PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR SEDE ESMERALDAS

INGENIERÍA DE SISTEMAS Y COMPUTACIÓN

TRABAJO DE GRADO:

SISTEMA DE BÚSQUEDA Y RECUPERACIÓN DE DOCUMENTOS, APLICADO A LOS INFORMES DE TESIS Y PROYECTOS DE GRADO DIGITALIZADOS DE LA PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR SEDE ESMERALDAS.

LÍNEA DE INVESTIGACIÓN:

PROGRAMACIÓN Y DESARROLLO DE SOFTWARE.

PREVIA A LA OBTENCIÓN DEL TÍTULO DE:

INGENIERÍA EN SISTEMA Y COMPUTACIÓN.

AUTORA:

RODRÍGUEZ HERNÁNDEZ CARLA FERNANDA

ASESOR:

MGT. JOSÉ LUIS CARVAJAL

JULIO / 2016

TRABAJO DE TITULACIÓN DE GRADO

Disertación aprobada luego de haber dado cumplimiento a los requisitos exigidos por el Reglamento de Grado de la PUCESE, previa a la obtención del Título de Ingeniero en Sistemas y Computación.

Mgt. José Luis Carvajal.

ASESOR.

Mgt. Guillermo Cedeño.

LECTOR 1.

Ing. Tania Posligua.

LECTOR 2.

Mgt. Xavier Quiñonez Ku

DIRECTOR DE ESCUELA.

AUTORÍA

Yo, Rodríguez Hernández Carla Fernanda, portadora de la cédula de ciudadanía Nº 080416414-3, declaro bajo juramento que la presente investigación es de total responsabilidad de la autora, y que se ha respetado las diferentes fuentes de información realizando las citas correspondientes.

Rodríguez Hernández Carla Fernanda.
AUTORA.

DEDICATORIA

A Dios nuestro creador, que como ser de luz ha sido mi acompañante iluminándome y guiándome durante todo el trayecto de mis años de estudio, dándome la fortaleza para poder culminar con éxito mi proyecto de grado y alcanzar esta gran meta como es la obtención de mi título profesional.

A mis padres, por su comprensión, paciencia y apoyo incondicional durante todo el transcurso de mi formación como profesional hasta culminar con mis estudios universitarios.

A mis hermanos, por ser parte fundamental de mi vida y por apoyarme anímicamente durante todo el lapso de esta importante etapa.

Y a todos los familiares y amigos que de una u otra forma contribuyeron para hacer este anhelo posible.

AGRADECIMIENTO

Mi agradecimiento a todas aquellas personas que han hecho posible la realización del presente proyecto de grado:

A mis padres, por depositar en mí su confianza, y darme su apoyo moral y económico, lo que hoy hace que pueda dar este gran paso hacia mi vida profesional.

Al Mgt. José Luis Carvajal, por brindarme su apoyo y ayuda necesaria durante toda la elaboración del proyecto de grado.

A los lectores del proyecto de grado, por cada una de sus sugerencias y observaciones, que sirvieron de mucha ayuda para dar finalización al proyecto final.

A la Pontificia Universidad Católica del Ecuador Sede Esmeraldas, por concederme todas las facilidades necesarias para que pudiera realizar este proyecto de grado.

RESUMEN

En toda organización o institución se maneja documentación valiosa, que debido a la evolución de la tecnología, se han visto en la obligación de digitalizarla para mayor organización y agilidad de los procesos que se llevan a cabo con respecto a esta.

Los sistemas de búsqueda son una herramienta factible para encontrar información idónea que se precisa obtener en una situación determinada, haciendo que el acceso a la información sea ágil y generando los resultados deseados.

Basándose en la problemática de que aún existen procesos en los cuales el manejo de la información y documentos no se realiza de manera que los usuarios o clientes tengan la facilidad de poder manejarlos o acceder a ellos, se planteó la propuesta a la Pontificia Universidad Católica Sede Esmeraldas de elaborar un sistema que permita realizar la búsqueda de los documentos de tesis o proyectos de grado digitalizados, de manera que los investigadores, es decir, los estudiantes, docentes y público en general puedan realizar consultas a estos documentos de forma que esté disponible en el momento que lo necesiten, y les permita ganar tiempo en la realización de sus investigaciones.

