definir y especificar claramente las metas y el alcance del proceso de evaluación. Una de las metas principales general de la evaluación de calidad del sistema Web, radica en comprender el grado de cumplimiento de un conjunto de características y subcaracterísitcas con respecto a los requerimientos de calidad establecidos.

Para el proceso de evaluación utilizado en el presente proyecto, la meta principal consiste en: "Comprender a la calidad global de un sitio Web que publica documentos académicos científicos a través de un servidor de metadatos OAI-PMH desde el punto de vista del autor que publica y el visitante general".

4.2.1.3 SELECCIÓN DEL PERFIL DE USUARIO

La formulación de metas y consecuentemente la relativa importancia de las características y atributos de calidad, varían conforme al perfil de usuario seleccionado.

El estándar ISO [ISO/IEC 9126], considera tres perfiles de usuario, a un alto nivel de abstracción para dominios Web, a saber: visitantes, autores, y bibliotecarios.

Para la definición de los requerimientos de calidad para el dominio, el perfil de usuario seleccionado es el de visitante general con el fin de evaluar la calidad global del sistema web. Este perfil se puede extender a usuarios casuales e intencionales, y dentro de éste último a usuarios registrado y a usuarios no registrados.

4.2.2 DEFINICIÓN Y ESPECIFICACIÓN DE LOS REQUERIMIENTOS DE CALIDAD

En esta fase se debe especificar las características y atributos de calidad y agruparlas en un árbol de requerimientos, conforme a las metas, el dominio, y el perfil de usuario seleccionados.

Es importante destacar que la calidad de un sistema Web se la puede cuantificar mediante la apropiada descomposición de algunas o todas las características de un conjunto prescrito, esto es, a partir de la medición directa de un conjunto de atributos (derivados a partir de esas características) y por medio de un proceso de agregación y cómputo de preferencias elementales. Para este conjunto prescrito de características el proceso se basa en las características de calidad mencionadas en la estándar ISO [130/IEC 9126] los cuales son: usabilidad, funcionalidad, confiabilidad, eficiencia, portabilidad, y mantenibilidad[20].

Para el perfil de usuario Visitante General y para la meta de evaluación anteriormente citada, de las seis características prescriptas en el estándar ISO solamente cuatro de ellas resultan de relevancia para el proceso de evaluación. Las características de significativa importancia son: usabilidad, funcionalidad, confiabilidad y eficiencia.

En la tabla 4.1 se puede apreciar el árbol de requerimientos de calidad del sistema Web donde se observa las características y atributos de calidad. Para la construcción de este árbol se tomó como base las características, subcaracterísitcas y atributos citados por Olsina, que es un modelo básicamente estándar para utilizarlo en distintos dominios.

[20] [Olsina 99]

1. Usabilidad	2. Funcionalidad
1.1 Comprensibilidad Global del Sitio 1.1.1 Esquema de Organización Global 1.1.1.1 Mapa del Sitio 1.1.1.2 Tabla de Contenidos 1.1.2 Calidad en el Sistema de Etiquetado 1.1.3 Visita Guiada Orientada al Visitante 1.2 Mecanismos de Ayuda y Retroalimentación en línea 1.2.1 Calidad de la Ayuda 1.2.1.1 Ayuda de la Búsqueda 1.2.2 Indicador de Ultima Actualización 1.2.2.1 Global (de todo el sitio Web) 1.2.2.2 Restringido (por subsitio o página) 1.2.3 Directorio de Direcciones 1.2.3.1 Directorio E-mail 1.2.3.2 Directorio TE-Fax 1.2.4 Facilidad FAQ 1.2.5 Retroalimentación 1.2.5.1 Cuestionarios/Encuestas 1.2.5.2 Comentarios/Sugerencias 1.3 Aspectos de Interfaces y Estéticos 1.3.1 Coercividad al Agrupar los Objetos de Control Principales 1.3.2 Permanencia y Estabilidad en la Presentación de los Controles Principales 1.3.2.1 Permanencia de Controles Directos 1.3.2.2 Permanencia de Controles Indirectos 1.3.2.3 Estabilidad 1.3.3 Aspectos de Estilo 1.3.3.1 Uniformidad en el Color de Enlaces 1.3.3.2 Uniformidad en el Estilo Global 1.4 Misceláneas 1.4.1 Soporte a Lenguaje Extranjero 1.4.2 Atributo "Qué es lo Nuevo" 1.4.3 Indicador de Resolución de Pantalla	2.1 Aspectos de Búsqueda y Recuperación 2.1.1 Mecanismo de Búsqueda en el Sitio Web 2.1.1.1 Búsqueda Restringida 2.1.1.1.1 Búsqueda de Documentos 2.1.1.2 Búsqueda Global 2.2 Aspectos de Navegación y Exploración 2.2.1 Navegabilidad 2.2.1.1 Orientación 2.2.1.1.1 Indicador del Camino 2.2.1.1.2 Etiqueta de la Posición Actual 2.2.1.2 Promedio de Enlaces por Página 2.2.2 Objetos de Control Navegacional 2.2.2.1 Permanencia y Estabilidad en la Presentación de los Controles Contextuales (Subsitio) 2.2.2.1.1 Permanencia de los Controles Contextuales 2.2.2.1.2 Estabilidad 2.2.2.2 Nivel de Desplazamiento 2.2.2.2.1 Desplazamiento Vertical 2.2.2.2.2 Desplazamiento Horizontal 2.2.3 Predicción Navegaciones 2.2.3.1 Enlace con Título (enlace con textoexplicatorio) 2.2.3.2 Calidad de la Frase del Enlace 2.3 Aspectos del Dominio orientados al Investigador 2.3.1 Relevancia de Contenido 2.3.1.1 Información de la colección 2.3.1.2 Información de solicitud de registro de publicación 2.3.1.2.1 Información de los Requerimientos de publicación 2.3.1.2.2 Formulario para Rellenar la solicitud de publicación 2.3.1.3 Información de los autores 2.3.1.3.1 información de los autores por

	publicación 2.3.1.3.2 Buscador de los autores 2.3.1.3.3 Información pública de los autores y documentos 2.3.1.3.4 Información del servidor OAI de la Biblioteca Central 2.3.1.4 Información de Servicios de publicación para los autores 2.3.1.4.1 índice de documentos 2.3.1.4.2 Información de nuevas publicaciones 2.3.1.4.3 Información de proveedores de metadatos académicos 2.3.1.4.4 Información de nuevas publicaciones de otros servidores OAI 2.3.2 Servicios On-line 2.3.2.1 Información de nuevas solicitudes 2.3.2.2 Información de las colecciones 2.3.2.3 Información de registro de solicitudes.
3. Confiabilidad	**4. Eficiencia**
.1 No Deficiencia 3.1.1 Errores de Enlaces 3.1.1.1 Enlaces Rotos 3.1.1.2 Enlaces Inválidos 3.1.1.3 Enlaces no Implementados 3.1.2 Errores o Deficiencias 3.1.2.1 Deficiencias o cualidades ausentes debido a diferentes navegadores (browsers) 3.1.2.2 Deficiencias o resultados inesperados independientes de browsers errores de búsqueda imprevistos,(p.ej. errores de búsqueda de improvistos , deficiencias con marcos (frames), etc.) 3.1.2.3 Nodos Destinos (inesperadamente) en Construcción 3.1.2.4 Nodos Web Muertos (sin enlaces de retorno)	4.1 Performancia 4.1.1 Páginas de Acceso Rápido 4.2 Accesibilidad 4.2.1 Accesibilidad de Información 4.2.1.1 Soporte a Versión sólo Texto 4.2.1.2 Legibilidad al desactivar la Propiedad Imagen del Browser 4.2.1.2.1 Imagen con Título 4.2.1.2.2 Legibilidad Global 4.2.2 Accesibilidad de ventanas 4.2.2.1 Numero de Vistas considerando Marcos. 4.2.2.2 Vistas sin Marcos

Tabla 4.1 Árbol de requerimientos de calidad

Fuente Elaboración propia

4.2.3 DEFINICIÓN E IMPLEMENTACIÓN DE LA EVALUACIÓN ELEMENTAL

Esta fase comprende actividades, modelos, técnicas, heurísticas y herramientas para determinar criterios de evaluación para cada atributo cuantificable y realizar el proceso de medición. Se consideran diferentes tipos de criterios de calidad elemental, escalas (y representación gráfica como escala de preferencia), valores y rangos críticos, y funciones para determinar la preferencia elemental, entre otros asuntos. Una vez definidos y consensuados los criterios para medir cada atributo, se debe ejecutar el proceso de medición, es decir, la recolección de datos, el cómputo de las variables y las preferencias elementales, y la documentación de los resultados[21].

4.2.3.1 CRITERIO DE EVALUACIÓN ELEMENTAL PARA ATRIBUTOS

A partir del árbol de requerimientos para cada atributo cuantificable Ai (u hoja del árbol) debemos asociar y determinar una variable Xi, que tomará un valor real a partir de un proceso de medición. Además para cada variable Xi, computada, por medio de un criterio elemental IEi. Este resultado final, elemental, se puede interpretar como el grado o porcentaje del requerimiento del usuario satisfecho para el atributo Ai.

4.2.3.1.1 ESPECIFICACIÓN PARA REPRESENTAR A LOS ATRIBUTOS DE CALIDAD

A seguir, presentamos un marco de especificación descriptivo y jerárquico para representar a los atributos cuantificables a partir de un árbol de requerimientos. Principalmente, se describen atributos con distintos tipos de criterios de preferencia elemental o mecanismos de recolección de datos.

Para esta representación se hace uso de las plantillas de características y Atributos de Calidad. Según Olsina y sus estudios.

[21] [Olsina, 1999]

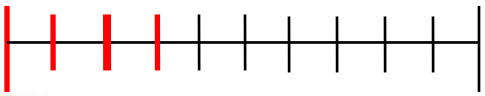
Título: Mapa del Sitio	**Código: 1.1.1.1**	**Tipo: Atributo**
Características de más Alto Nivel: Usabilidad		
Super-caracteristica: Esquema de Organización Global		
Definición/Comentario: Un mapa del sitio es una representación con componentes gráficos, que muestra la estructura o arquitectura global (a menudo jerárquica) del sitio Web como un todo.		
Tipo de criterios elemental: es un criterio binario, discreto y absoluto; solo se pregunta si está disponible (1) o si no está disponible (0).	**Escala de Preferencias** :  0% 40% 60% 100%	
Tipo de Recolección de Datos: Manual Observacional		

Tabla 4.2 Plantilla del atributo de Calidad – Mapa del sitio
Fuente: Elaboración Propia

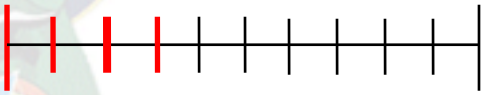
Título: Facilidad FAQ	**Código: 1.2.4**	**Tipo: Atributo**
Características de más Alto Nivel: Usabilidad		
Super-caracteristica: Mecanismos de ayuda y retroalimentación en línea.		
Definición/Comentario: Este atribulo representa a un conjunto de preguntas (agrupadas y enlazadas) que se realizan con mayor frecuencia, y que están ya publicadas en el sitio con sus respectivas respuestas. A su vez las respuestas pueden estar enlazadas a otros contenidos. Esto favorece al mecanismo de aprendizaje y/o ayuda, evitando potencialmente la demora cognitiva de los visitantes.		
Tipo de criterios elemental: es un criterio multi-nivel discreto y absoluto definido como subconjunto; en donde sí se evalúa la disponibilidad y tipo de mecanismo entonces; 0= no disponible; 1= FAQ básico, permite acceder a las respuestas, generalmente no enlazadas a otros contenidos; 2=1 + permite acceder a las respuestas generalmente enlazadas otros contenidos.	**Escala de Preferencias** :  0% 40% 60% 100%	
Tipo de Recolección de Datos: Manual Observacional		

Tabla 4.3 Plantilla del atributo de Calidad – Factibilidad FAQ
Fuente: Elaboración Propia

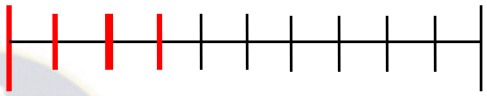
Título: Soporte a lenguaje a Facilidad FAQ	**Código: 1.2.4**	**Tipo: Atributo**
Características de más Alto Nivel: Usabilidad		
Super-caracteristica: Misceláneas		
Definición/Comentario: Este atribulo modela el número de lenguajes extranjeros soportados por un sitio, Además especifica el nivel de soporte para cada lenguaje no se computa obviamente el lenguaje nativo del sitio Web, como lenguaje extranjero.		
Tipo de criterios elemental: es un criterio multi-variable, continuo y absoluto.	**Escala de Preferencias** :  0% 40% 60% 100%	
Tipo de Recolección de Datos: Manual Observacional		

Tabla 4.4 Plantilla del atributo de Calidad – Soporte a lenguaje extranjero
Fuente: Elaboración Propia

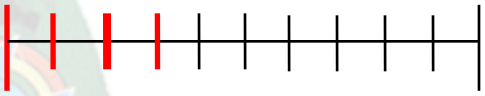
Título: Búsqueda de Documentos	**Código: 2.1.1.1.1**	**Tipo: Atributo**
Características de más Alto Nivel: Funcionalidad		
Super-caracteristica: Búsqueda restringida		
Definición/Comentario: Algunas veces, tiene sentido brindar al usuario con la facilidad de búsqueda restringida a un subsitio o parte de un sitio, debido a que el mismo es altamente cohercitivo o distintivo del resto de la información del sitio Web global.		
Tipo de criterios elemental: es un criterio multi-variable, continuo y absoluto, definido como subconjunto. Podemos decir que; 0 = no disponible mecanismos de búsqueda restringija ; 1= búsqueda básica; mecanismo de búsqueda por nombre apellido ;2=1+ búsqueda extendida o avanzada; mecanismo de búsqueda (con filtros).	**Escala de Preferencias** :  0% 40% 60% 100%	
Tipo de Recolección de Datos: Manual Observacional		

Tabla 4.5 Plantilla del atributo de Calidad – Búsqueda de Documentos
Fuente: Elaboración Propia

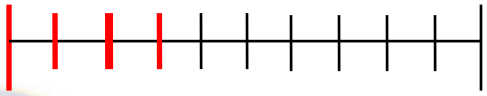
Título: Indicador del Camino	**Código: 2.2.1.1.1**	**Tipo: Atributo**
Características de más Alto Nivel: Funcionalidad		
Super-caracteristica: Búsqueda restringida		
Definición/Comentario: Los usuarios de la Web , al Navegar, debe tener pistas visuales (con elementos de diseño consistentes), que les indique con presicion en donde se encuentran posicionados , dentro de la estructura del espacio de información del sitio.		
Tipo de criterios elemental: es un criterio multi-nivel y absoluto; en donde si se observa la presencia del camino (en las paginas del sitio), ya sea en modo grafico o en modo textual, entonces; o= no disponible; 1= parcialmente disponible (sugerido por controles sombreados o texto); 2= totalmente disponible.	**Escala de Preferencias** :  0% 40% 60% 100%	
Tipo de Recolección de Datos: Manual Observacional		

Tabla 4.6 Plantilla del atributo de Calidad – Indicador del Camino
Fuente: Elaboración Propia

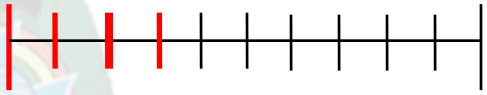
Título: Promedio de enlaces de pagina	**Código: 2.2.1.2**	**Tipo: Atributo**
Características de más Alto Nivel: Funcionalidad		
Super-caracteristica: Navegabilidad		
Definición/Comentario: Cada página o nodo tiene generalmente dos tipos principales de enlaces (o puntos de partida): enlaces semánticos (distribuidos en la información de las páginas del sitio o contenido), y enlaces estructurales (a cerca de controles principales y navegacionales).		
Tipo de criterios elemental: es un criterio multi-nivel, discreto y absoluto en donde entre 10 y 14 anchors en promedio, se corresponde a una preferencia de 100% 9 o 15 produce el 90%; 8 o 18 produce el 80 %; 7 o 17 produce el 60% 8 o 18 produce el 50% ; 5 o 19 produce el 30%; y < = 4 o >=20 anchors en promedio , produce el 0%.	**Escala de Preferencias** :  0% 40% 60% 100%	
Tipo de Recolección de Datos: Automatizado		

Tabla 4.7 Plantilla del atributo de Calidad – Promedio de enlace de pagina
Fuente: Elaboración Propia

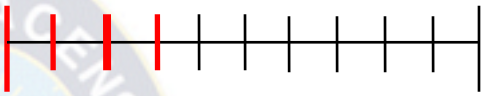
Título: Información de los autores por publicación	**Código: 2.3.1.3.1**	**Tipo: Atributo**
Características de más Alto Nivel: Funcionalidad		
Super-caracteristica: Información de los autores		
Definición/Comentario: Este mecanismo generalmente implementado en el subsitio (no en la página principal), permite estructurar por medio de enlaces a través de un ordenamiento de los autores, la navegación al contenido de las mismas.		
Tipo de criterios elemental: es un criterio binario, discreto y absoluto : solo se pregunta si está disponible (1) o si no está disponible (0).	**Escala de Preferencias** :  0% 40% 60% 100%	
Tipo de Recolección de Datos: Manual Observacional		

Tabla 4.8 Plantilla del atributo de Calidad – Información de los autores por publicación
Fuente: Elaboración Propia

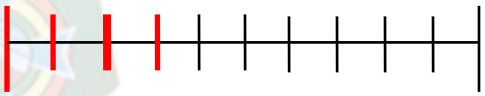
Título: Enlaces rotos	**Código: 3.1.1.1**	**Tipo: Atributo**
Características de más Alto Nivel: Confiabilidad		
Super-caracteristica: Errores de Enlaces		
Definición/Comentario: Este atributo representa básicamente a los enlaces encontrados que conducen a nodos destinos ausentes (también llamados enlaces ausentes o pendientes).		
Tipo de criterios elemental: es un criterio binario de variable normalizada, continuo y absoluto; en donde el BL=Numero de enlaces rotos encontrados, TL = número total de enlaces del sitio. La fórmula para computar la variable es: X=100-(BL*100/TL)*10; donde, si X<0 entonces X=0.	**Escala de Preferencias** :  0% 40% 60% 100%	
Tipo de Recolección de Datos: Automatizado		

Tabla 4.9 Plantilla del atributo de Calidad – Enlaces Rotos
Fuente: Elaboración Propia

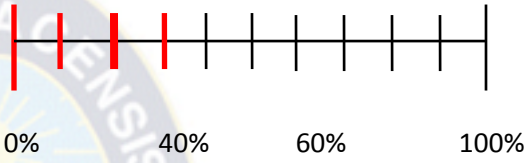
Título: Pagina de Acceso Rapido	**Código**: 4.1.1	**Tipo**: Atributo
Características de más Alto Nivel: Eficiencia		
Super-caracteristica: Performancia		
Definición/Comentario: Para este atributo, se mide el tamaño de todas las paginas (estáticas) del sitio Web considerando todos sus componentes gráficos, tabulares y textuales. El tamaño de cada página se especifica como una función del tiempo de espera y de la velocidad mínima establecida para una línea de comunicación dada espera y de la velocidad mínima establecida para una línea de comunicación dada.		
Tipo de criterios elemental: es un criterio multi-variable, continuo y absoluto.	**Escala de Preferencias** : 	
Tipo de Recolección de Datos: Automatizado		

Tabla 4.10 Plantilla del atributo de Calidad – Paginas de Acceso Rápido
Fuente: Elaboración Propia

4.2 .3.1.2 IMPLEMENTACIÓN DE LAS MEDICIONES ELEMENTALES

Una vez que los criterios de calidad para los atributos han sido definidos y acordados, es preciso recolectar los datos y computar las variables y preferencias elementales.