Para la realización de esta investigación se aplicaron los métodos Inductivo – Deductivo, que permitieron recopilar los datos y aspectos vinculados con el proyecto. Por lo cual se utilizaron técnicas como la entrevista, aplicada al encargado de la Biblioteca y al encargado del Departamento de Sistemas de la PUCESE. También se realizaron encuesta a 144 personas entre ellos estudiantes, profesores y miembros del CID (Centro de Investigación y Desarrollo).

De este modo se pudo captar la incomodidad de algunos al tener que realizar préstamos de estos documentos en Biblioteca, porque no pueden acceder a ellos con libertad.

Por lo que se considera que el proyecto es factible para la institución, porque proporcionara la facilidad de acceso que los investigadores necesitan al momento de disponerse hacer una consulta a este tipo de documentos.

ABSTRACT

In any organization or institution is handled valuable documentation that due to the evolution of technology, have been seen in the obligation to digitalize it for greater organization and agility of the processes that are carried out with regard to this.

The search systems are a feasible tool to find suitable information that we need to get in a given situation, making access to information to be agile and generating the desired results.

Based on the problem of that there are still processes in which the management of the information and documents are not conducted in a way that users or customers have the ease of being able to manage them or access them you put forward the proposal to the Pontificia Universidad Católica del Ecuador Sede Esmeraldas (PUCESE) develop a system that allows the search for documents of thesis or projects of degree digitized, so that the researchers, i.e. the students, teachers and the general public can perform queries to these Documents so that is available at the time they need it, and allows them to gain time in the conduct of its investigations.

For the realization of this research the inductive methods - which allowed a deductive collect data and aspects associated with the project. By which techniques were used as the interview, applied to the maintainer of the Library and the Department of the PUCESE Systems. It was also conducted survey of 144 people including students, professors and members of the CID (Center for Research and Development).

Thus it is considered that the project is feasible for the institution, because it provides ease of access that researchers need when arranged a consultation to such documents.

PRESENTACIÓN

Los sistemas informáticos, han contribuido de manera positiva al avance tecnológico y económico; solucionando diversas problemáticas que se presentan fundamentalmente en el campo empresarial, facilitando la gestión de los procesos administrativos y disminuyendo el manejo de documentación física, que en ocasiones presenta inconvenientes, en cuanto a organización y seguridad se refiere.

Actualmente la aplicación de los sistemas de búsqueda es de gran utilidad en distintos ambientes donde son necesarios para el orden y la administración, como empresas e instituciones acostumbradas a manejar cantidades considerables de documentación.

A través del análisis y recopilación de datos necesarios, se pudo constatar entre varios aspectos la existencia de diversas técnicas y algoritmos de búsqueda que permiten la recuperación de información, datos que se obtienen de la base de datos para posteriormente hacer uso pertinente de los mismos.

Durante el desarrollo del proyecto, se necesitó de información correspondiente a los documentos de tesis o proyectos de grado de la PUCESE, esta fue proporcionada por la institución, la cual facilitó una copia de la base de datos contenida en el Gestor de Base de Datos PostgreSQL. Lo que garantizó que el sistema pudiera ser desarrollado con información verídica y confiable.

Para determinar la problemática se aplicaron métodos y técnicas de investigación, por medio de las cuales se pudo realizar un diagnóstico que permitió constatar los inconvenientes que tienen los investigadores para realizar consultas a los documentos de tesis o proyectos de grado de la PUCESE, y los requisitos de búsqueda que precisan de ellos. Es por ello que la propuesta se considera adecuada, porque cubre una necesidad inminente de la comunidad investigadora.

ÍNDICE DE CONTENIDOS

TRABAJO DE TITULACIÓN DE GRADO ..... **ivi**

AUTORÍA ..... **ivii**

DEDICATORIA ..... **iv**

AGRADECIMIENTO ..... **v**

RESUMEN ..... **vi**

PRESENTACIÓN ..... **viii**

ÍNDICE DE CONTENIDOS ..... **ix**

CAPÍTULO I: MARCO TEÓRICO **1**

1.1. PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR SEDE ESMERALDAS (PUCESE) 1

1.1.1. Filosofía Institucional 1

1.1.2. Objetivos 2

1.1.3. Estructura Orgánica 2

1.1.4. Departamento de Biblioteca 3

1.1.4.1. Estructura 3

1.1.4.2. Manejo de Inventario 4

1.2. REPOSITORIOS DIGITALES 5

1.3. BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN 6

1.3.1. Definición 6

1.3.2. Problemática de la Recuperación de información (RI) 6

1.3.3. Técnicas de Búsqueda 8

1.4. ALGORITMOS DE BÚSQUEDA 9

1.4.1. Definición 9

1.4.2. Tipos 10

1.4.2.1. Búsqueda secuencial 10

1.4.2.2. Búsqueda Dicotómica (Binaria) 11
1.5. DESARROLLO DE SOFTWARE 14
1.5.1. Herramienta de Desarrollo 14
1.5.1.1. Xampp 14
1.5.1.2. PHP 14
1.5.1.3. JQuery 15
1.5.1.4. Sistema Gestor de Base de Datos PostgreSQL. 15
1.5.2. Metodologías de Desarrollo de Software 17
1.5.2.1. Programación Extrema XP 18
1.5.3. Modelos de Desarrollo de Software 20
1.5.3.2 Lenguaje Unificado de Modelado (UML) 20
1.5.3.2.1 Diagramas UML 21
1.5.4 Arquitecturas de Sistemas de Información 22
1.5.5 Fases metodológicas 23
CAPÍTULO II: DIAGNÓSTICO 25
2.1. ANTECEDENTES DIAGNÓSTICOS 25
2.1.1. Objetivos Diagnósticos 26
2.1.2. Variables Diagnósticas 26
2.1.3. Indicadores 26
2.2. MATRIZ DE RELACIÓN DIAGNÓSTICA 27
2.3. MECÁNICA OPERATIVA 30
2.3.1. Población 30
2.3.2. Muestra 30
2.3.3. Información Primaria 31
2.3.4. Información Secundaria 31
2.3. ANÁLISIS Y TABULACIÓN DE DATOS 32
2.5. F.O.D.A. 42

CAPÍTULO III: PROPUESTA 45
3.1 JUSTIFICACIÓN 45
3.2 OBJETIVOS 46
3.2.1 General 46
3.2.2 Específicos 46
3.3 SOLUCIÓN INFORMÁTICA. 47
3.3.1 Metodología de Sofware. 47
3.3.2 Diseño 47
3.3.2.3 Diagramas de Caso de Uso. 50
3.3.4 Modelado de Datos 58
3.3.4.1 Modelo Conceptual 58
3.3.4.2 Modelo Entidad Relación 59
3.3.4.3 Modelo Relacional 60
3.3.5 Plataformas de Desarrollo. 60
3.3.6 Módulos del Sistema 62
3.3.7 Manual de Usuario. 62
CAPÍTULO IV: IMPACTOS 68
4.1 ANÁLISIS DE IMPACTOS 68
4.1.1 Impacto Tecnológico 69
4.1.2 Impacto Económico 70
4.1.3 Impacto Educativo 71
4.1.4 Impacto General 72
CAPÍTULO V:CONCLUSIONES Y RECOMENDACIONES 73
5.1 CONCLUSIONES 73
5.2 RECOMENDACIONES 744
REFERENCIAS BIBLIOGRÁFÍCAS 77
TRABAJOS CITADOS 80

ANEXO 85

ÍNDICE DE TABLAS

Tabla 1: Matriz Diagnóstica 29
Tabla 2: F.O.D.A. 43
Tabla 3: Descripción de tablas en Postgresql 57
Tabla 4: Matriz de Impacto 68
Tabla 5: Análisis de Impacto Tecnológico 69
Tabla 6: Análisis de Impacto Económico 70
Tabla 7: Análisis de Impacto Educativo 71
Tabla 8: Análisis de Impacto General 72

ÍNDICE DE ILUSTRACIONES

Ilustración 1: Estructura organizacional de la PUCESE. 3
Ilustración 2: Sistema de Recuperación de Información (SRI) 7
Ilustración 3: Componentes de un Sistema PostgreSQL 16
Ilustración 4: Fases Metodológicas 24
Ilustración 5: Pregunta 1-Investigadores 32
Ilustración 6: Pregunta 2-Investigadores 33
Ilustración 7: Pregunta 3-Investigadores 34
Ilustración 8: Pregunta 4-Investigadores 35
Ilustración 9: Pregunta 5-Investigadores 36
Ilustración 10: Pregunta 6-Investigadores 37
Ilustración 11: Pregunta 7-Investigadores 38
Ilustración 12: Pregunta 8-Investigadores 39
Ilustración 13: Pregunta 9-Investigadores 40
Ilustración 14: Diagrama Funcional del Sistema 48
Ilustración 15: DFD - Nivel 0 49
Ilustración 16: DFD Nivel 1 49
Ilustración 17: DFD Nivel 2 - Investigadores 49
Ilustración 18: DFD Nivel 2 – Administrador 50
Ilustración 19: Diagrama de caso de uso. Primera etapa - ingreso al sistema 51