Muchos de los valores para las variables de los atributos especificados, fueron obtenidos manual u observacionalmente. Sin embargo, la recolección de datos automática es en muchos casos el único mecanismo para obtener resultados de un modo confiable y efectivo.

A continuación se detalla y se describen los atributos cuantificables obtenidos para las referencias de calidad elemental, para el sistema Web.

Código	Código	Definición	Tipo de Criterio Elemental	IE%
1 .1 .1 .1	Mapa del Sitio	Un mapa del sitio es una representación con componentes gráficos, que muestra la estructura o arquitectura global del sitio Web como un todo	Criterio binario, discreto y absoluto	100
1.1.1.2	Tabla de Contenidos	Es un mecanismo, disponible generalmente en la página principal, que permite estructurar el contenido de todo el sitio.	Criterio binario, discreto y absoluto	0
1.1.2	Calidad en el Sistema de Etiquetado	Se Refiere a las etiquetas que debe tener en la pagina, como ser enlaces, fotos, títulos, etc	Criterio de Preferencia Directa	80
1.1.3	Visita Guiada Orientada al Visitante	Permite recorrer páginas de un sitio en forma estructurada. Es un contexto de navegación arbitrario.	Criterio binario, discreto y absoluto	O
1 .2.1 .1	Ayuda de la Búsqueda	La ayuda explicativa debe estar basada en la riqueza, concisión, y oportunidad del texto. Y debe estar relacionada al mecanismo de búsqueda.	Criterio Muftí-nivel discreto, absoluto	60
1.2.2.1	Ultima Actualización Global (de todo el sitio Web)	Permite conocer al visitante si se ha realizado alguna modificación o agregado en el sitio Web. Se debe encontrar en la página principal.	Criterio binario, discreto y absoluto	0
1.2.2.2	Última Actualización Restringido (por subsitio o página)	Permite conocer al visitante si se ha realizado alguna modificación o agregado en alguna página del subsitio (la que está visitando).	Criterio Muftí-nivel discreto, absoluto	100
1.2.3.1	Directorio E-mail	Es el lugar en el sitio donde se agrupan las direcciones electrónicas para enviar E-mail (como mecanismo de retroalimentación)	Criterio binario, discreto y absoluto	100
1.2.3.2	Directorio TE-Fax	Es el lugar en el sitio donde se agrupan las direcciones de números para contactarse vía telefónica.	Criterio binario, discreto y absoluto	100
1.2.4	Facilidad FAQ	La disponibilidad FAQ es un conjunto de preguntas que se realizan con mayor frecuencia, y que están ya publicadas en el sitio con sus respuestas. Esto le permite aprender y/o ayudar a los visitantes (evitando la demora cognitiva)	Criterio muftí-nivel, discreto y absoluto	0
1.2.5.1	Cuestionarios/ Encuestas	Si la página web tiene este método de retroalimentación para con los usuarios.	Criterio binario, discreto y absoluto	100
1.2.5.2	Comentarios/ Sugerencias	Este atributo permite hacer conocer las inquietudes u otros intereses que el usuario formule mediante el llenado de un formulario con mayor o menor nivel de	Criterio binario, discreto y absoluto	100

		estructuración.		
1.3.1	Cohesividad al Agrupar los Objetos de Control Principales	Si los enlaces que posee el sitio web están agrupados del la pagina o menú principal.	Criterio Muftí-nivel, discreto, absoluto	80
1.3.2.1	Permanencia de Controles Directos	Este atributo representa la permanencia directa de los controles del menú principal del sitio que permiten navegación	Criterio binario, discreto y absoluto	100
1.3.2.2	Permanencia de Controles Indirectos	Es un control o referencia indirecta a la página principal (en donde se encuentran los controles a los subsitios)	Criterio binario, discreto y absoluto	100
1.3.2.3	Estabilidad	Se refiere a la ubicación de los controles principales directos o indirectos en los nadas del sitio. Un control es estable si se encuentra siempre en la misma ubicación al navegar por los nodos	Criterio Muftí-nivel, discreto, absoluto	100
1.3.3.1	Uniformidad en el Color de Enlaces	Los enlaces de cada nodo deben tener un mismo color, enlaces visitados, etc.	Criterio Multi-nivel, discreto, absoluto	70
1.3.3.2	Uniformidad en el Estilo Global	Todo el sitio debe tener una misma apariencia es decir, mismos colores, letra, tamaños, etc.	Criterio Multi-nivel, discreto, absoluto	80
1.4.1	Soporte a Lenguaje Extranjero	Este atributo modela la disponibilidad parcial o total de lenguajes extranjeros soportados por el sitio Web. No se computa el lenguaje nativo como lenguaje extranjero	Criterio muftí-variable, continuo y absoluto;	0
1.4.2	Atributo "Qué es lo Nuevo"	Este atributo indica que la página está constantemente actualizada, al publicar siempre todo lo novedoso.	Criterio binario, discreto y absoluto	100
1.4.3	Indicador de Resolución de Pantalla	Para tener la información de que calidad de resolución esta en el usuario.	Criterio binario, discreto y absoluto	0

Tabla 4.11 Resultados de las preferencias de calidad elemental de la característica Usabilidad
Fuente: Elaboración propia

Codigo	Atributo	Descripción	Tipo de Criterio Elemental	IE%
2.1.1.1.1	Búsqueda de Autores	Búsqueda restringida a un subsitio o parte de un sitio, debido a que el mismo es altamente cohesivo o distintivo del resto de la información del sitio Web en este caso los autores.	Criterio multi-nivel, discreto y absoluto	100
2.1.1.2	Búsqueda Global	Este atributo permite realizar búsquedas sobre cualquier palabra en el sitio.	Criterio multi-nivel, discreto y absoluto	60
2.2.1.1.1	Indicador del Camino	Los usuarios al navegar por el sitio, deben tener pistas visuales, que les indique con precisión en dónde se encuentran posicionados dentro de la estructura del espacio de información del sitio. Este atributo trata con la orientación del usuario.	Criterio multi-nivel, discreto y absoluto	70
2.2.1.1.2	Etiqueta de la Posición Actual	Este atributo permite hacer conocer a los usuarios en qué nodo están posicionados, por medio de una etiqueta o título.	Criterio multi-nivel, discreto y absoluto	60
2.2.1 .2	Promedio de Enlaces por Página	Cada página o nodo tiene generalmente dos tipos principales de enlaces (o puntos de partida): enlaces semánticos (distribuidos en la información de las páginas del sitio o contenido), y enlaces estructurales (a cerca de controles principales y navegacionales)	Criterio multi-nivel, discreto y absoluto	80
2.2.2.1 .1	Permanencia de los Controles Contextuales	Es la permanencia de los controles de navegación de un subsitio	Criterio multi-nivel, discreto y absoluto	100
2.2.2.1 .2	Estabilidad	Se refiere a la ubicación de los controles contextuales en los nodos del subsitio. Un control es estable si se encuentra siempre en la misma ubicación durante la navegación de los nodos del subsitios.	Criterio multi-nivel, discreto y absoluto	100
2.2.2.2.1	Desplazamiento Vertical	Nivel de desplazamiento vertical que el visitante debe realizar para ajustar la interface.	Criterio multi-nivel, discreto y absoluto	70
2.2.2.2.2	Desplazamiento Horizontal	Nivel de desplazamiento horizontal que el visitante debe realizar para ajustar la interface (considerando 640x480 como la mínima resolución)	Criterio multi-nivel, discreto y absoluto	100
2.2.3.1	Enlace con Título (enlace con texto explicatorio)	Este atributo trata de predecir los temas o contenidos que están asociados al enlace.	Criterio multi-nivel, discreto y absoluto	70
2.2.3.2	Calidad de la Frase del Enlace	Valoración de Preferencia Directa de la calidad de la frase del enlace. Cuando los enlaces no son suficientemente descriptivos, el visitante podría no tener	Criterio de Preferencia Directa	50

		una buena pista respecto de lo que dichos enlaces significan, principalmente cuando no hay un buen contexto		
2.3.1 .1	Información Institucional	Información acerca de la institución.	Criterio binario, discreto y absoluto	100
2.3.1 .2.1	Información de los Requerimientos para publicación de documentos	Información de los requisitos y condiciones que necesita el autor para publicar su documento.	Criterio binario, discreto y absoluto	100
2.3.1.2.2	Formulario para Registro	Formulario de registro	Criterio binario, discreto y absoluto	100
2.3.1 .3.1	Lista de autores	Una lista de todos los autores registrados que publicaron documentos.	Criterio binario, discreto y absoluto	0
2.3.1.3.2	Buscador de Documentos	Es un buscador e identificador de documentos, un buscador con parámetros avanzados como búsqueda por título, autor, palabra clave, etc.	Criterio binario, discreto y absoluto	100
2.3.1.3.3	Información de los documentos publicados	Información de los documentos publicados en el servidor OAI que está a disponible a todos los cosechadores OAI.	Criterio binario, discreto y absoluto	100
2.3.1.3.4	Información de la solicitud de registros de publicación	Información de los registros de solicitudes de publicaciones, accediendo a información que el propio autor autoriza su publicación del documento	Criterio binario, discreto y absoluto	100
2.3.1 .4.1	Índice de Servicios	Muestra todos los servicios para acceder a alguno de ellos.	Criterio muftí-nivel, discreto absoluto	50
2.3.1 .4.2	Información de Nuevas Publicaciones	Información de nuevas publicaciones que la biblioteca debe brindar a sus usuarios.	Criterio binario, discreto y absoluto	100
2.3.1 .4.3	Información de los autores	Información de los autores que solicitan la publicación de sus documentos.	Criterio binario, discreto y absoluto	100
2.3.1 .4.4	Información de Eventos Académicos de los servidores OAI	Información de eventos académicos de los cosechadores de servidores OAI.	Criterio binario, discreto y absoluto	100
2.3.2.1	Catalogación del documento	Información de catalogación del documento en el servidor OAI para todo usuario que esté interesado.	Criterio binario, discreto y absoluto	100
2.3.2.2	Información de la solicitud de publicaciones	Información de todos los autores que solicitan la publicación en el servidor OAI de la biblioteca.	Criterio binario, discreto y absoluto	100
2.3.2.3	Información de los	Información que pone al tanto a los	Criterio binario,	100

	derechos de propiedad intelectual del documento.	cosechadores de servidores OAI sobre la propiedad intelectual.	discreto y absoluto	

Tabla 4.12 Resultados de las preferencias de calidad elemental de característica Funcionalidad
Fuente: Elaboración propia

Código	Atributo	Descripción	Tipo de Criterio Elemental	IE%
3.1.1.1	Enlaces Rotos	Este atributo representa básicamente a los enlaces encontrados que conducen a nodos destino ausentes (también llamados enlaces ausentes o pendientes)	Criterio de variable normalizada, continuo y absoluto	100
3.1.1.2	Enlaces Inválidos	Representa básicamente a los enlaces que llevan a otro destino del esperado.	Criterio de variable normalizada, continuo y absoluto	100
3.1.1.3	Enlaces no Implementados	Son los enlaces que debido a problemas en el desarrollo no fueron implementado o puestos a disposición.	Criterio de variable normalizada, continuo y absoluto	90
3.1.2.1	Deficiencias o cualidades ausentes debido a diferentes navegadores (browsers)	Representa algunos problemas relacionados a estilos formas, tamaño, lugares, etc.	Criterio multi- variable, continuo y absoluto	70
3.1.2.2	Deficiencias o resultados inesperados independientes de browsers.	Problemas o deficiencias que se presentan inesperadamente.	Criterio multi- variable, continuo y absoluto	100
3.1.2.3	Nodos Destinos en Construcción	Son Nodos que aún no ha sido desarrollado completamente y se muestra que están en construcción.	Criterio de variable normalizada, continuo y absoluto	100
3.1.2.4	Nodos Web Muertos (sin enlaces de retorno)	Son nodos que no pueden volver al nodo anterior del cual estaban viniendo.	Criterio de variable normalizada, continuo y absoluto	100

Tabla 4.13 Resultados de las preferencias de calidad elemental de lancaracterística Confiabilidad
Fuente: Elaboración propia

Código	Atributo	Descripción	Tipo de Criterio Elemental	IE%
4.1.1	Páginas de Acceso Rápido	El tamaño de cada página se especifica como una función del tiempo de espera y de la velocidad mínima establecida para una línea de comunicación dada	Criterio multi-variable, continuo y absoluto	90
4.2.1.1	Soporte a Versión sólo Texto	Este atributo representa la accesibilidad a la información que está en las páginas, principalmente para las personas con invalidez o cuando la velocidad es un problema. Es de relevancia que el sitio entero sea editado en una versión sólo texto; sin embargo, a veces, una disponibilidad parcial también puede ser deseable	Criterio multi-nivel, discreto y absoluto	60
4.2.1.2.1	Imagen con Título	Se debe proveer texto alternativo para cada imagen u objeto gráfico para que comuniquen información visual. Este atributo mide el nivel de disponibilidad del texto alternativo a la imagen, útil principalmente cuando se desactiva la propiedad de ver imágenes del navegador	Criterio de variable normalizada, continuo y absoluto	95
4.2.1.2.2	Legibilidad Global	Este atributo representa la preferencia de calidad en consideración del nivel de legibilidad global del sitio, cuando se desactiva la propiedad de ver imágenes del navegador.	Criterio de preferencia directa	60
4.7.2.1	Numero de Vistas considerando Marcos.	Los frames or marcos organizan a una ventana en diferentes áreas o sub vistas tanto de control como de contenido. Cuanto mayor es la cantidad de frames, menor es la accesibilidad de la ventana, principalmente para personas con discapacidades.	Criterio multi-nivel, discreto y absoluto	100
4.2.2.2	Versión sin Marcos	Cuando un sitio utiliza la estrategia de frames marcos es deseable que cuente con una opción de versión del sitio sin Trames, en consideración de la accesibilidad	Criterio binario, discreto y absoluto	100

Tabla 4.14 Resultados de las preferencias de calidad elemental de la característica Eficiencia
Fuente: Elaboración propia

4.2.4 DEFINICIÓN E IMPLEMENTACIÓN DE LA EVALUACIÓN GLOBAL

En la fase de Definición e Implementación de la Evaluación Global se trata con actividades, modelos, procedimientos y herramientas para determinar los criterios de agregación de las preferencias de calidad elemental (obtenidas en la fase anterior, a partir del árbol de requerimientos), para producir la preferencia global para cada sistema de información interviniente[22].

4.2.4.1 AGREGACIÓN DE LAS PREFERENCIAS ELEMENTALES

Aplicando un mecanismo de agregación paso a paso, las preferencias de calidad elementales deben estructurarse y agregarse de un modo de abajo hacia arriba (estrategia bottom-up) para permitir el ulterior cálculo de las preferencias parciales respectivas. A su vez, repitiendo el proceso de agregación recursivo al final puede obtenerse la estructura de agregación de todo el sistema.

El evaluador puede valorar la relativa importancia de los elementos de entrada a las funciones de acuerdo a su intuición y experticia, o puede utilizar mecanismos como encuestas y establecer fórmulas de relativa importancia para computar los pesos.

En este caso se usó como base los datos proporcionados por Olsina que para muchos de los atributos se utilizó una estrategia intuitiva, (dado el esfuerzo que implicaba realizar la encuesta para todos los elementos intervinientes). Del mismo modo que para los pesos, se utilizó una estrategia intuitiva para la determinación de los tipos de funciones y los niveles de polarización lógica.