Ilustración 20: Diagrama de caso de uso. Segunda etapa – Búsqueda. 51
Ilustración 21: Diagrama de caso de uso. Búsqueda por categoría. 52
Ilustración 22: Diagrama de caso de uso. Búsqueda por contenido. 52
Ilustración 23: Diagrama de Caso de Uso - Vista del administrador..... 53
Ilustración 24: Modelo Entidad Relación 59
Ilustración 25: Modelo Relacional..... 60
Ilustración 26: Página Principal del Sistema 62
Ilustración 27: Opciones de búsqueda 63
Ilustración 28: Búsqueda por título..... 63
Ilustración 29: Búsqueda por Escuela..... 64
Ilustración 30: Búsqueda por año 64
Ilustración 31: Búsqueda por autor..... 65
Ilustración 32: Búsqueda por contenido 65
Ilustración 33: Inicio de sesión 66
Ilustración 34: Ingreso de Tesis 66
Ilustración 35: Ingreso de usuarios 67
Ilustración 36: Editar usuario..... 67

CAPÍTULO I:
MARCO TEÓRICO

1.1. PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR SEDE ESMERALDAS (PUCESE)

1.1.1. Filosofía Institucional

La Pontificia Universidad Católica del Ecuador Sede Esmeraldas (PUCESE) es una institución cuya fundación fue propuesta por la iglesia, por lo que inculca valores católicos, como son la solidaridad, armonía, justicia, conciencia y responsabilidad social, entre otros, los cuales apoyan a la formación de profesionales competentes.

La razón de ser de la PUCESE se refleja claramente en su misión, la cual establece que la institución pretende:

"Formar continua, personalizada e integralmente a seres humanos con sentido emprendedor social, ético, crítico y autocrítico, a la luz del evangelio, capaces de liderar y generar transformaciones en orden a una provincia solidaria, justa, pacífica y que respeta la biodiversidad, desarrollando propuestas científicas, innovadoras y sostenibles." (PUCESE, 2014)

La universidad tiene metas futuras que espera cumplir a largo plazo, esto consta en su visión la misma que dispone que:

"La PUCESE será una institución educativa en búsqueda permanente de la excelencia académica, con carreras acreditadas, apoyada en la estructura de trabajo por áreas de conocimiento; estrechamente vinculada a organizaciones de los sectores educativos, productivos, de salud y medioambientales de Esmeraldas, como provincia costera; participando en redes de investigación, intercambio y formación de estudiantes y

docentes con instituciones de educación superior nacionales e internacionales, a través de trabajo cooperativo en propuestas de transformación social." (PUCESE, 2014)

1.1.2. Objetivos

Dentro de la planeación correspondiente a los años 2012 - 2016, la PUCESE se plantea los siguientes objetivos estratégicos. (PUCESE, 2014)

- Fortalecer la estructura y los procesos administrativos y académicos.
- Elevar el nivel académico del ingreso y egreso, y la permanencia en la universidad.
- Impulsar la investigación y la innovación.
- Mejorar los canales de comunicación.
- Implementar las tics en todos los niveles de la gestión académica y administrativa.
- Ser una universidad emprendedora con capacidad para generar proyectos empresariales y procesos de liderazgo orientados a los planes de desarrollo provincial y nacional.
- Fortalecer los vínculos con la comunidad que generen transformaciones efectivas en la provincia.

1.1.3. Estructura Orgánica

El máximo ente de la PUCESE es el Consejo Directivo que al 2016 está conformado por diferentes miembros: Mons. Eugenio Arellano Fernández (VICARIO APOSTÓLICO), Lic. Aitor Urbina García de Vicuña (PRO-RECTOR), Mgt. María de los Lirios Bernabé (DIRECTORA ACADÉMICA), Lic. Lilian Vera Pinargote (DIRECTORA DE ESTUDIANTES), Ing. Marjorie Segovia Aguilar (DIRECTORA ADMINISTRATIVA), Mgt. María Elena Vélez Loor, (DIRECTORA FINANCIERA), Ing. Marítza Demera Mejía (SECRETARIA GENERAL), Ing. Orlín Álava (REPRESENTANTE ASOCIACIÓN DE PROFESORES), Soc. Ecuador Olaya (REPRESENTANTE ASOCIACIÓN DE TRABAJADORES), Srta. Katiuska Arboleda Quinde (REPRESENTANTE ASOCIACIÓN DE ESTUDIANTES), Dr.

Walter Mosquera (REPRESENTANTE DE LOS DOCENTES DESIGNADO POR EL PRO-RECTOR), Ing. Boris Estupiñán Ortíz (REPRESENTANTE DE LA ASOCIACIÓN DE GRADUADOS), todas estas dignidades cumplen con un rol fundamental dentro de la institución. La estructura orgánica comprende jerarquías que se muestran en la ilustración 1.

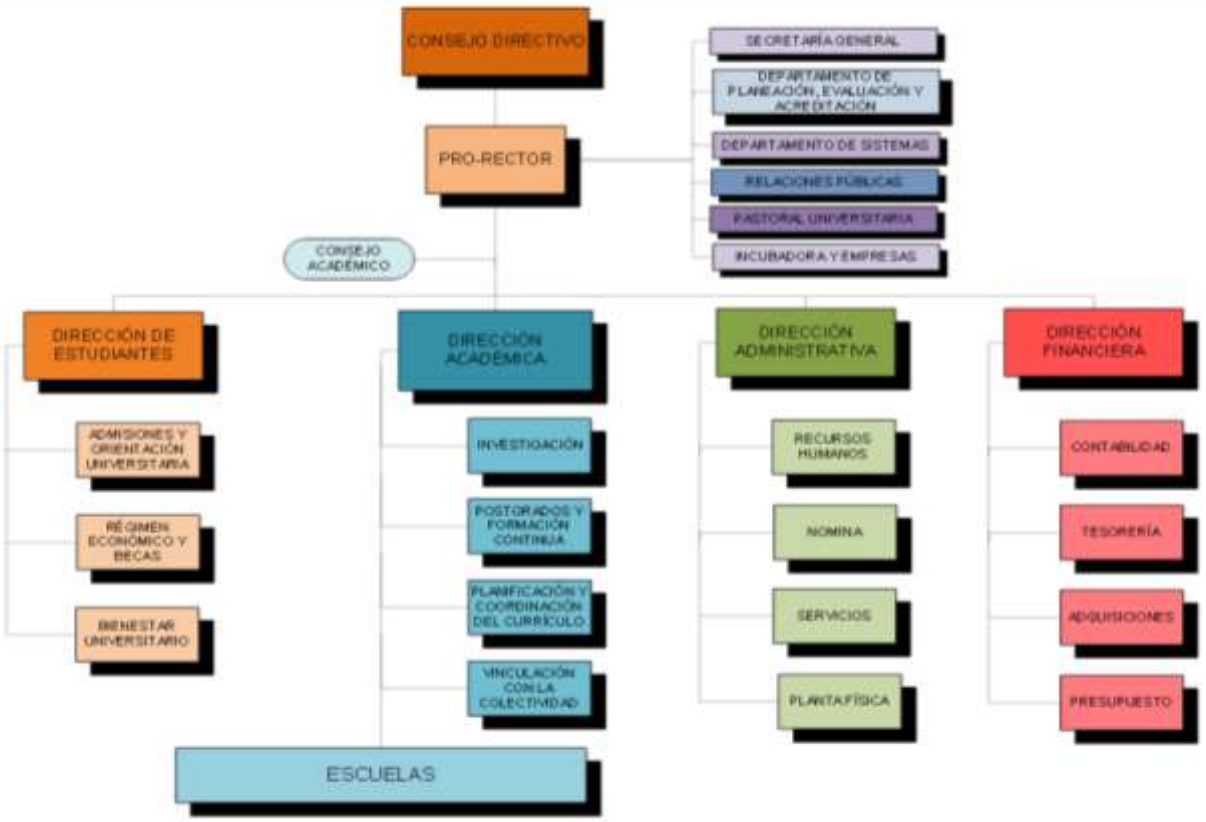


Ilustración 1: Estructura organizacional de la PUCESE.

Fuente: (PUCESE, 2016)

1.1.4. Departamento de Biblioteca

1.1.4.1. Estructura

Para una institución de educación superior es crucial el aspecto investigativo, por lo que la Biblioteca de la PUCESE brinda sus servicios a estudiantes, docentes y público en

general, realizando préstamos de tomos bibliográficos, contribuyendo al desarrollo de investigaciones y proyectos.