[22] [Olsina 99]

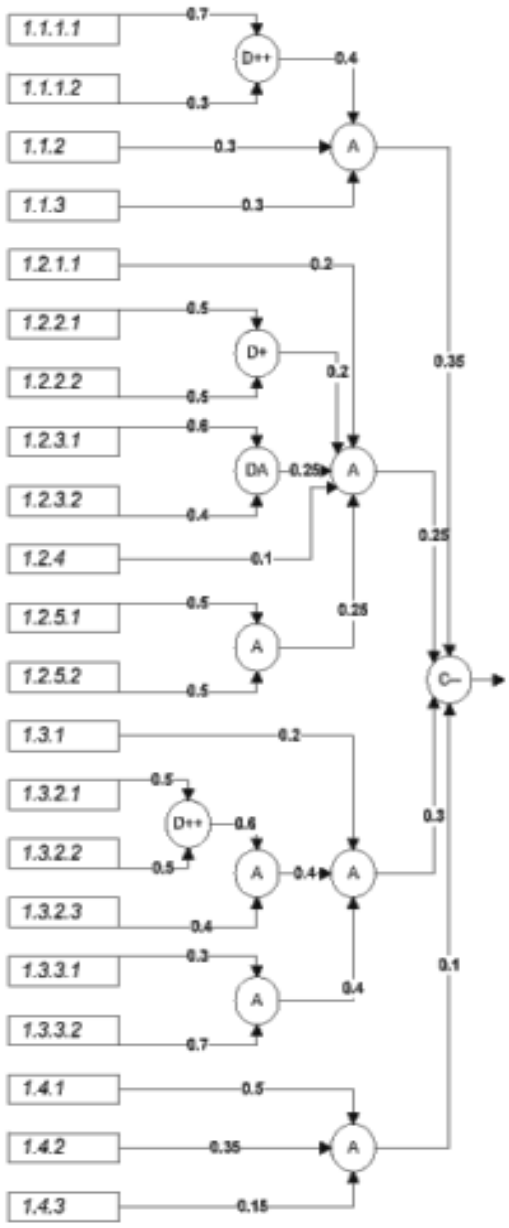


Figura 4.1 Estructura de agregación de preferencias parciales para la característica Usabilidad
Fuente Elaboración propia

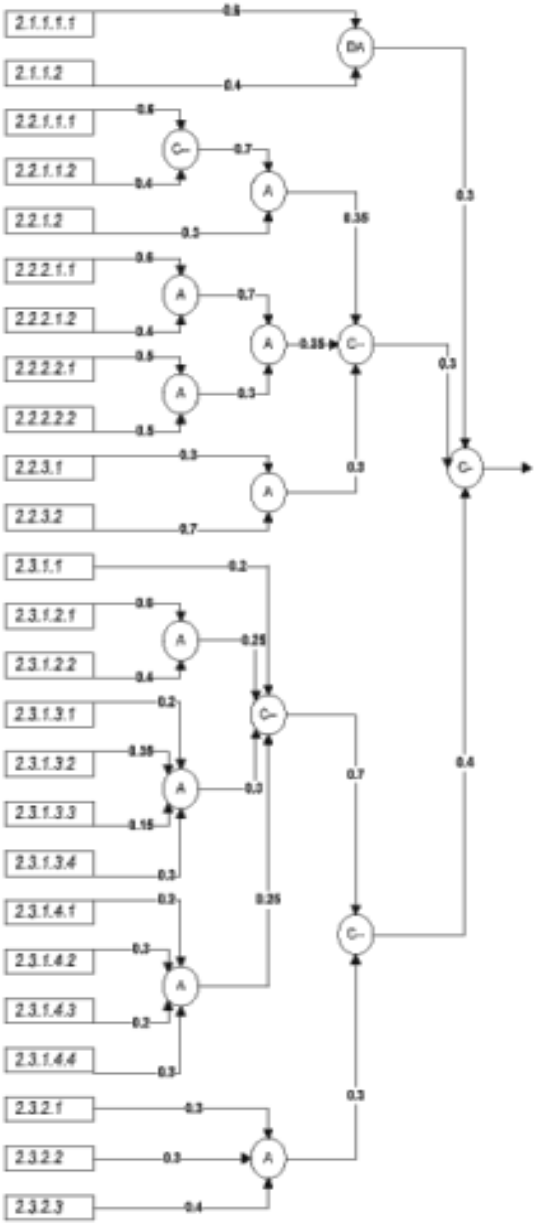


Figura 4.2 Estructura de agregación de preferencias parciales para la característica Funcionalidad
Fuente Elaboración propia

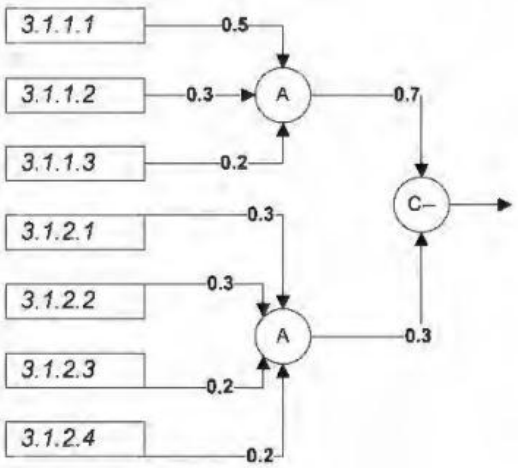


Figura 4.3 Estructura de agregación de preferencias parciales para la característica Confiabilidad
Fuente Elaboración propia

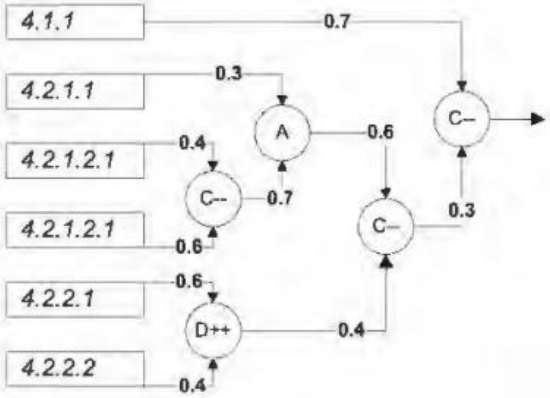


Figura 4.4 Estructura de agregación de preferencias parciales para la característica Eficiencia
Fuente Elaboración propia

En la figura 4.5 se puede apreciar la estructura de agregación de preferencias parciales para las características de más alto nivel. Esta estructura producida es útil para cuantificar el indicador de calidad global IG para el sistema Web.

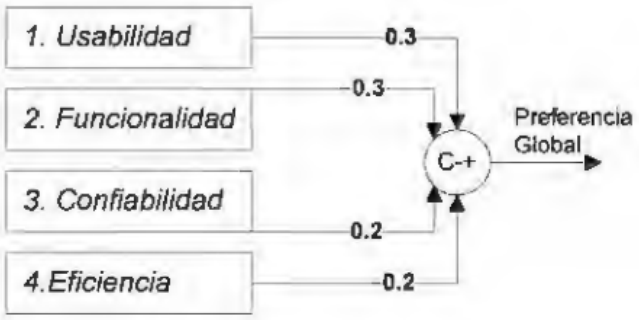


Figura 4.5 Estructura de agregación de características de alto nivel
Fuente Elaboración propia

4.2.4.2 CÓMPUTO DE LAS PREFERENCIAS DE CALIDAD PARCIALES Y GLOBALES

Finalmente, una vez establecida la estructura, se debe llevar a cabo el proceso de cómputo de las preferencias de calidad. Para esta actividad se hace uso de herramientas de cálculo automáticas. La preferencia de calidad global representa el grado de satisfacción de todos los requisitos involucrados.

Los valores obtenidos IE_1 de la evaluación elemental, permitirán obtener la evaluación global. Aplicando el modelo LSP juntamente con la función media potencia pesada, que se muestra a continuación;

$$IG_{(r)} = (P_1 IE_1^r + P_2 IE_2^r + ... + P_m IE_m^r)^{1/r}$$

Donde:

$-\infty <= r <= +\infty;$

$0 <= IE_i <= 1;$

$(P_1 + P_2 + ... + P_m) = 1;$

$i = 1.....m;$

r : es el exponente de la ecuación, guarda el valor real conforme al operador lógico y a la cantidad de entradas seleccionadas para una función de agregación dada. La siguiente tabla muestra los valores parametrizados de r para una función CDG de 17 niveles, para 2, 3, 4 y 5 entradas respectivamente.

Operador L.SP	Abreu	c	d	r(2)	r(3)	r(4)	r(5)	Mandat
Disyunción	D	0.0000	1.0000	$-\infty$	$-\infty$	$-\infty$	$-\infty$	No
CD Fuerte(+)	D++	0.0625	0.9375	20.63	24.30	27.11	30.09	No
CD Fuerte	D+	0.1250	0.8750	9.521	11.095	12.27	13.235	No
CD Fuerte (-)	D+-	0.1875	0.8125	5.802	6.675	7.316	7.819	No
CD Media	DA	0.2500	0.7500	3.929	4.450	4.825	5.111	No
CD Debil (+)	D-+	0.3125	0.6875	2.792	3.101	3.318	3.479	No

CD Debil	D-	0.4375	0.6250	2.018	2.187	2.302	2.384	No
CD Debil (-)	D--	0.5000	0.5625	1.449	1.519	1.569	1.596	No
Media Aritmet	A	0.5625	0.5000	1.000	1.000	1.000	1.000	No
CC Debil (+)	C--	0.6250	0.4375	0.4375	0.619	0.573	0.546	No
CC Debil	C-	0.6250	0.3750	0.261	0.192	0.153	0.129	No
CC Debil (-)	C-+	0.6875	0.3125	-0.148	-0.208	-0.235	0.251	Si
CC Media	CA	0.7500	0.2500	-0.720	-0.732	-0.721	-0.707	Si
CC Fuerte (-)	C+-	0.8125	0.1875	-1.655	-1.550	-1.455	-1.380	Si
CC Fuerte	C+	0.8750	0.1250	-3.510	-3.114	-2.823	-2.606	Si
CC Fuerte (+)	C++	0.9375	0.0625	-9.060	-7.639	-6.689	-6.013	Si
Conjunción	C	1.0000	0.0000	$-\infty$	$-\infty$	$-\infty$	$-\infty$	Si

Tabla 4.15 Valores parametrizados de "r"
Fuente [Dujmovic 74]

Evaluación Global para la usabilidad

A continuación se calcula los IG(r) parciales para la característica Usabilidad

Para la sub-característica 1.1.1

$$IG(2) = (0.7*100^r + 0.3*0^r)^{1/r} = 98.28 \qquad ;r=20.63$$

Para la sub-característica 1.1

$$IG(3) = (0.4*98.28^r + 0.3*80^r + 0.3*0^r)^{1/r} = 63.31 \qquad ;r=1.0$$

Para la sub-característica 1.2.2

$$IG(2) = (0.5*0^r + 0.5*100^r)^{1/r} = 92.28 \qquad ;r=9.521$$

Para la sub-característica 1.2.3

$$IG(2) = (0.6*100^r + 0.4*100^r)^{1/r} = 100$$;r=3.929

Para la sub-característica 1.2.5

$$IG(2) = (0.5*100^r + 0.5*100^r)^{1/r} = 100$$;r=1.0

Para la sub-característica 1.2

$$IG(5) = (0.5*60^r + 0.2*92.98^r + 0.25*100^r + 0.25*100^r + 0.1*0^r + 0.25*100^r)^{1/r}$$
; r=1.0

IG(5)=80.60

Para la sub-característica 1.3.2

$$IG(2) = (0.5*100^r + 0.5*100^r)^{1/r} = 100$$;r=20.63

$$IG(2) = (0.6*100^r + 0.4*100^r)^{1/r} = 100$$;r=1.0

Para la sub-característica 1.3.3

$$IG(2) = (0.3*70^r + 0.7*80^r)^{1/r} = 77$$;r=1.0

Para la sub-característica 1.3

$$IG(3) = (0.2*80^r + 0.4*100^r + 0.4*77^r)^{1/r} = 86.8$$;r=1.0

Para la sub-característica 1.4

$$IG(3) = (0.5*0^r + 0.35*100^r + 0.15*0^r)^{1/r} = 35$$;r=1.0

Reemplazando los valores de las sub-características, calculamos el IG(r) para la característica Usabilidad

$$IG(4) = (0.35*63.31^r + 0.25*80.60^r + 0.3*86.8^r + 0.1*35^r)^{1/r}$$; r=10.546

IG(4)=70.93

Evaluación Global para la funcionalidad

A continuación se calcula los IG(r) parciales para la característica Funcionalidad

Para la sub-característica 2.1

$IG(2) = (0.6*100^r + 0.4*60^r)^{1/r} = 89.75$; r=10.546

Para la sub-característica 2.2.1.1

$IG(2) = (0.6*70^r + 0.4*60^r)^{1/r} = 865.92$; r=0.619

Para la sub-característica 2.2.1

$IG(2) = (0.7*65.92^r + 0.3*80^r)^{1/r} = 70.14$; r=1.0

Para la sub-característica 2.2.2.1

$IG(2) = (0.7*65.92^r + 0.3*80^r)^{1/r} = 70.14$; r=1.0

Para la sub-característica 2.2.2.2

$IG(2) = (0.5*70^r + 0.5*100^r)^{1/r} = 85$; r=1.0

Para la sub-característica 2.2.2

$IG(2) = (0.7*100^r + 0.3*85^r)^{1/r} = 95.5$; r=1.0

Para la sub-característica 2.2.3

$IG(2) = (0.3*70^r + 0.7*50^r)^{1/r} = 56$; r=1.0

Para la sub-característica 2.2

$IG(3) = (0.35*70.14^r + 0.35*95.5^r + 0.3*56^r)^{1/r} = 74.03$; r=0.573

Para la sub-característica 2.3.1.2

$IG(2) = (0.3*70^r + 0.7*50^r)^{1/r} = 56$; r=1.0

Para la sub-característica 2.3.1.3

$IG(4) = (0.2*0^r + 0.35*100^r + 0.15*100^r + 0.3*100^r)^{1/r} = 80$; r=1.0

Para la sub-característica 2.3.1.4

$$IG(4) = (0.2*50^r + 0.3*100^r + 0.2*100^r + 0.3*100^r)^{1/r} = 90$$; r=1.0

Para la sub-característica 2.3.1

$$IG(4) = (0.2*100^r + 0.25*100^r + 0.3*80^r + 0.25*90^r)^{1/r} = 91.32$$; r=0.546

Para la sub-característica 2.3.2

$$IG(3) = (0.3*100^r + 0.3*100^r + 0.4*100^r)^{1/r} = 100$$; r=1.0

Para la sub-característica 2.3

$$IG(2) = (0.2*91.32^r + 0.7*100^r)^{1/r} = 93.89$$; r=0.619

Reemplazando los valores de las sub-características, calculamos el IG(r) para la característica Funcionalidad

$$IG(3) = (0.3*89.75^r + 0.3*74.03^r + 0.4*93.89^r)^{1/r}$$; r=0.192

IG(3)=86.34

Evaluación Global para la confiabilidad

A continuación se calcula los IG(r) parciales para la característica Confiabilidad

Para la sub-característica 3.1.1

$$IG(3) = (0.5*100^r + 0.3*100^r + 0.2*90^r)^{1/r} = 98$$; r=1.0

Para la sub-característica 3.1.2

$$IG(3) = (0.3*70^r + 0.3*100^r + 0.2*100^r + 0.2*100^r)^{1/r} = 91$$; r=1.0

Reemplazando los valores de las sub-características, calculamos el IG(r) para la característica Confiabilidad

$$IG(2) = (0.7*98^r + 0.3*91^r)^{1/r}$$; r=0.619

IG (2)=95.88

Evaluación Global para la eficiencia

A continuación se calcula los IG(r) parciales para la característica Eficiencia.

Para la sub-característica 4.2.1.2

$IG(2) = (0.4*90^r + 0.6*60^r)^{1/r} = 71.44$; r=0.619

Para la sub-característica 4.2.1

$IG(2) = (0.3*60^r + 0.7*71.44^r)^{1/r} = 68.01$; r=1.0

Para la sub-característica 4.2.2

$IG(2) = (0.6*100^r + 0.4*100^r)^{1/r} = 100$; r=20.63

Para la sub-característica 4.2

$IG(2) = (0.6*68.01^r + 0.4*100^r)^{1/r} = 80.24$; r=0.619

Reemplazando los valores de las sub-características, calculamos el IG(r) para la característica Eficiencia

$IG(2) = (0.7*90^r + 0.3*80.24^r)^{1/r}$; r=0.619

IG (2)=87.02

Los valores parciales obtenidos de las características de más alto nivel no permitirán determinar el valor de la calidad global, haciendo uso de la misma función tenemos:

$IG(4) = (0.3*70.93^r + 0.3*86.34^r + 0.2*95.88^r + 0.2*87.02^r)^{1/r}$; r=-0.235

IG (4)=83.13

Calidad Global **83.13**

4.2.5 ANÁLISIS DE RESULTADOS

El proceso de evaluación en consideración a los usuario que acceden a los documentos a través de los cosechadores de metadatos académicos de los " Servidores de Metadatos con el Protocolo OAI-PMH de la Biblioteca central de la Universidad Mayor de San Andrés" presenta los siguientes resultados

Característica	IG(r)
1. Usabilidad	70.93
2. Funcionalidad	86.34
3. Confiabilidad	95.88
4. Eficiencia	87.02
Calidad Global	**83.13**

Tabla 4.16 Resultados de la evaluación global
Fuente: Elaboración propia

Teniendo en cuenta, un nivel de aceptabilidad satisfactorio, indica que los valores están en al rango de 60 a 100 %. Un nivel de aceptabilidad marginal indica que los valores de la preferencia caen en el rango de 40 a 60%, Finalmente, un valor de preferencia es insatisfactorio si cae dentro del rango de 0 % y el 40%[23].

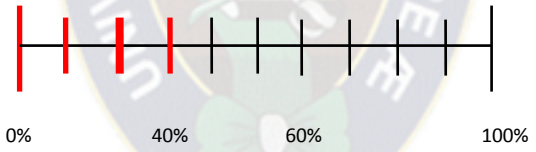


Figura 4.6 Rango de aceptabilidad de preferencia de calidad
Fuente: [Olsina 991]

Las valores obtenidos de las características de más alto nivel están dentro de los márgenes de satisfacción.

[23] [Olsina 99]

Para finalizar podernos concluir que la valoración que se obtuvo después de aplicar la metodología Web-site QEM en el "Sistema Servidores de Metadatos con el Protocolo OAI-PMH de la Biblioteca central de la Universidad Mayor de San Andrés" y orientado a los usuario que acceden a los documentos en general, nos da una calificación o un porcentaje de la calidad global de 83.13%, el cual está dentro del grado **satisfactorio.**

4.3 SEGURIDAD

El incidente más sonado actualmente es sin duda, la vulnerabilidad Heartbleed, que sacudió los cimientos de la seguridad en Internet. En este caso, el quid del problema no era la sagacidad de los ciberdelincuentes, sino otra verdad incómoda: que ningún programa es infalible y, por lo tanto, hay que cuidar al máximo la seguridad por tratarse de un sistema de gestión de documentos digitales académicos de acceso libre, comprometen a los autores por los registros de propiedad intelectual, la seguridad de los sitios web con sistemas más avanzados y eficaces.

Para registrar una variable de sesión en java se usa uno de los conceptos a trabajar con aplicaciones web en Java es el concepto de java session (HttpSession) que sirve para almacenar información entre diferentes peticiones HTTP ya que este protocolo es stateless (sin estado). Así pues en este sistema encontraremos con el estado (datos usuario) entre un conjunto amplio de páginas de nuestra Aplicación.

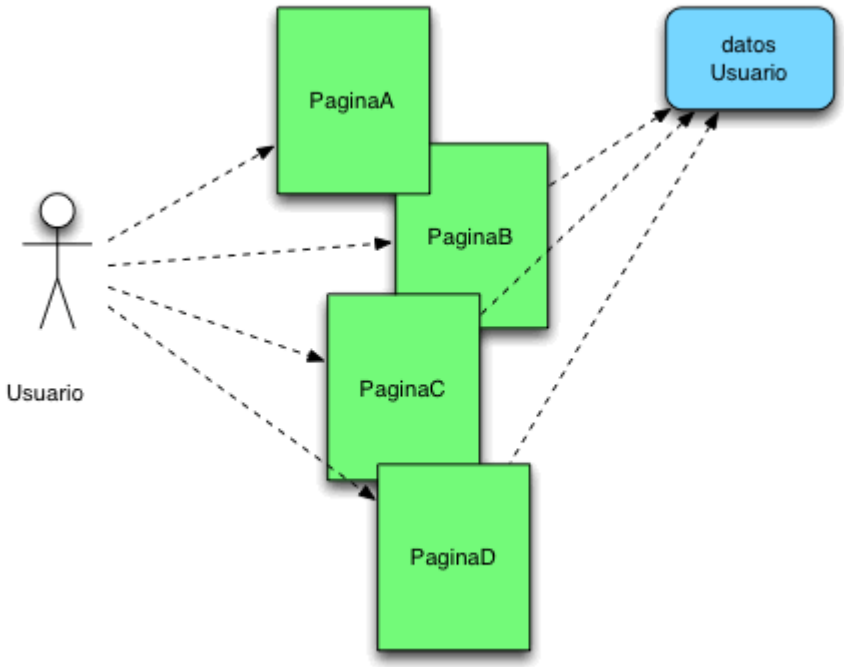


Figura 4.7 Compartir nuestros datos de usuario de
varias páginas de acceso
Fuente: Elaboración propia

Para registrar una variable de sesión y establecer la disposición de los datos del usuario para varias páginas en la plataforma Java EE se usa de forma muy habitual la clase HttpSession que tiene una estructura de HashMap (Diccionario) y permite almacenar cualquier tipo de objeto en ella de tal forma que pueda ser compartido por las diferentes páginas que como usuarios utilizan para acceder a los documentos digitales desde distintos sitios de internet.

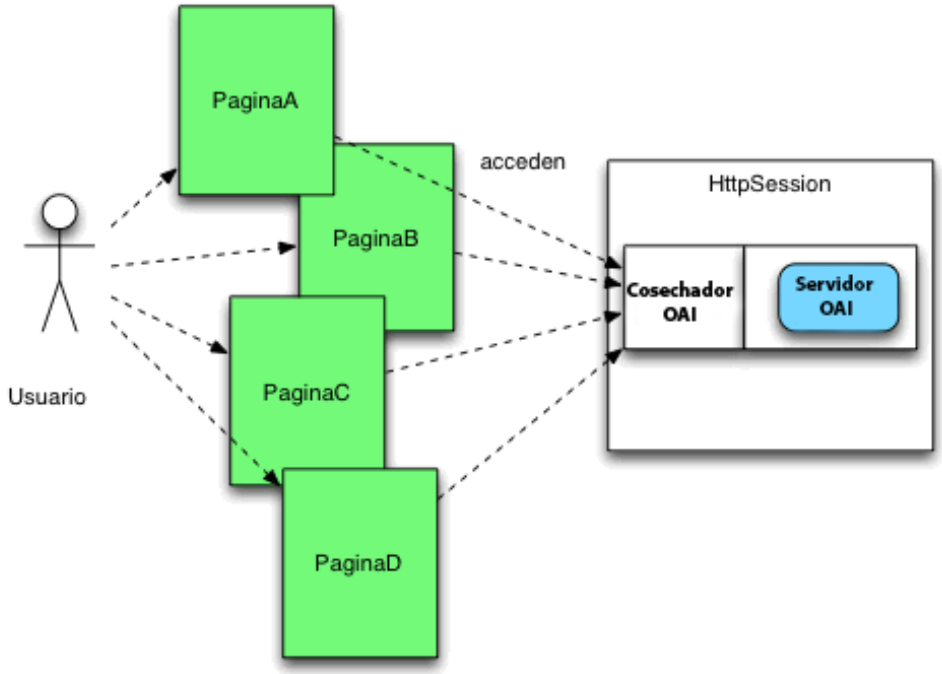


Figura 4.8 Usuario accede por diferentes paginas
al Cosechador OAI
Fuente: Elaboración propia

El funcionamiento del sistema de sessiones es relativamente sencillo. Cada vez que un usuario crea una session accediendo a una página (que la genere) se crea un objeto a nivel de Servidor con un HashMap vacío que nos permite almacenar la información que necesitamos relativa a este usuario. Realizado este primer paso se envía al navegador del usuario una Cookie que sirve para identificarle y asociarle el HashMap que se acaba de construir para que pueda almacenar información en él. Este HashMap puede ser accedido desde cualquier otra página permitiéndonos compartir información.

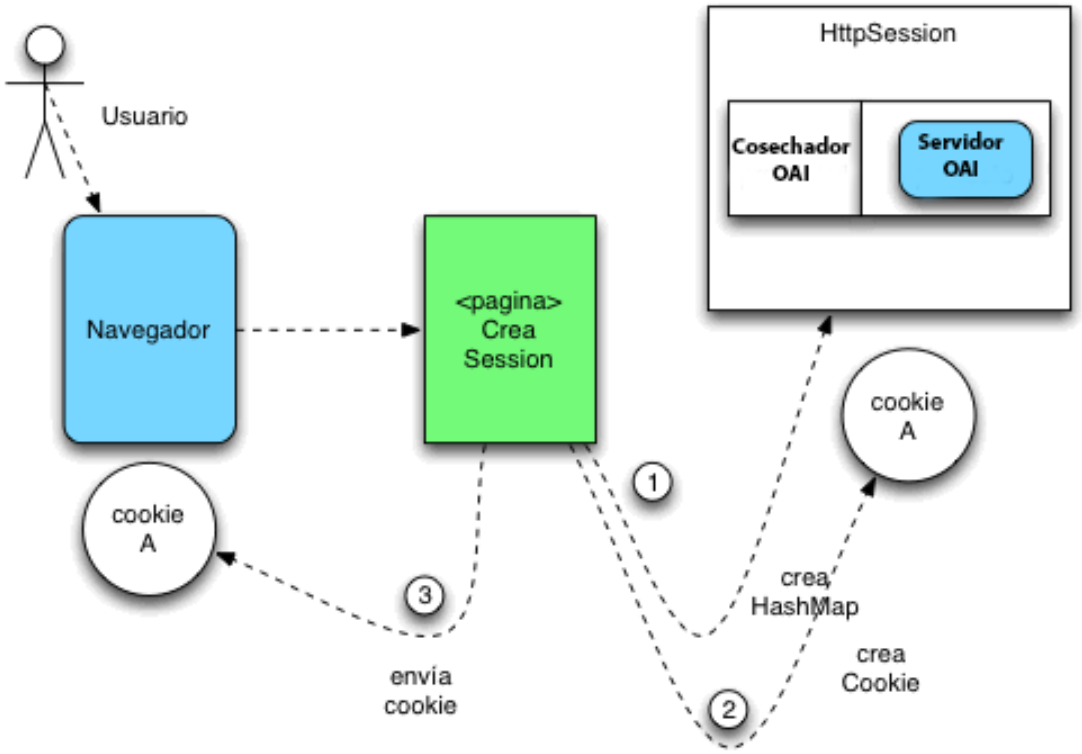


Figura 4.9 Usuario recibe un envía al navegador del usuario una Cookie que sirve para identificarle y asociarle el HashMap
Fuente: Elaboración propia

El ejemplo anterior nos sirve también como método para realizar el inicio de sesión (validación) de un usuario en nuestro sistema web, mostramos la consulta SQL que se ejecutará y la comprobación de si el usuario existe y si la contraseña introducida en el formulario (txtcontraseña) coincide con la guardada en la base de datos (contraseña una vez obtenido su HashMap).

Cada usuario y sesión es individual de cada usuario que se conecta a nuestro servidor de metadatos académicos OAI y la información no es compartida entre ellos. Así pues cada

usuario dispondrá de su propio HashMap en donde almacenar la información que resulte útil entre páginas y resultado de las búsquedas.

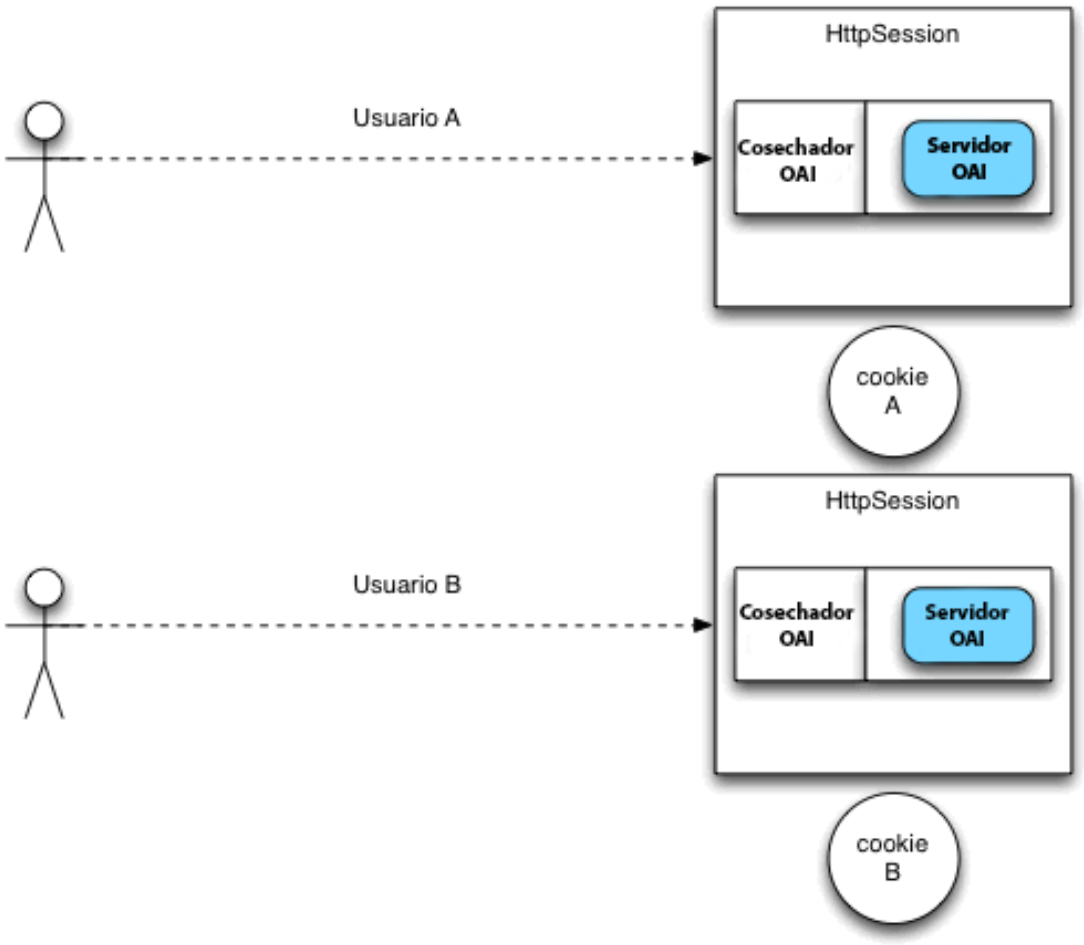


Figura 4.10 Usuario recibe un envía al navegador del usuario una Cookie que sirve para identificarle y asociarle el HashMap
Fuente: Elaboración propia

Cuando el usuario ha iniciado sesión, siempre es recomendable añadir un enlace para que éste pueda cerrar la sesión en cualquier momento de forma segura. Así evitaremos accesos indebidos por otros usuarios que usen el mismo equipo que el que inició sesión.

En teoría, siempre que se cierre el navegador web se cerrará la sesión automáticamente, pero es aconsejable añadir este enlace para que el usuario decida cuándo cerrar la sesión.

Vamos a ver el código do los Servlets en el cual un servlet almacena datos en la sesión (concretamente un objeto producto) y otro servlet lee los datos de la session y los muestra por pantalla al usuario.

```
package com.arquitecturajava;

public class LogSistema {

private int id;
private String concepto;
private double importe;
public int getId() {
return id;
}
public void setId(int id) {
this.id = id;
}
public String getConcepto() {
return concepto;
}
public void setConcepto(String concepto) {
this.concepto = concepto;
}
public double getImporte() {
return importe;
}
public void setImporte(double importe) {
this.importe = importe;
}
public LogSistema(int id, String concepto, double importe) {
super();
this.id = id;
this.concepto = concepto;
this.importe = importe;
```

```
}
public LogSistema() {
super();
}

}

package com.arquitecturajava;

import java.io.IOException;
import java.io.PrintWriter;

import javax.servlet.ServletException;
import javax.servlet.annotation.WebServlet;
import javax.servlet.http.HttpServlet;
import javax.servlet.http.HttpServletRequest;
import javax.servlet.http.HttpServletResponse;
import javax.servlet.http.HttpSession;

@WebServlet("/CrearSession")
public class CrearSession extends HttpServlet {
private static final long serialVersionUID = 1L;

protected void doGet(HttpServletRequest request, HttpServletResponse
response) throws ServletException, IOException {

HttpSession misession= request.getSession(true);
LogSistema miproducto= new LogSistema(1,"Usuario",300);
misession.setAttribute("producto",miproducto);
PrintWriter pw= response.getWriter();
pw.println("<html><body>Inicio de Sesion del Sistema OAI</body></html>");
pw.close();

}
}

package com.arquitecturajava;

import java.io.IOException;
import java.io.PrintWriter;
```

```
import javax.servlet.ServletException;
import javax.servlet.annotation.WebServlet;
import javax.servlet.http.HttpServlet;
import javax.servlet.http.HttpServletRequest;
import javax.servlet.http.HttpServletResponse;
import javax.servlet.http.HttpSession;

@WebServlet("/VerSession")
public class VerSession extends HttpServlet {
private static final long serialVersionUID = 1L;

protected void doGet(HttpServletRequest request, HttpServletResponse
response) throws ServletException, IOException {
HttpSession misession= (HttpSession) request.getSession();

LogSistema miproducto= (LogSistema) misession.getAttribute("producto");

PrintWriter pw= response.getWriter();
pw.println("<html><body>"+ miproducto.getId()+ ","
+miproducto.getConcepto()+","+ miproducto.getImporte());
pw.close();
}
}
```

Protección en la Base de Datos

La seguridad de la base de datos esta implementada en varios niveles:

Protección de los ficheros de la base de datos. Todos los ficheros almacenados en la base de datos están protegidos contra escritura por cualquier cuenta que no sea la del superusuario de Postgres.

Las conexiones de los clientes al servidor de la base de datos están permitidas, por defecto, únicamente mediante sockets Unix locales y no mediante sockets TCP/IP. Ha de arrancarse el demonio con la opción -i para permitir la conexión de clientes no locales.

Las conexiones de los clientes se pueden restringir por dirección IP y/o por nombre de usuario mediante el fichero pg_hba.conf situado en PG_DATA.

Las conexiones de los clientes pueden ser autentificadas mediante otros paquetes externos.

A cada usuario de PostgreSQL se le asigna un nombre de usuario y (opcionalmente) una contraseña de 8 caracteres. Por defecto, los usuarios no tienen permiso de escritura a bases de datos que no hayan creado.

Los usuarios pueden ser incluidos en grupos, y el acceso a las tablas puede restringirse en base a esos grupos.

Autentificación de Usuarios

Autentificación es el proceso mediante el cual el servidor de la base de datos y el postmaster se aseguran de que el usuario que está solicitando acceso a la base de datos es en realidad quien dice ser. Todos los usuarios que quieren utilizar PostgreSQL se comprueban en la tabla pg_user para asegurarse que están autorizados a hacerlo. Actualmente, la verificación de la identidad del usuario se realiza de distintas formas:

Desde la shell del usuario

Un demonio que se lanza desde la shell del usuario anota el id original del usuario antes de realizar un setuid al id del usuario postgres. El id original del usuario se emplea como base para todo tipo de comprobaciones.

Desde los equipos de digitalización de la red

Si PostgreSQL se instala como distribuido en los equipos que digitalizan los documentos, el acceso al puerto TCP del postmaster está disponible para todo el mundo. El ABD configura el fichero pg_hba.conf situado en el directorio PG_DATA especificando el sistema de autentificación a utilizar en base al equipo que realiza la conexión y la base de datos a la que se conecta. Ver pg_hba.conf para obtener una descripción de los sistemas de autentificación disponibles.

Protección de Sistema Operativo del Servidor Debian 8

Asegurar nuestro Sistema Operativo Como Servidor es importante para proteger nuestros datos y propiedad intelectual de la mano de los Hackers.