La Biblioteca realiza un análisis de adquisiciones, es decir, las posibilidades de ingreso de ejemplares al departamento, dependiendo de las necesidades de información y actualización de los tomos, los cuales se obtienen por donaciones, por compras directas o mediante convenios.

Entre los servicios que ofrece está el préstamo de libros requerido por los investigadores, respondiendo a peticiones de temas, autores, especialidades, entre otros.

La Biblioteca PUCESE además presta sus servicios a través de su Biblioteca Virtual, donde los investigadores pueden acceder desde la página principal de la institución y realizar consultas de libros alojados en biblioteca.

1.1.4.2. Manejo de Inventario

Para llevar el adecuado control del material bibliográfico que la biblioteca administra, es necesario llevar un inventario que arrastra un proceso minucioso en cuanto al registro de adquisiciones y préstamos.

El inventario es quizás una de las actividades más minuciosas y pesadas que deben cumplir los bibliotecarios; sin embargo, resulta imprescindible para el control de los materiales del acervo de la biblioteca. De su omisión se deriva un deterioro progresivo en la efectividad del catálogo topográfico como herramienta de trabajo y, por lo tanto, de la calidad del servicio. (Ministerio del poder popular de cultura de Venezuela, 2008)

Los objetivos principales que se manejan para la realización de un inventario son conocer, identificar, ajustar, determinar, facilitar y posibilitar la consulta inmediata percatándose de la existencia real del material de la biblioteca.

La frecuencia con que se realiza el inventario depende del tipo, tamaño y movimiento de la biblioteca. Cuando se revisa frecuentemente las estanterías y si las pérdidas son

pocas no es necesario realizar el inventario regularmente, por el contrario si la biblioteca es pequeña y tiene mucho movimiento con muchas pérdidas será preciso hacer un inventario lo más frecuente posible.

Para una correcta realización del inventario se debe suspender el servicio de préstamo externo con anterioridad, con el fin de recuperar los libros prestados y revisar cuidadosamente en otros lugares de la biblioteca que no se quede ningún material sin inventariar y así garantizar que todas las obras estén devueltas y localizadas en las estanterías. (Dorado, 2012)

1.2. REPOSITORIOS DIGITALES

Los sistemas donde se almacena todo tipo de información académica como revistas, tesis, proyectos, artículos, videos, imágenes y demás disponibles en la web, se denominan repositorios digitales. Las universidades u organizaciones disponen de los servicios de los repositorios para difundir esta información y facilitar el acceso a las mismas a la comunidad investigativa, en otros casos los archivos están protegidos y la autenticación del usuario para acceder a ellos es necesaria. Dichos archivos son clasificados por categorías, temáticas, contenidos científicos y de acuerdo al año en que fueron desarrollados. (Universidad Interamericana para el Desarrollo (UNID), 2014)

Un repositorio cuenta con varias características entre ellas está que pueden almacenar diferentes tipos de archivos, cada objeto contiene metadatos, ofrece un sistema de gestión y validación de las publicaciones, cuenta con un sistema de búsqueda que agiliza la localización de los objetos, incrementa la difusión y visualización de contenidos. (Universidad Interamericana para el Desarrollo (UNID), 2014)

En la actualidad lo habitual es almacenar toda la información de manera digital con el software pertinente y este es necesario que, entre otras cosas, ofrezca personalización e interoperatividad.

Una universidad puede crear un repositorio de investigaciones científicas en Internet. Para esto, comienza a publicar documentos en un portal de para que cualquier persona tenga la posibilidad de acceder a la información. Al cabo de unos meses, el portal puede

albergar cada vez más documentos y se convertirá en un gran repositorio del saber. (Pérez, 2014)

1.3. BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN

1.3.1. Definición

El proceso de recuperación se lleva a cabo mediante consultas a la base de datos donde se almacena la información estructurada. Es necesario tener en cuenta los elementos clave que permiten hacer la búsqueda, determinando un mayor grado de pertinencia y precisión, como son: los índices, palabras clave y los fenómenos que se pueden dar en el proceso como son el ruido y silencio documental. Uno de los problemas que surgen en la búsqueda de información es si lo que recuperamos es "mucho o poco" es decir, dependiendo del tipo de búsqueda se pueden recuperar multitud de documentos o simplemente un número muy reducido. A este fenómeno se denomina Silencio o Ruido documental. (Pinto, 2011)