Todos los datos que se transmiten por una red pueden ser interceptada, por eso es de vital importancia encriptar los datos donde sea posible usando claves, llaves o certificados.

En muchas configuraciones de red la mejor práctica es usar OpenSSH, SFTP o FTPS (FTP sobre SSL), los cuales agregan niveles de encriptación haciendo mucho mas difícil la captura de datos que están viajando.

Nuestro sistema operativo corre los servicios de red en servidores separados. Esto limita el número de los servicios que pueden ser comprometidos.

Si un hacker logra accesar al sistema por una vulnerabilidad de algún software o servicio, como apache, este también tendría acceso a otros servicios, digamos Postgres, Servidor OAI, Etc, si estos corren en el mismo servidor.

Deshabilitamos servicios innecesarios que corren en Segundo plano y que no estan siendo utilizados. También deshabilitamos en el booteo. Si queremos un listado de los servicios que están activos en un Runlevel específico corremos el siguiente comando como root.

```
#chkconfig –list | grep ’3:on’
```

Cerramos cualquier puerto de comunicación que no sea necesario

Corremos los siguientes commandos para listar los puertos abiertos en nuestro sistema y los programas asociados.

#netstat -tulpn o #nmap -sT -O localhost

No dejamos abierto los puertos que estemos utilizando y los que no lo cerramos y filtramos por Iptables. Mientras menos puertos tengamos abiertos menos posibilidad tenemos que puedan accesar a nuestro servidor.

Configurar a profundidad la Herramienta NMAP y filtrar las conexiones a estos puertos.

Copias de Seguridad

En un repositorio digital lo más importante es realizar las copias de seguridad de los documentos digitales, nuestro servidor de metadatos de documentos digitales debe ser resguardado periódicamente como una política de seguridad primaria.

Basado en la realización de duplicados de los datos almacenados. Esta técnica no puede trasladarse a un largo periodo de tiempo, ya que, suele utilizarse en pérdidas de datos debidas a fallos en la computadora, ya sea por causas técnicas o ajenas (como virus), o fallos en el hardware. Suele utilizarse en centros de procesamiento de datos o en situaciones domésticas para evitar la pérdida de estos datos importantes. Lo que se olvida es que los métodos de almacenamiento también quedan obsoletos, por lo que debe realizarse un rejuvenecimiento y una transferencia a nuevos medios.

4.4 ANALISIS COSTO - BENEFICIO

Entre los distintos métodos de estimación de costes de desarrollo de software, el modelo COCOMO (COnstructive COst MOdel) desarrollado por Barry M. Boehm se engloba en el grupo de los modelos algorítmicos que tratan de establecer una relación matemática la cual permite estimar el esfuerzo y tiempo requerido para desarrollar un producto.

Dentro de la jerarquía de modelos COCOMO, el modelo intermedio es el que se utilizara, dado que realiza las estimaciones con más precisión, es un modelo univariable estático que calcula el esfuerzo del desarrollo de software en función del tamaño del programa y un conjunto de "conductores de coste".

Este modelo presenta las siguientes ecuaciones:

$$E = a * (KLDC)^b * FAE$$

$$T = c * E^d$$

Donde:

E: Es el esfuerzo estimado expresado en hombres-mes.
KLDC: Número estimado de líneas de código distribuidas en miles para el proyecto.
FAE: Es el Factor de ajuste de esfuerzo, producto de 15 factores de costo.
T: Tiempo de duración del desarrollo.
a,b,c,d: Son constantes determinadas por el modo de desarrollo, obtenidas de la siguiente tabla.

MODO DE DESARROLLO	a	b	c	d
Organico	2.40	1.05	2.50	0.38
Semiacoplado	3.00	1.12	2.50	0.35
Empotrado	3.60	1.20	2.50	0.32

Tabla 4.17 Coeficiente Modelo COCOMO
Fuente: [Bedini, 2004]

El número estimado de líneas de código del proyecto aproximadamente son de 120000 de las cuales se estima que un 40% del código es reutilizable, entonces el total de líneas de código es:

KLDC = (total LDC — LDC reutilizable) / 1000

KLDC = (110000 — 44000) / 1000

KLDC = 66

Es así que el modo de desarrollo que optaremos será el semiacoplado ya que el número de líneas de código es mayor a 50 KLDC y menor a 300 KLDC, y que se utiliza en proyectos intermedios en complejidad y tamaño y donde la experiencia en este tipo de proyectos es variable, y las restricciones intermedias.

Por otro lado para hallar el FAE, se obtendrá mediante la multiplicación de los valores evaluados en los diferentes 15 conductores de coste que se muestra a continuación.

CONDUCTORES DE COSTE	VALORACIÓN					
	Muy bajo	Bajo	Nominal	Alto	Muy alto	Extra alto
Fiabilidad requerida del software	0,75	0,88	1.00	1,15	1,40	-
Tamaño de la base de datos	-	0,94	1.00	1,08	1,16	-
Complejidad del producto	0,70	0,85	1.00	1,15	1,30	1,65
Restricciones del tiempo de ejecución	-	-	1.00	1,11	1,30	1,66
Restricciones del almacenamiento principal	-	-	1.00	1,06	1,21	1,56
Volatilidad de la máquina virtual	-	0,87	1.00	1,15	1,30	-
Tiempo de respuesta del ordenador	-	0,87	1.00	1,07	1,15	-
Capacidad del analista	1,46	1,19	1.00	0,86	0,71	-
Experiencia en la aplicación	1,29	1,13	1.00	0,91	0,82	
Capacidad de los programadores	1,42	1,17	1.00	0,86	0,70	-
Experiencia con la máquina virtual	1,21	1,10	1.00	0,90	-	-
Experiencia en el lenguaje de programación	1,14	1,07	1.00	0,95	-	-
Prácticas de programación modernas	1,24	1,10	1.00	0,91	0,82	-
Utilización de herramientas software	1,24	1,10	1.00	0,91	0,83	-
Limitaciones de planificación del proyecto	1,23	1,08	1.00	1,04	1,10	-

Tabla 4.18 Ponderaciones para los 15 atributos de COCOMO Intermedio
Fuente: [Bedini, 2004]

Entonces tenemos:

FAE=(0.75)*(0.94)*(0.70)*(1.00)*(1.00)*(0.87)*(0.87)*(0.71)*(0.91)*(0.86)*(0.90)*(0.95)*(0.91)*(0.91)*(1.00)

FAE = 0.14695107

Justificación de los valores

Atributos de software

- ✓ Fiabilidad requerida del software: Indica las posibles consecuencias para el usuario en el caso que todavía existan defectos en el producto. (Valoración Muy Baja).
- ✓ Tamaño de la base de datos: La base de datos de nuestro producto será de tipo estándar (Valoración Baja).
- ✓ Complejidad del producto: La aplicación no va a realizar cálculos complejos (Valoración Muy Baja).

Atributos de hardware

- ✓ Restricciones del tiempo de ejecución: En los requerimientos se exige alto rendimiento (Valoración Nominal).
- ✓ Restricciones del almacenamiento principal: No hay restricciones al respecto (Valoración Nominal).
- ✓ Volatilidad de la máquina virtual: (Valoración Baja).
- ✓ Tiempo de respuesta del ordenador: Deberá ser interactivo con el usuario (Valoración Baja).

Atributos del personal

- ✓ Capacidad del analista: Capacidad alta relativamente, debido a la experiencia en análisis en proyecto similar (Valoración Muy Alta)

- ✓ Experiencia en la aplicación: Se tiene cierta experiencia en aplicaciones de esta envergadura (Valoración Alta).
- ✓ Capacidad de los programadores: Teóricamente deberá tenerse una capacidad alta por la experiencia en anteriores proyectos similares (Valoración Alta).
- ✓ Experiencia con la maquina virtual: (Valoración Alta).
- ✓ Experiencia en el lenguaje de programación: Es relativamente alta, dado que se controlan las nociones básicas y las propias del proyecto (Valoración Alta).

Atributos del proyecto

- ✓ Prácticas de programación modernas: Se usarán prácticas de programación mayormente convencional (Valoración Alta).
- ✓ Utilización de herramientas software: Se usarán herramientas estándar que no exigirán apenas formación, de las cuales se tiene cierta experiencia (Valoración Alta).
- ✓ Limitaciones de planificación del proyecto: Existen pocos límites de planificación. (Valoración Nominal).

Cálculo del esfuerzo del desarrollo:

$$E = a * (KLDC) \quad (KLDC)^{b} * FAE$$

$$E = 3 * (66)^{1,12} * 0.14695107$$

E = **48.1 hombres/mes**

Cálculo tiempo de desarrollo:

$$T = c * E^{d}$$

$$T = 2{,}5 * (48{,}1)^{0.35}$$

$$\boldsymbol{T} = \mathbf{9.7\ meses}$$

Productividad:

PR = LDC/Esfuerzo

PR = 66000/48.1

PR = 1372.14 LDC/personas mes

Personal promedio:

P = E/T

P = 48.1/9.7

$\boldsymbol{P}$ **= 4.96 personas**

Costo del Provecto:

Costo = (P * T) * Salario promedio de programadores y analistas. Donde el salario promedio de un programador es de $us 700. Entonces:

Costo = (4.96 * 9.7) * 700

Costo Total del Proyecto= 33.678.4 $us

Finalizando la utilización de la metodología se puede concluir que es necesario un equipo de 5 personas, trabajando alrededor de 10 meses, teniendo como costo total final del proyecto 33.768.4 $us.

CAPITULO VI

CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

Al haber concluido con la implementación del sistema la "Generación de Servidores de Metadatos con el Protocolo OAI-PMH Caso: Biblioteca Central de la Universidad Mayor de San Andrés" se ha alcanzado a realizar en su totalidad los objetivos planeados en el presente Proyecto de Grado por lo tanto se puede dar las siguientes conclusiones:

- ✓ Gracias a la flexibilidad provista por el protocolo OAP-PMH OAI 2.0 fue posible generar este proyecto para la Universidad Mayor de San Andrés vistas de un repositorio de metadatos académicos, tarea que no era posible previamente y que limitaba las vías de interoperabilidad para servidores OAI basados en la idea de OpenAcces.

- ✓ Cumplimos con la calidad y formato establecidos por una o más directrices para la implementación de un servidor OAI.

- ✓ Se logró integrar a la Universidad Mayor de San Andrés a la red OAISTER, WorldCAT de la OCLC redes internacionales de metadatos académicos.

- ✓ Con la implementación del sistema se garantiza la protección de los derechos de autoría del documento en la universidad.

- ✓ Se construyó el repositorio de colecciones de manera ordenada de acuerdo a las áreas del conocimiento por facultades y carreras existentes.

- ✓ Compartiremos recursos con otros repositorios de materiales académicos. Por ejemplo se podría construir un contexto que exponga todos los recursos de una

temática particular para que lo coseche un agregador temático o incluso de una tipología en particular para que los coseche un agregador de proyectos de grado , Tesis, Revistas o Libros.

- ✓ Una conclusión importante es el resguardo de los derechos de autor, al ser publico el documento generado por el autor en el estándar de Dublin Core especifica la fecha de creación del documento.
- ✓ Finalmente, con la metodología Website QEM se pudo lograr la formalización de la evaluación de la calidad del sistema. Tras la valoración obtenida aplicando la metodología se obtuvo un 83.13%, podemos asegurar que el producto elaborado es satisfactorio y comprensible para los usuarios del sistema.

5.2 RECOMENDACIONES

Con la implementación del sistema y analizando los alcances obtenidos se recomienda:

- ✓ Se recomienda realizar copias de seguridad de la base de datos así mismo el directorio donde se almacena los documentos digitales.
- ✓ También se recomienda que nuestro servidor de metadatos con el protocolo OAI-PMH sea registrado principalmente en OAISTER y WORLCAT, como institución académica con resolución univertitaria.
- ✓ Se recomienda organizar las colecciones por facultades y/o carreras, y también por áreas del conocimiento.
- ✓ Desarrollar un módulo para la integración del Servidor de Metadatos OAI-PMH con el sistema KOHA sistema de gestión de bibliotecas para exponer los libros y referencia de las bibliotecas virtuales de la universidad a través de un WebService.

BIBLIOGRAFIA

[1] Rusch-Feja, D. “The Open Archives Initiative and the OAI Protocol for Metadata Harvesting: rapidly forming a new tier in the scholarly communication infrastructure.” Learned, Publishing, 15(3), 179-186.

[2] R, Wolfe, W. Reilly. DSpace METS document profile for submission information packages (SIP).2006.

[3] Schatz BR, Chen H. Introduction to the Special Issue on Building Large-Scale Digital Libraries. IEEE Computer 1996; 29(5):22-7.

[4] Graves, M. 2012. Designing XML Databases. Prentice Hall, New Jersey.

[5] Fernández, L., Sánchez, J. A. (2003). “Community Tales: An infrastructure for the collaborative construction of digital theses repositories”. Proceedings of the Sixth International Conference on Electronic Theses and Dissertations (ETD 2003, Berlin, Germany, May).

[6] García, P. 2012. Consulta a textos digitalizados: implementación y análisis en el contexto de las colecciones especiales de la UDLA. B. Eng. Thesis. Department of Computer Systems Engineering, Universidad de las Américas Puebla.

[7] Feria Basurto L (comp.). Bibliotecas digitales. Colima: Universidad de Colima, 2002. p. 8-12.

[8] Vega Belmonte A. Web de calidad. La Habana: Editorial Científico-Técnica, 2001. p. 14-5.

[9] Coutín Domínguez A. Arquitectura de información para sitios web. Madrid: Anaya Multimedia, 2002. p. 259-61.

[10] Bates MJ. Idea tactics. Journal of the American Society for Information Science 1979;30(5):280-9.

[11] Sequeira D. Bibliotecas y archivos virtuales. Argentina: Nuevo Paradigma, 2011. p. 105.

[12] Harper,Corey A. Dublin Core metadata InItIatIve:beyond the element set. National

information standards Organization(NISO), winter 2010.

[13] Norman, Walsh. Extensible Markup Language (XML) 1.0 (Fifth Edition). W3C Recommendation, 26 November 2013.

[14] R, Jones, Stuart Lewis. Sword 2.0 Profile. swordapp.org, 20 de Marzo de 2012.

REFERENCIAS WEB

[1] [OAI - Protocol for Metadata Harvesting]
https://www.openarchives.org/pmh/

[2] Open Archives In a Box
http://dlt.ncsa.uiuc.edu/oaib/

[3] OAISTER
http://www.oaister.org

[4] Dublin Core Referencia
http://www.hipertexto.info/documentos/dublin_core.htm

[5] OCLC
http://www.oclc.org/

[6] WorlCAT
https://www.worldcat.org/

[7] Virginia Tech OAI
http://www.dlib.vt.edu/projects/OAI/software/vtoai/vtoai.html/

[8] OAIbiblio
http://www.ibiblio.org/oaibiblio/

[9] Rapid Visual OAI Tool
http://rvot.sourceforge.net/

[10] OAICat
http://www.oclc.org/research/software/oai/cat.htm/

[11] Dublín Core Metadata Element Set
http://dublincore.org

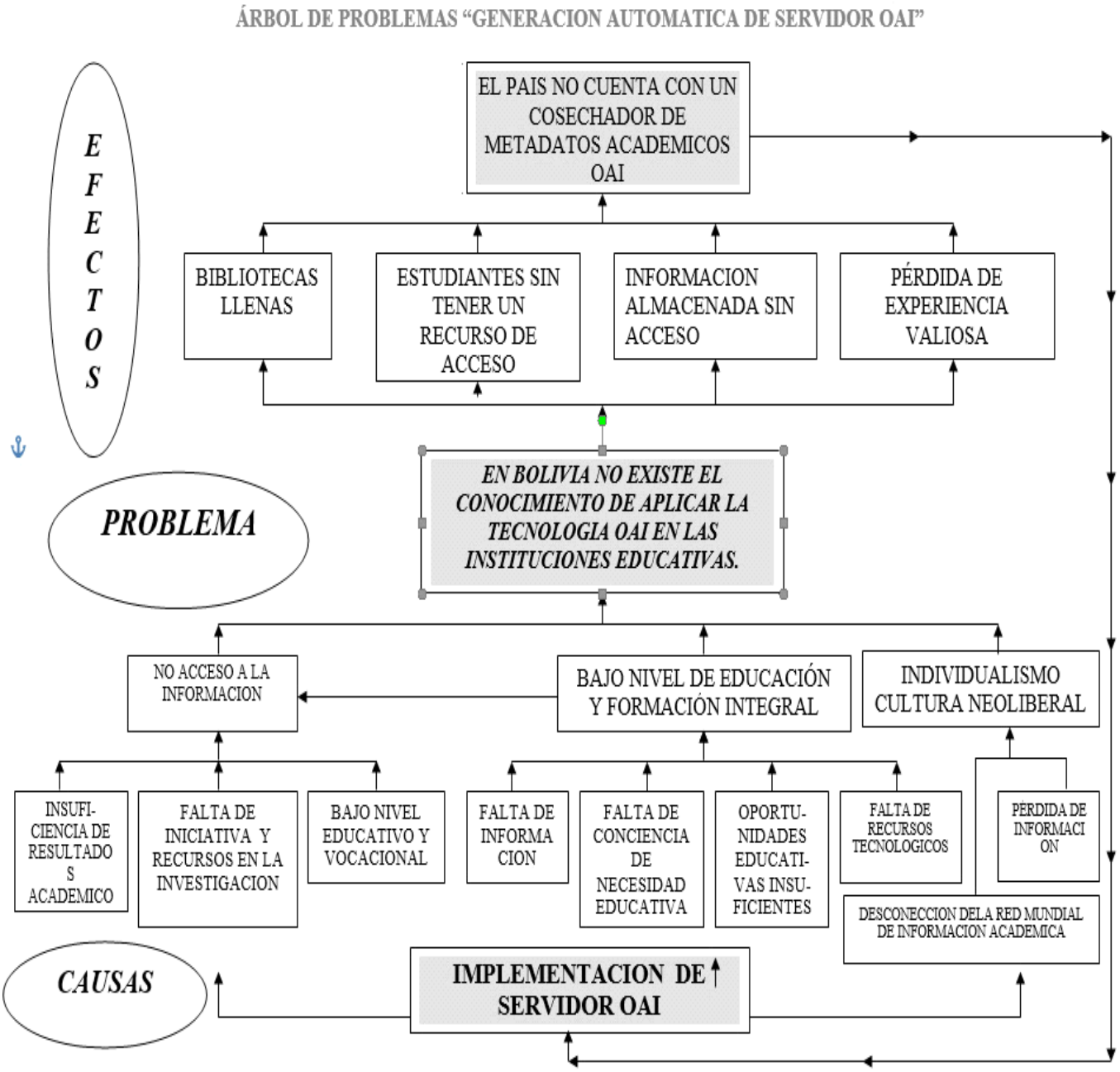
[12] Apache Jakarta Tomcat
http://jakarta.apache.org/tomcat/

[13] Dublin Core - María Jesús Lamarca,
http://www.hipertexto.info/documentos/dublin_core.htm

ANEXOS

ANEXO A

ARBOL DE PROBLEMAS



ANEXO B

MARCO LOGICO

RESUMEN NARRATIVO DE OBJETIVOS	INDICADORES VERIFICABLES OBJETIVAMENTE	MEDIOS DE VERIFICACION	SUPUESTOS
FIN El PRYECTO AYUDARA A LA BIBLIOTECA CENTRAL DE LA UMSA A CONTAR CON UN SERVIDOR OAI	LA CANTIDAD DE VISITA A CADA DOCUMENTO.	EN LOS BUSCADORES DE METADATOS, DEBERA REFLEJAR NUESTROS DOCUMENTOS PUBLICADOS. MEDIOS COMO **OCLC** **OAISTER**	LA CONTINUIDAD DE ESTE PROYECTO ES RESPONSABILIDAD DE LA BIBLIOTECA CENTRAL, YA QUE UNA VEZ REGISTRADO A LA RED INTERNACIONAL COMO INSTITUCION OTRO NO PODRA REGISTRARLO DE NUEVO SOLO SE HACE UNA VEZ..
PROPOSITO EL PROPOSITO FINAL DEL PROYECTO ES PUBLICAR DOCUMENTOS DE INVESTIGACION ACADEMICA A TRAVES DEL SERVIDOR OAI	LA INTEGRACION A REDES INTERNACIONALES DE METADATOS.	NUESTRO SERVIDOR DEBE SER RECONOCIDO E INTEGRADO A LA RED INTERNACIONAL DE COSECHADORES DE METADATOS COMO OCLC Y AOISTER.	EL PROYECTO DEBE DESARROLLARSE DE ACUERDO A LA PLANIFICACION FUIURA DE LA BIBLIOTECA CENTRAL PARA EL SOPORTE Y LA SOSTENIBILIDAD DEL PROYECTO.
COMPONENTES LOS COMPONENETS DE ESTE PROYECTO SERAN DISPUESTAS POR LA BIBLIOTECA CENTRAL, COMO EQUIPOS DE COMPUTACION, DOCUMENTOS A PUBLICAR.	CANTIDAD DE DOCUMENTOS PUBLICADOS CANTIDAD DE VISITAS POR DOCUMENTOS CANTIDAD DE REDES QUE INTEGRAN A NUESTRO SERVIDOR OAI. CANTIDAD DE SERVIDORES GENERADOS. .	EL QUE CALIFICA ESTE PROYECTO PODRA VERIFICAR A TRAVES DE LA RED DE INTERNET LA EFECTIVIDAD DEL PROYECTO.	LOS RECURSOS QUE DISPONGA LÑA BIBLIOTECA CENTRAL AYUDARAN AL LOGRO DEL OBJETIVO DEL PROYECTO PARA LA GENERACION DEL SERVIDOR OAI.
ACTIVIDADES (TAREAS) LAS ACTIVIDADES QUE DESARROLLAREMOS CON EL PROYECTO ES GENERAR UN DIRECTORIO DE DOCUMENTOS A PUBLICAR , ESTRUCTURAR LA BASE DE ESTRUCTURA PARA LA GENERACION DEL SERVIDOR OAI	EL PRESUPUESTO DEL PROYECTO SERA DISPUESTO POR LA BIBLIOTECA CENTRAL, COMO GASTOS DE OJAS, GASTO DE ENERGIA, GASTO DE PASAJES PARA ASISTENCIA A LA BIBLIOTEAC CENTRAL.	EL EVALUADOR CONSULTARA A LA BIBLIOTECA CENTRAL SOBRE LOS GASTOS Y RECURSOS USADOS PARA LA EJECUCION DEL PROYECTO.	EL SERVIDOR OAI SE INSTALARAY GENERARA POR UNA SOLA VEZ EN LA BIBLIOTECA CENTRAL EL MISMO DEBERA TENER RESPALDOS PARA LA SOSTENIBILIDAD.

ANEXO C

FORMATO DE METADATOS DUBLIN CORE

ETIQUETA DEL ELEMENTO DUBLIN CORE http://dublincore.org/documents/dcmi-terms/	DESCRIPCIÓN
DC. Title	**Título: El nombre dado a un recurso.** Típicamente, un título es el nombre formal por el que es conocido el recurso.
DC. Creator	**Autor: La entidad primariamente responsable de la creación del contenido intelectual del recurso.** Entre los ejemplos de un creador se incluyen una persona, una organización o un servicio. Típicamente, el nombre del creador podría usarse para indicar la entidad.
DC. Subject	**Materias y palabras clave: El tema del contenido del recurso.** Un tema será expresado como palabras clave, frases clave o códigos de clasificación que describan el tema de un recurso. Se recomienda seleccionar un valor de un vocabulario controlado o un esquema de clasificación formal.
DC. Description	**Descripción: La descripción del contenido del recurso.** La descripción puede incluir, pero no se limita a: un resumen, tabla de contenidos, referencia a una representación gráfica de contenido o una descripción de texto libre del contenido.
DC. Publisher	**Editor: La entidad responsable de hacer que el recurso se encuentre disponible.** Ejemplos de editores son una persona, una organización o un servicio. Típicamente, el nombre de un editor podría usarse para indicar la entidad.
DC. Contributor	**Colaborador. La entidad responsable de hacer colaboraciones al contenido del recurso.** Ejemplos de colaboradores son una persona, una organización o un servicio. Típicamente, el nombre del colaborador podría usarse para indicar la entidad.
DC. Date	**Fecha: Una fecha asociada con un evento en el ciclo de vida del recurso.** Típicamente, la fecha será asociada con la creación o disponibilidad del recurso. Se recomienda utilizar un valor de datos codificado definido en el documento "Date and Time Formats",http://www.w3.org/TR/NOTE-datetime que sigue la norma ISO 8601 que sigue el formato YYYY-MM-DD.
DC. Type	**Tipo: la naturaleza o categoría del contenido del recurso.** El tipo incluye términos que describen las categorías generales, funciones, géneros o niveles de agregación del contenido. Se recomienda seleccionar un valor de un vocabulario controlado (por ejemplo, el *DCMI Vocabulary* -DCMITYPE-

	http://dublincore.org/documents/dcmi-type-vocabulary/). Para describir la manifestación física o digital del recurso, se usa el elemento Formato.
DC. Format	**Formato: la manifestación física o digital del recurso.** El formato puede incluir el tipo de media o dimensiones del recurso. Podría usarse para determinar el *software*, *hardware* u otro equipamiento necesario para ejecutar u operar con el recurso. Ejemplos de las dimensiones son el tamaño y la duración. Se recomienda seleccionar un valor de un vocabulario controlado (por ejemplo, la lista de Internet Media Types (MIME) que define los formatos de medios de ordenador).
DC. Identifier	**Identificación: Una referencia no ambigua para el recurso dentro de un contexto dado.** Se recomienda identificar el recurso por medio de una cadena de números de conformidad con un sistema de identificación formal, tal como un URI (que incluye el Uniform Resource Locator -URL, el Digital Object Identifier (DOI) y el International Standard Book Number (ISBN).
DC. Source	**Fuente**: **Una referencia a un recurso del cual se deriva el recurso actual.** El recurso actual puede derivarse, en todo o en parte, de un recurso fuente. Se recomienda referenciar el recurso por medio de una cadena o número de conformidad con un sistema formal de identificación.
DC. Language	**Lengua**: **La lengua del contenido intelectual del recurso.** Se recomienda usar RFC 3066 http://www.ietf.org/rfc/rfc3066.txt en conjunción con la ISO 639 [ISO639] http://www.loc.gov/standards/iso639-2/, que define las etiquetas de dos y tres letras primarias para lenguaje, con subetiquetas opcionales. Ejemplo: "en" u "eng" para Inglés, "akk" para Acadio, y "en-GB" para inglés usado en Reino Unido.
DC. Relation	**Relación**: **Una referencia a un recurso relacionado.** Se recomienda referenciar el recurso por medio de una cadena de números de acuerdo con un sistemas de identificación formal.
DC. Coverage	**Cobertura: La extensión o ámbito del contenido del recurso.** La cobertura incluiría la localización espacial (un nombre de lugar o coordenadas geográficas), el período temporal (una etiqueta del período, fecha o rango de datos) o jurisdicción (tal como el nombre de una entidad administrativa). Se recomienda seleccionar un valor de un vocabulario controlado (por ejemplo, del Thesaurus of Geographic Names (TGN) y que, donde sea apropiado, se usen preferentemente los nombres de lugares o períodos de tiempo antes que los identificadores numéricos tales como un conjunto de coordenadas o rangos de datos.
DC. Rights	**Derechos**: **La información sobre los derechos de propiedad y sobre el recurso.** Este elemento podrá contener un estamento de gestión de derechos para el recurso, o referencia a un servicio que provea tal información. La información sobre derechos a menudo corresponde a los derechos de propiedad intelectual, copyright y otros derechos

	de propiedad.

ANEXO D

CASOS DE USO

Solicitar Publicación

Casos de uso:	Solicitar publicación
Actor	Autor y Bibliotecario
Propósito:	Realizar la solicitud de la publicación del documento.
Tipo:	Primario
Descripción:	El autor solicita la Publicación del documento en el servidor OAI, luego el bibliotecario elabora el registro para su posterior publicación.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- El caso de uso tiene su inicio cuando el autor solicita la publicación del documento. 3.- luego llena el formulario de registro con sus respectivos datos verificando que este correctamente llenado con los datos completos.	2.- el bibliotecario registra lo documento y datos del autor, selecciona la colección correcta para su publicación.

Registra Datos del Autor

Casos de uso:	Registra datos del autor
Actor	Bibliotecario
Propósito:	Registra al autor que quiere la publicación del documento.
Tipo:	Primario
Descripción:	Se realiza el registro del autor para una mejor información y referencia del autor.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- El caso de uso tiene su inicio cuando el bibliotecario realiza el registro del autor con sus datos completos y referencias	2.- los datos son registrados manualmente en libros de registro que

autor solicita la publicación del documento. 3.- El bibliotecario guarda la copia como registro y al autor se extiende un comprobante de registro original.	son hojas con copia que sirven para tener un registro del título y el autor y sirve para el control de inventarios.

Registrar datos del documento

Casos de uso:	Registra datos del documento
Actor	Bibliotecario
Propósito:	Se realiza el registro del documento.
Tipo:	Primario
Descripción:	Se realiza el registro del documento para un mejor control de datos.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- El caso de uso tiene su inicio cuando el bibliotecario realiza el registro del documento: título del documento autor al que corresponde y otros datos. 3.- El bibliotecario se encarga de guardar los datos del documento y su autor en copias para un inventario adecuado.	2.- los datos del documento son registrados manualmente con sus respectivos datos, en formularios que contienen copias.

Realizar la Catalogación

Casos de uso:	Realizar la Catalogación
Actor	Bibliotecario
Propósito:	Realiza por el titulo la catalogación en la colección al que corresponde.
Tipo:	Primario
Descripción:	Se realiza la catalogación del documento en la colección seleccionada.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- En este caso de uso el bibliotecario elabora la catalogación del documento. 3.- Luego de realizar la catalogación se guarda una copia de respaldo por el bibliotecario.	2.- Se elabora la catalogación del documento, se busca la colección al que pertenece el documento y se llena con los respectivos datos del documento.

Consultar Duplicidad o Plagio del Titulo

Casos de uso:	Consultar Duplicidad o Plagio del Titulo
Actor	Bibliotecario
Propósito:	Realiza por el titulo la verificación del plagio o duplicidad del contenido.
Tipo:	Primario
Descripción:	Se realiza la consulta a la base de datos de la existencia de otro título similar y contenido plagiado.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- En este caso de uso el bibliotecario verifica la existencia de otro título Similar. 3.- Luego de realizar la verificación se procede a la catalogación.	2.- Se elabora la verificación del documento, se busca en la colección al que pertenece el documento.

Catalogación del Documento

Casos de uso:	Catalogación del Documento
Actor	Bibliotecario
Propósito:	Realiza por el titulo la catalogación del documento en Dublín Core.
Tipo:	Primario
Descripción:	Se realiza la verificación de la solicitud de publicación y se describen los datos del documento en el formato Dublín Core.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- En este caso de uso el bibliotecario verifica la solicitud de la publicación. 3.- Luego se envía un correo al autor informando que su documento ha sido dado de alta en el servidor OAI.	2.- Se elabora el registro de los datos del documento en formato Dublín Core. Esto se registra en el sistema.

Registro de Solicitud de Publicación

Casos de uso:	Registro de solicitud de publicación
Actor	Bibliotecario y Autor
Propósito:	Verifica la solicitud de la publicación
Tipo:	Primario
Descripción:	El bibliotecario verifica el documento de publicación y los datos de la solicitud para luego proceder con el llenado del formulario de solicitudes.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- El bibliotecario introduce su login y password. 3.- Elige la opción crear solicitud de publicación. 5.- El bibliotecario escoge por orden las solicitudes de publicación y los documentos a registrar.	2.-El sistema verifica si los datos introducidos son válidos y son los correctos. a) Si los datos son correctos se muestra la interfaz del menú principal de usuario, b) Si los datos no son correctos el sistema solicita nuevamente los datos y muestras los errores. 4.- Se muestra el menú de registro de solicitudes.6.- Si los datos son correctamente llenados se acepta la solicitud de publicación. 7.- Luego de realizar la solicitud completa se guarda y actualiza la base de datos.

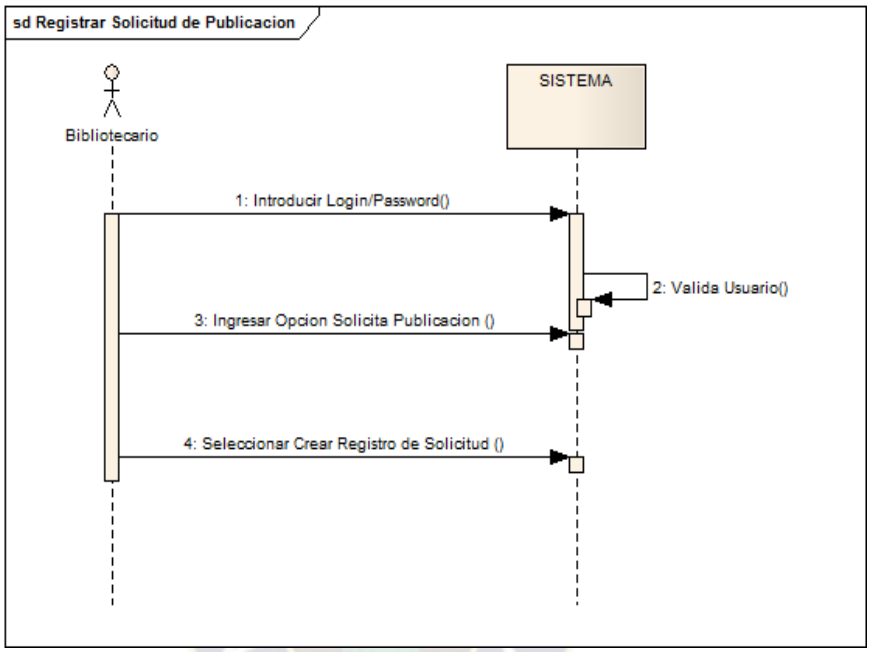
Realizar la catalogación del documento en Dublín Core

Casos de uso:	Realizar la catalogación del documento en Dublín Core
Actor	Bibliotecario
Propósito:	Catalogar todos los documentos.
Tipo:	Primario
Descripción:	Se realiza la catalogación de todos los documentos en Dublín Core de todas las solicitudes realizadas.
Curso Normal de Eventos	
Acción del actor	**Respuesta del sistema**
1.- El bibliotecario introduce su login y password. 3.- Elige la opción catalogar documento. 5.- El bibliotecario llena todos los datos del nuevo documento con reglas de catalogación e indización en el formato Dublín Core.	2.-El sistema verifica si los datos introducidos son válidos y son los correctos. a) Si los datos son correctos se muestra la interfaz del menú principal de usuario, b) Si los datos no son correctos el sistema solicita nuevamente los datos y muestras los errores. 4.- Muestra el formulario de catalogación de documentos. 6.- Si los datos de la catalogación son correctamente llenados se acepta la publicación. 7.- Luego de realizar la catalogación completa se guarda y publica el documento.