- Silencio documental: Son aquellos documentos almacenados en la base de datos pero que no han sido recuperados, debido a que la estrategia de búsqueda ha sido demasiado específica o que las palabras clave utilizadas no son las adecuadas para definir la búsqueda.
- Ruido documental: Son aquellos documentos recuperados por el sistema pero que no son relevantes. Esto suele ocurrir cuando la estrategia de búsqueda se ha definido demasiado genérica. (Pinto, 2011)

1.3.2. Problemática de la Recuperación de información (RI)

Si se analiza la problemática de la RI desde un alto nivel de abstracción se puede establecer que: (Tolasa & Bordignon, 2008)

- Existe una colección de documentos que contienen información de interés(sobre uno o varios temas)

- Existen usuarios con necesidades de información, quienes las plantean al Sistema de Recuperación de Información (SRI) en forma de una consulta (en inglés, query. En adelante, ambas palabras se utilizarán indistintamente)

- Como respuesta, el sistema retorna – de forma ideal – referencias a documentos "relevantes", es decir aquellos que satisfacen la necesidad expresada, generalmente en forma de una lista.

La respuesta "ideal" de un SRI como se refleja en la ilustración 2, está formada solamente por documentos relevantes a la consulta, pero en la práctica esta no es aún alcanzable.

Esto se debe a que entre otros motivos existe el problema de compatibilizar la expresión de la necesidad de información y el lenguaje y de los documentos. A demás, hay una carga de subjetividad subyacente y depende de los usuarios. Entonces, el SRI recupera la mayor cantidad posible de documentos relevantes, minimizando la cantidad de documentos no relevantes (ruido) en la respuesta. En términos de eficiencia, se plantea la idea de precisión de la respuesta, es decir, cuando más documentos relevantes contenga el conjunto solución (para una consulta dada), más preciso será. (Tolasa & Bordignon, 2008)

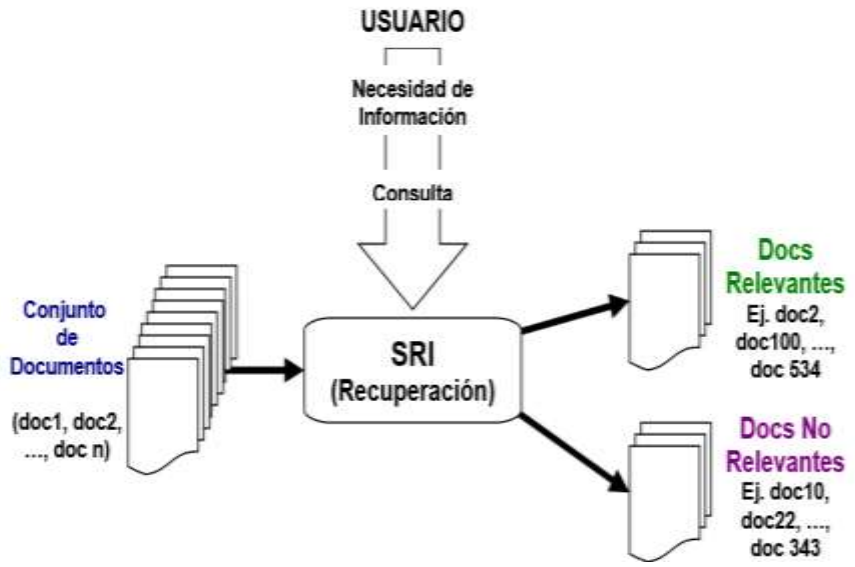


Ilustración 2: Sistema de Recuperación de Información (SRI)

Fuente: (Tolasa & Bordignon, 2008)

1.3.3. Técnicas de Búsqueda

Existen diferentes técnicas que ayudan a la búsqueda efectiva de aspectos requeridos en distintas instancias, de entre las cuales tenemos las siguientes: (Pinto, 2011)

- Sistemas de recuperación de lógica difusa: Esta técnica permite establecer consultas con frases normales, de forma que la máquina al realizar la búsqueda elimina signos de puntuación, artículos, conjunciones, plurales, tiempos verbales, palabras comunes (que suelen aparecer en todos los documentos), dejando sólo aquellas palabras que el sistema considera relevantes. La recuperación se basa en proposiciones lógicas con valores de verdadero y falso, teniendo en cuenta la localización de la palabra en el documento.