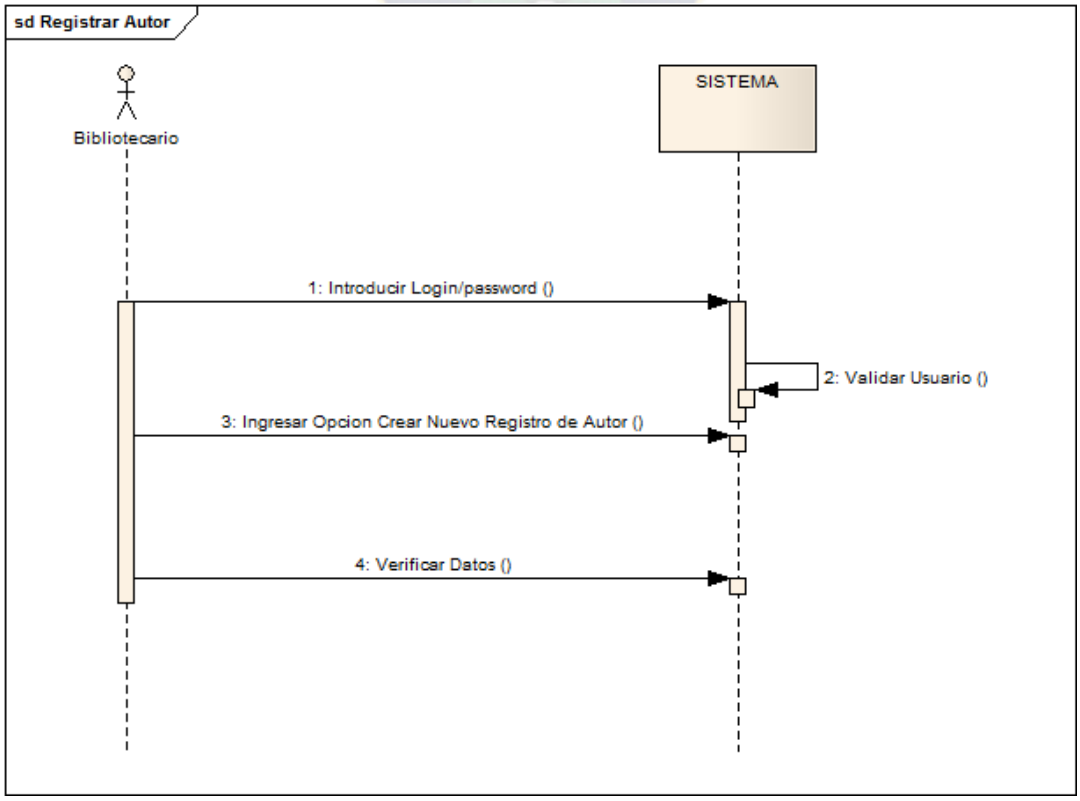
ANEXO E

DIAGRAMA DE SECUENCIA

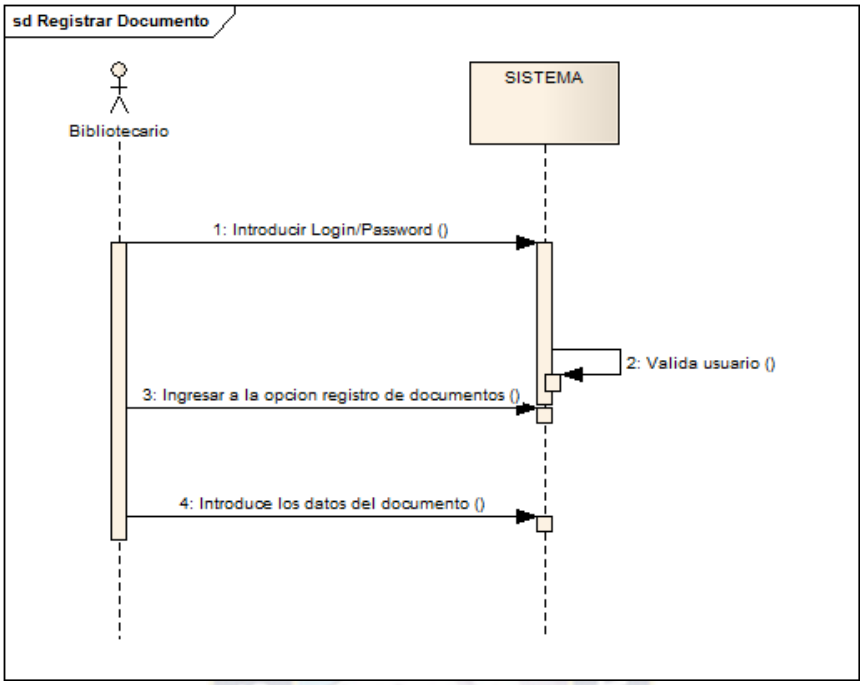
Registrar Solicitud de Publicación.



Registrar Autor

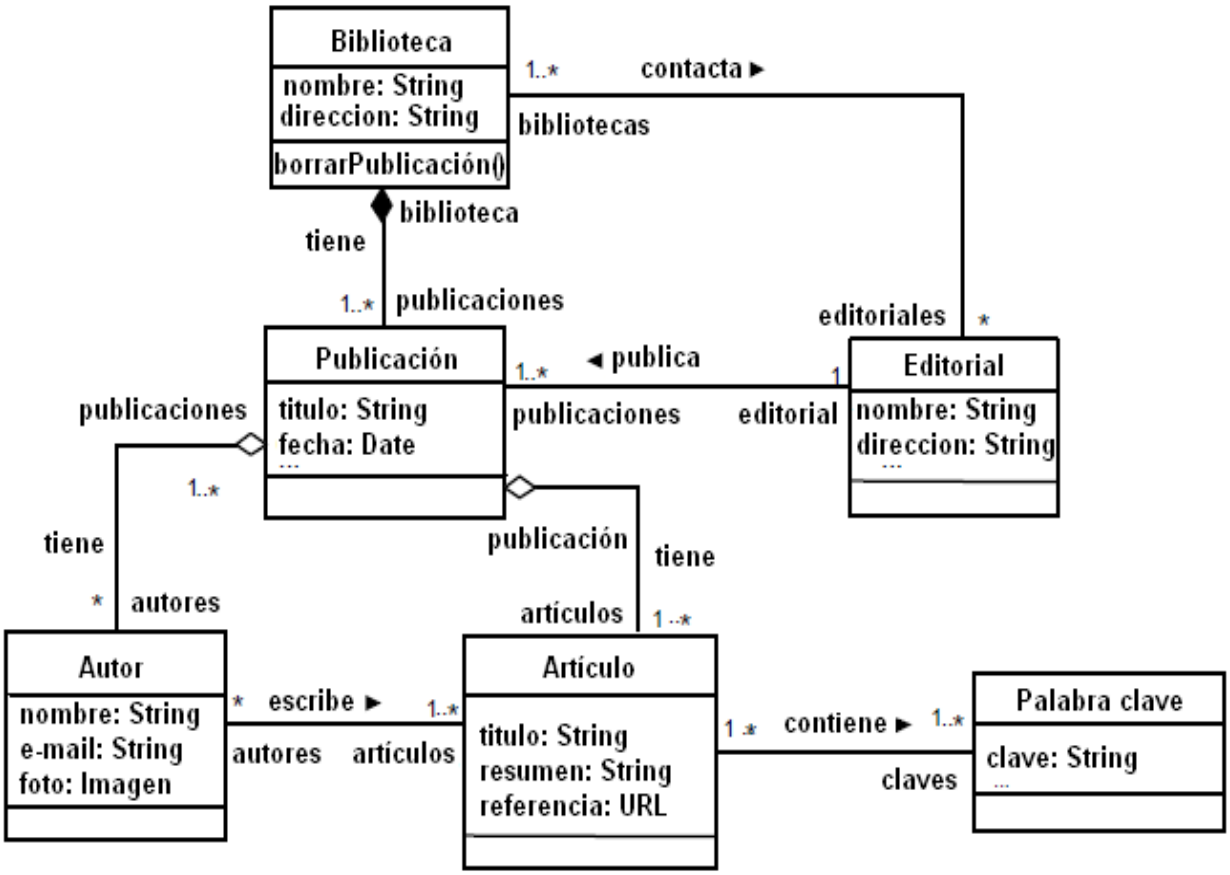


Registrar Documento



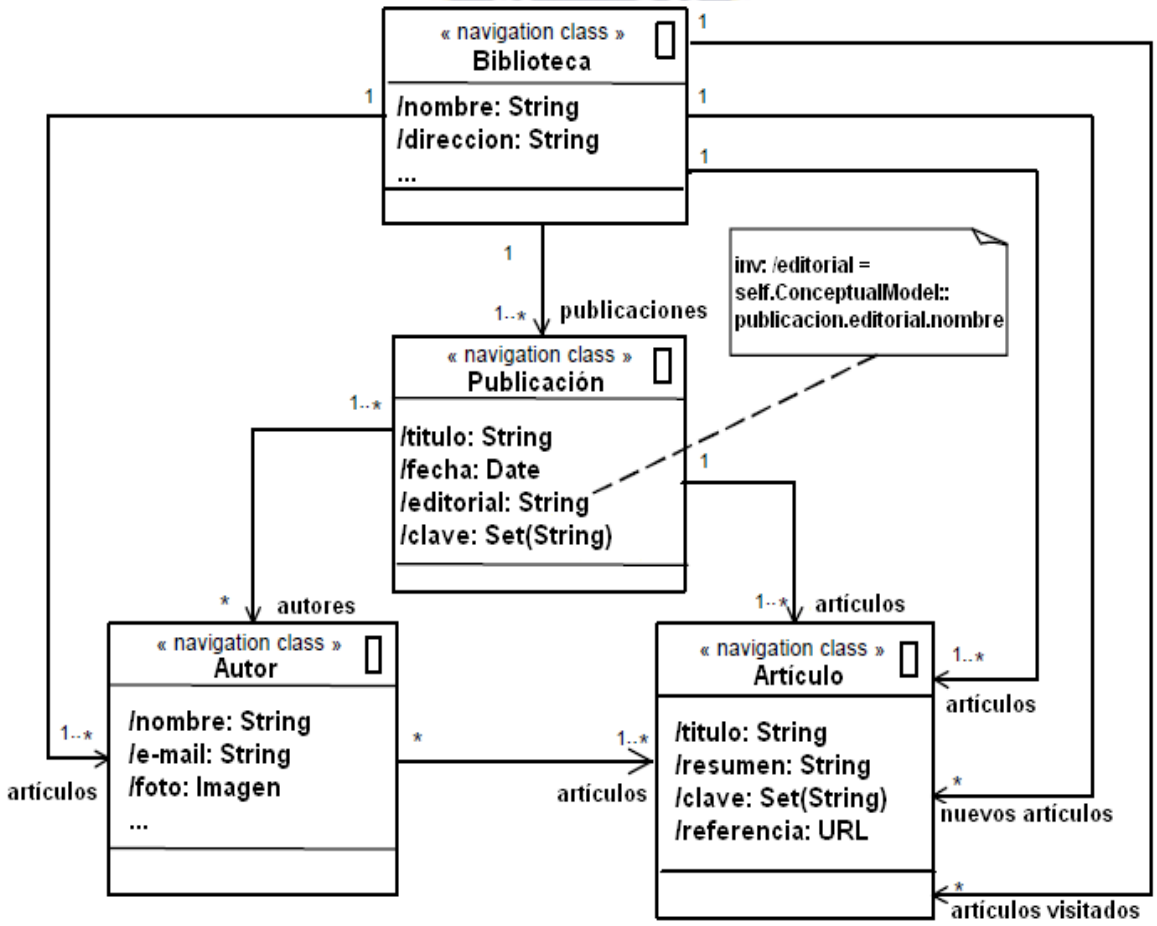
ANEXO F

MODELO CONCEPTUAL

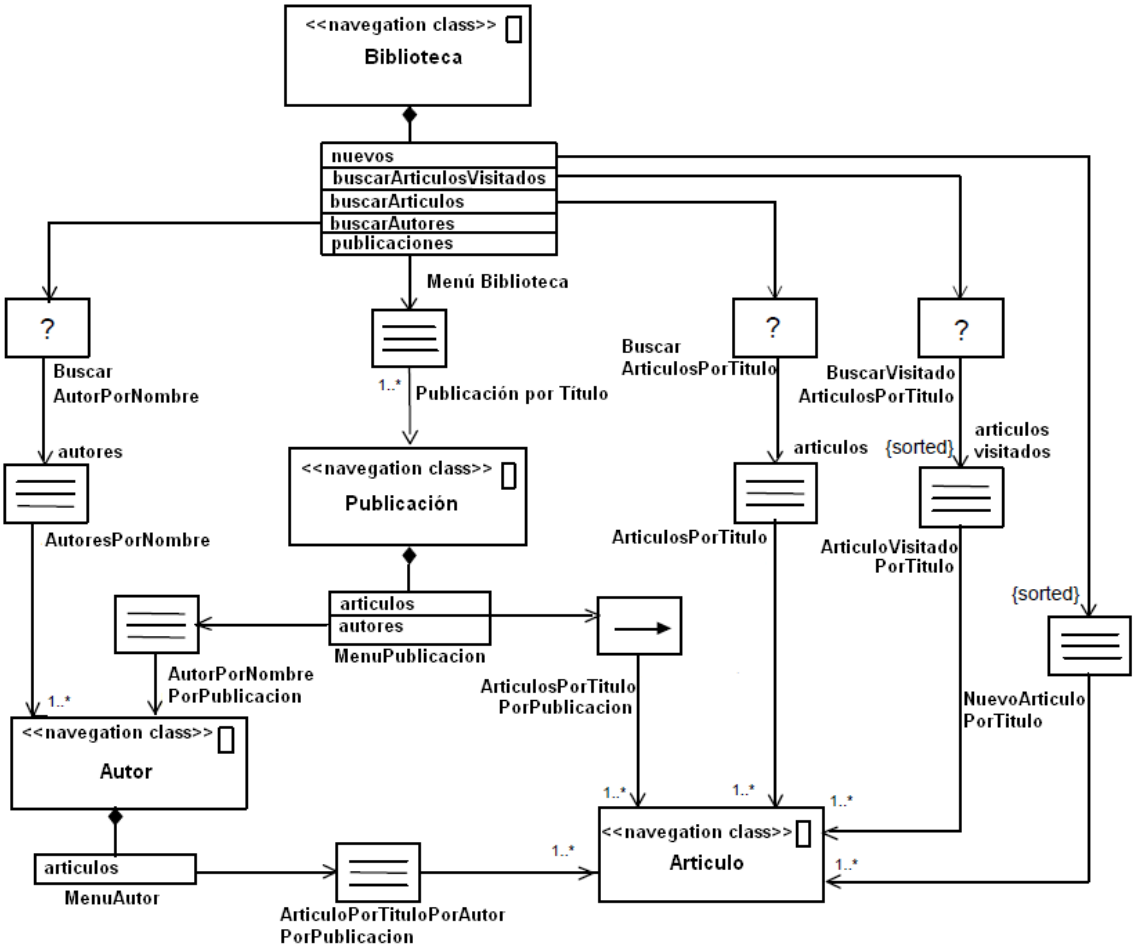


ANEXO G

MODELO DE NAVEGACIÓN

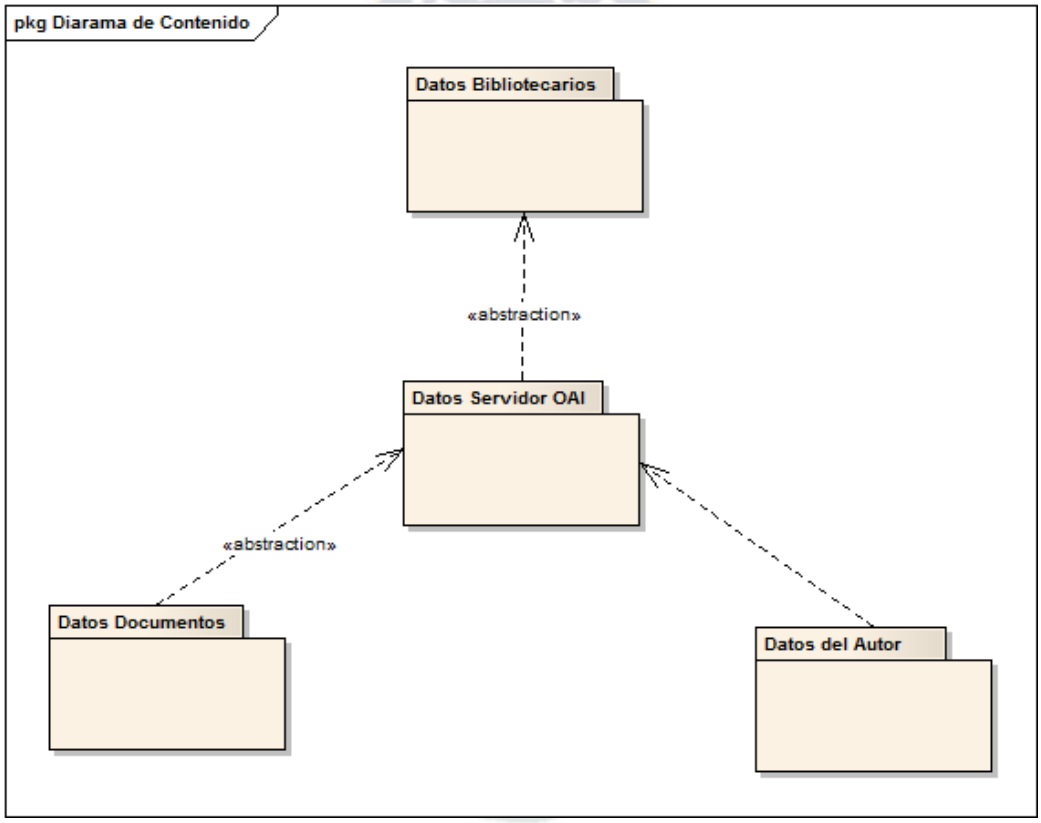


ANEXO G
MODELO DE ESTRUCTURA DE NAVEGACIÓN



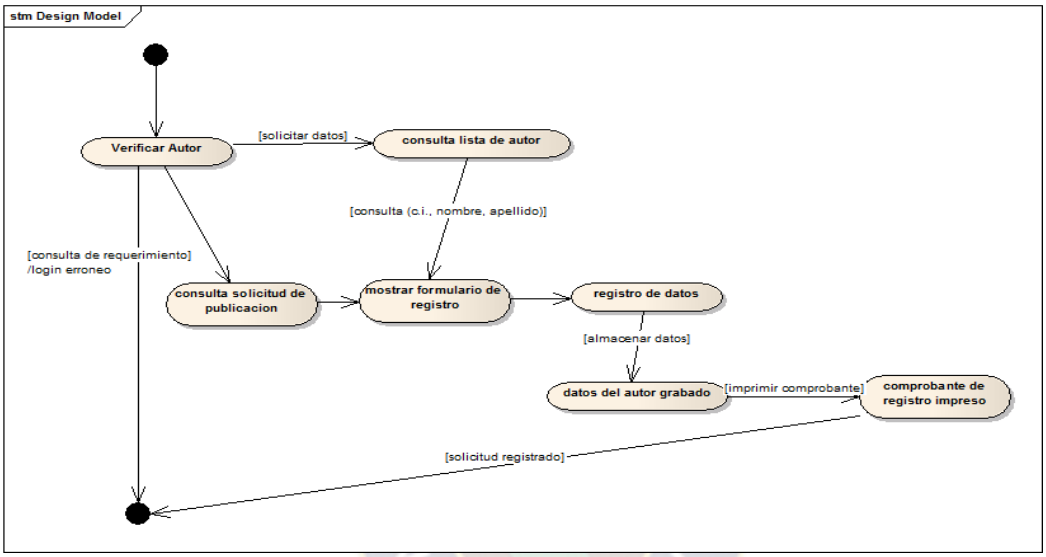
ANEXO H

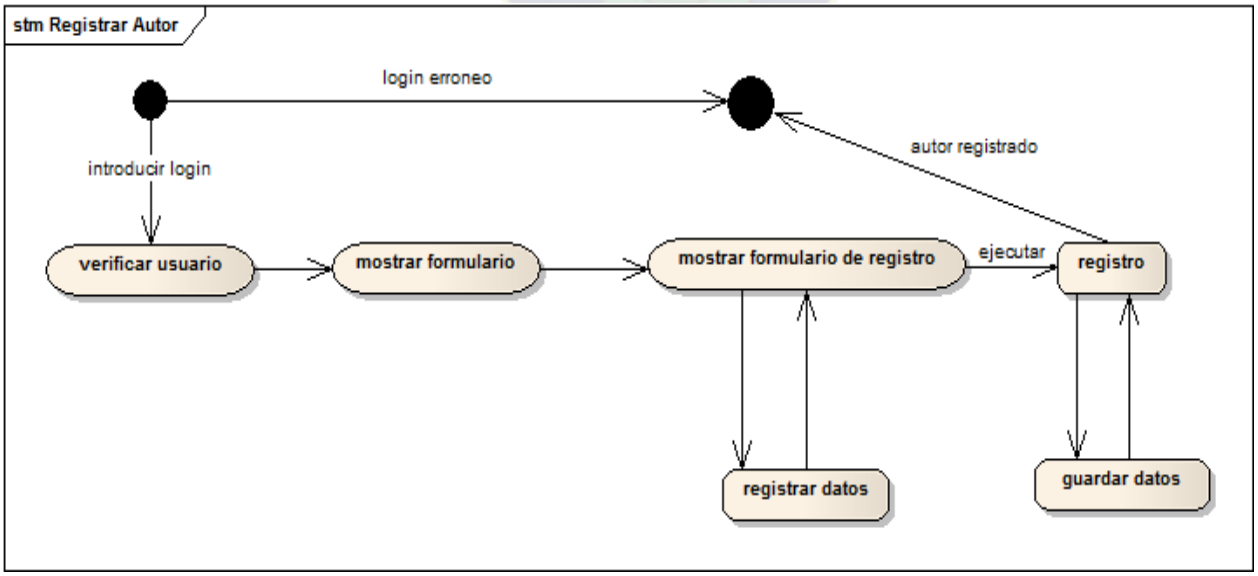
DIAGRAMA DE CONTENIDOS

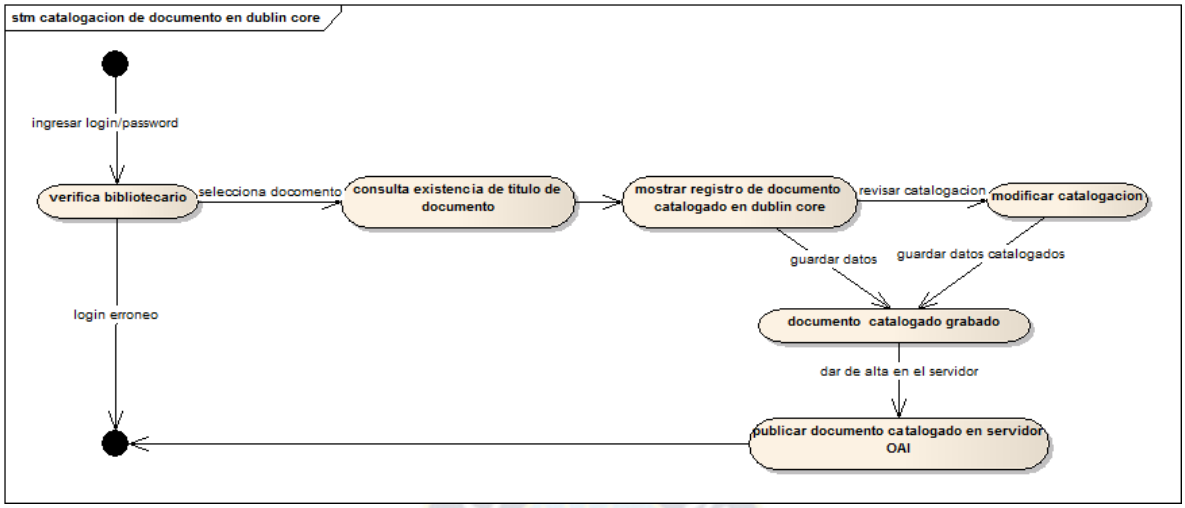


ANEXO I

DIAGRAMA DE ESTADO

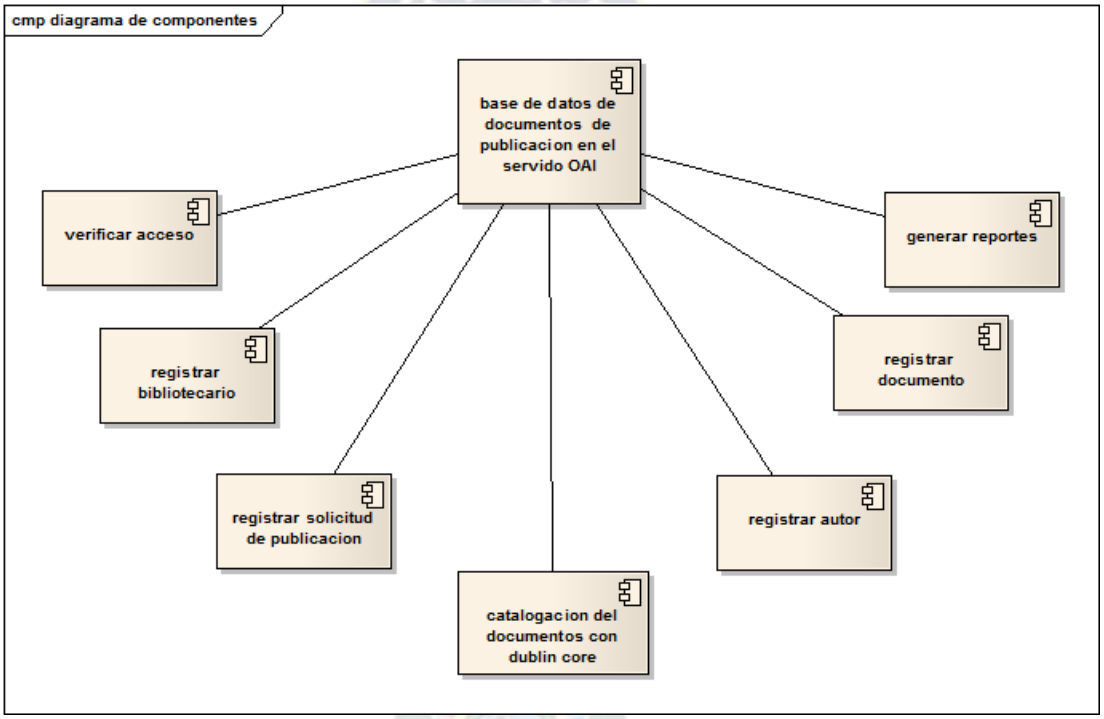





stm catalogacion de documento en dublin core
ingresar login/password
verifica bibliotecario
selecciona docomento
consulta existencia de titulo de documento
mostrar registro de documento catalogado en dublin core
revisar catalogacion
modificar catalogacion
guardar datos
guardar datos catalogados
login erroneo
documento catalogado grabado
dar de alta en el servidor
publicar documento catalogado en servidor OAI

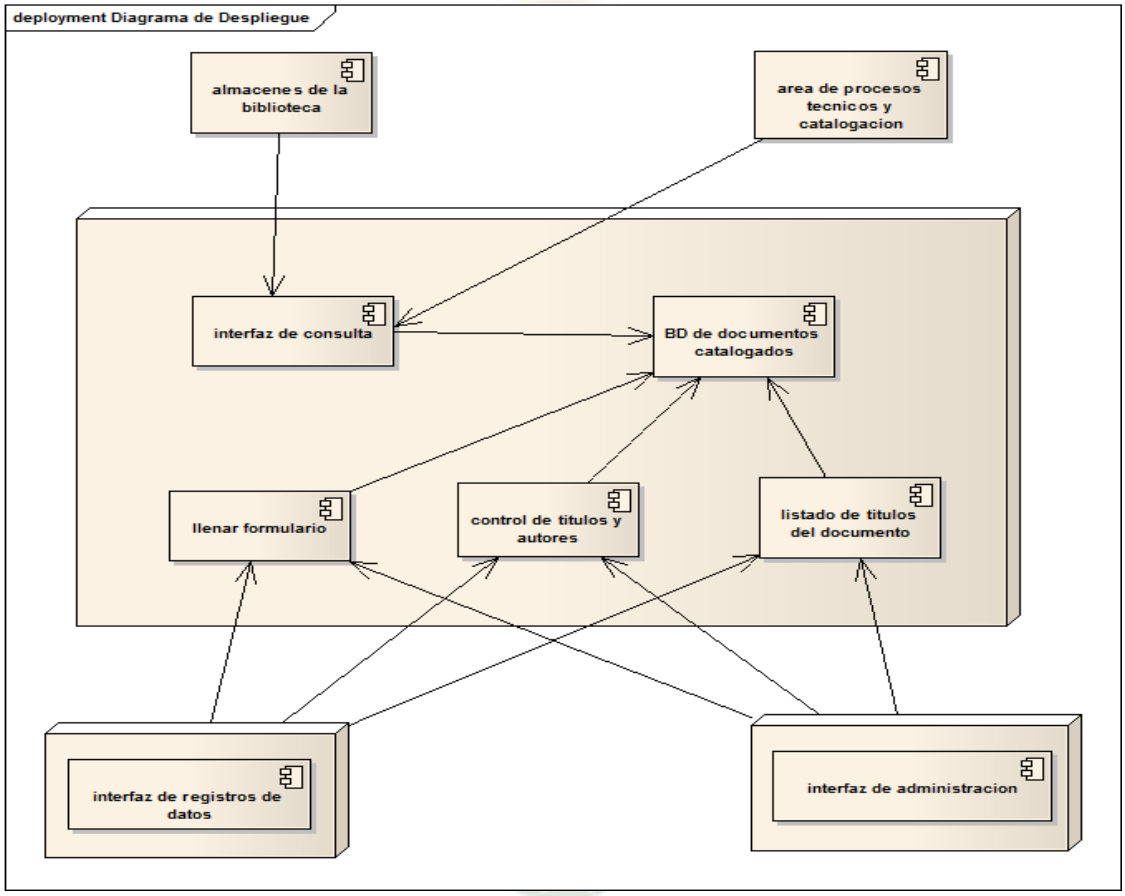
ANEXO I

DIAGRAMA DE COMPONENTES



ANEXO J

DIAGRAMA DE DESPLIEGUE



ANEXO K

CONTRATO DE OPERACIONES

Nombre:	Introducir login y password
Responsabilidades:	Despliega y confirma autenticación
Tipo:	Manual
Referencias:	Entrega la solicitud
Precondición:	el bibliotecario debe estar registrado en el sistema
Postcondición:	Proporciona la comprobante de registro de solicitud.

Nombre:	Solicita Publicación
Responsabilidades:	Despliega el formulario de solicitud de publicación
Tipo:	Sistema
Referencias:	Detalles del formulario de solicitud de publicación
Precondición:	El bibliotecario para registrar nueva solicitud de publicación
Postcondición:	Proporciona el formulario

Nombre:	Introducir login y password
Responsabilidades:	Despliega y confirma autenticación
Tipo:	Manual
Referencias:	Entrega de registro del autor
Precondición:	El usuario debe estar registrado en el sistema
Postcondición:	Proporciona el formulario de registro

Nombre:	Crear nuevo registro de solicitud
Responsabilidades:	Despliega nuevo registro de solicitud
Tipo:	Manual
Referencias:	Registra los datos para almacenar en la base de datos
Precondición:	El autor debe tener un comprobante de la solicitud
Postcondición:	Proporciona una información correcta del autor

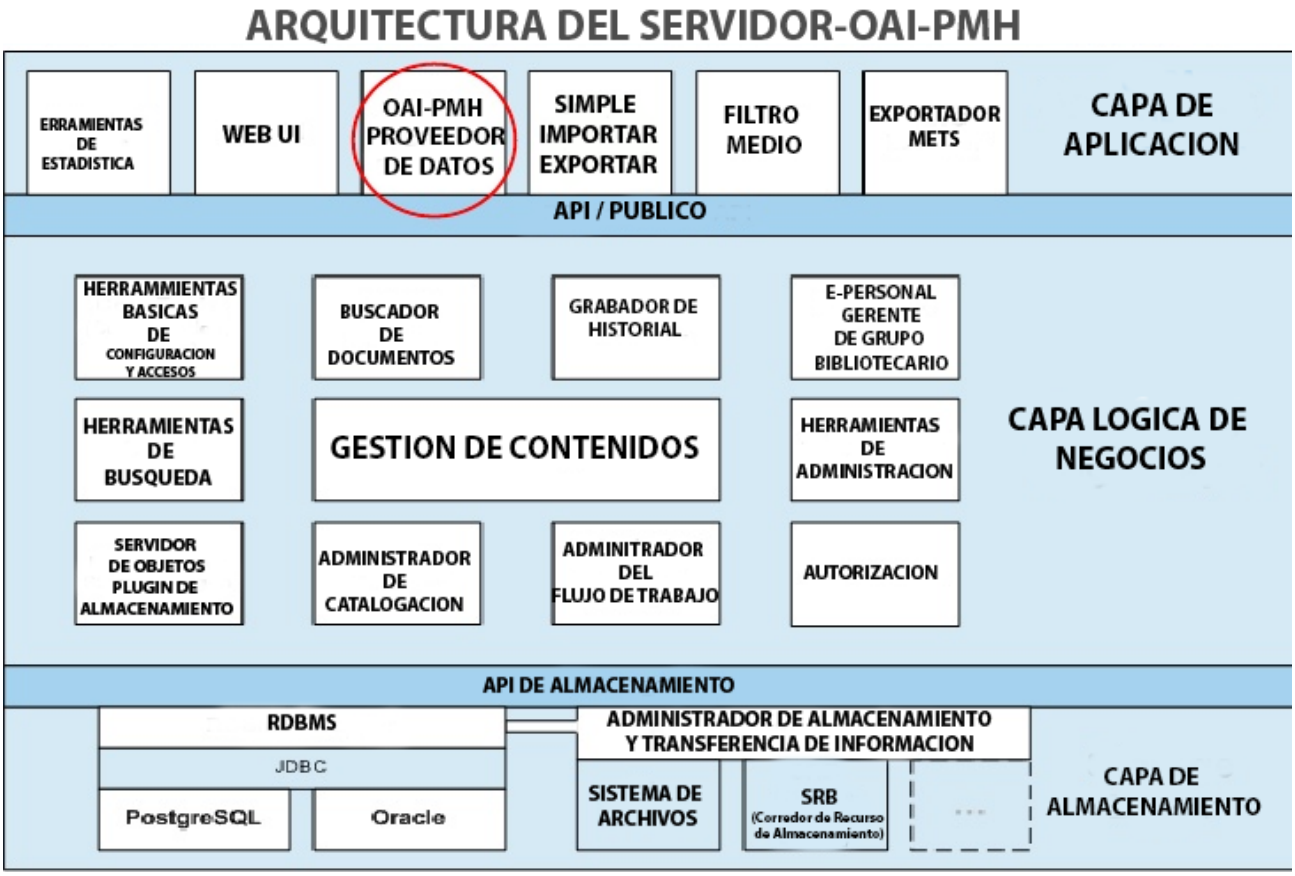
Nombre:	Verifica datos
Responsabilidades:	Se encarga de verificar los datos completos
Tipo:	Manual
Referencias:	Se verifica mediante el sistema los datos
Precondición:	EL sistema proporciona una base de datos actualizados
Postcondición:	Proporciona la actualización de los datos del autor

Nombre:	Introducir login y password
Responsabilidades:	Despliega y confirma autenticación
Tipo:	Manual
Referencias:	Entrega registro del autor
Precondición:	el bibliotecario debe estar registrado en el sistema
Postcondición:	Proporciona la proforma de registro

Nombre:	Registra el Documento
Responsabilidades:	Despliega el formulario del nuevo registro de documento
Tipo:	Sistema
Referencias:	Registra los datos del documento en la base de datos
Precondición:	El autor del documento debe tener un comprobante
Postcondición:	Proporciona una información correcta del documento

Nombre:	Realiza catalogación del documento en Dublin Core
Responsabilidades:	Se encarga de catalogar los datos completos del documento
Tipo:	Manual
Referencias:	Se verifica mediante el sistema datos completos y en norma
Precondición:	El sistema proporciona una base de datos actualizados
Postcondición:	Proporciona la actualización de los datos de la catalogación

ANEXO L
ARQUITECTURA DEL SISTEMA SERVIDOR-OAI



ANEXO M

DOCUMENTACION