- Técnicas de ponderación de términos: Es común que unos criterios en la búsqueda tenga más valor que otros, por tanto la ponderación pretende darle un valor adecuado a la búsqueda dependiendo de los intereses del usuario. Los documentos recuperados se encuentran en función del valor obtenido en la ponderación. El valor depende de los términos pertinentes que contenga el documento y la frecuencia con que se repita. De forma que, el documento más pertinente de búsqueda sería aquel que tenga representado todos los términos de búsqueda y además el que más valor tenga repetidos más veces, independientemente de donde se localice en el documento.

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

- Algoritmos utilizados para realizar la categorización (clúster):

  - ✓ Algoritmo K-means
  - ✓ COBWEB
  - ✓ Algoritmo EM

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

- Técnicas de stemming: Morfológicamente las palabras están estructuradas en prefijos, sufijos y la raíz. La técnica de Stemming lo que pretende es eliminar las posibles confusiones semánticas que se puedan dar en la búsqueda de un concepto, para ello trunca la palabra y busca solo por la raíz.

- Algoritmos utilizados para desechar prefijos y sufijos:

  - ✓ Paice/Husk
  - ✓ S-stemmer / n-gramas
  - ✓ Técnicas lingüísticas (Pinto, 2011)

Realizando un análisis de todas las técnicas de búsqueda descritas anteriormente, se escogió la técnica de recuperación de lógica difusa para ser aplicada en el sistema, debido a que esta técnica permite que los usuarios puedan ingresar palabras claves que consideren referenciales para obtener resultados óptimos, sin tener que preocuparse por escribir la palabra o frase en mayúsculas, minúsculas, plurales, acentos, etc. Lo que proporciona cierta confiabilidad en que los resultados no se verán afectados por este tipo de situación en particular.

1.4. ALGORITMOS DE BÚSQUEDA

1.4.1. Definición

Los algoritmos son similares a las ecuaciones algebraicas: se obtiene una respuesta diferente para cada valor de "x", o variable de tu plug in. Sin embargo, un algoritmo es una ecuación mucho más compleja, que trata con más variables de lo que puedes

contar. En la tecnología de los motores de búsqueda, los algoritmos se utilizan para ordenar a través de los miles de millones de documentos y páginas web en el ciberespacio para ofrecer al usuario un resultado específico. (Storti, D'Elía, Paz, Lizandro, & Martín, 2014)

1.4.2. Tipos

Con mucha frecuencia los programadores trabajan con grandes cantidades de datos almacenados en arrays y registros, y por ello será necesario determinar si un array contiene un valor que coincida con un cierto valor clave. El proceso de encontrar un elemento específico de un array se denomina búsqueda. En esta sección se examinarán dos técnicas de búsqueda: búsqueda lineal o secuencial, la técnica más sencilla, y búsqueda binaria o dicotómica, la técnica más eficiente. (Storti, D'Elía, Paz, Lizandro, & Martín, 2014)

1.4.2.1. Búsqueda secuencial

Consiste en recorrer y examinar cada uno de los elementos del array hasta encontrar el o los elementos buscados, o hasta que se han mirado todos los elementos del array. (Días, 2006)

Este es el método de búsqueda más lento, pero si nuestra información se encuentra completamente desordenada es el único que podrá ayudar a encontrar el dato que se busca. El siguiente algoritmo muestra un esquema de implementación del algoritmo de búsqueda secuencial:

1. Desde 0 hasta que el incremento sea mayor o igual al valor del array repetir.
2. Si el elemento del array es igual al elemento a buscar.
3. Se ha obtenido la solución.

Este algoritmo se puede optimizar cuando el array está ordenado, en cuyo caso la condición de salida cambiaría a:

1. Desde 0 repetir mientras le valor del array sea menor o igual al elemento.

Cuando sólo interesa conocer la primera ocurrencia del elemento en el array.

2. Desde 0 hasta que el incremento sea mayor o igual al valor del array repetir
3. Si el elemento del array es igual al elemento a buscar.

En este último caso, cuando sólo interesa la primera posición, se puede utilizar un centinela, esto es, dar a la posición siguiente al último elemento de array el valor del elemento, para estar seguro de que se encuentra el elemento, y no tener que comprobar a cada paso si seguimos buscando dentro de los límites del array: se recorre el array.

1. Si se encuentra con que uno de los valores del array es igual al elemento a buscar.
2. Se obtiene la solución.

Si al acabar de recorrer el array este termina con el bucle, i vale N esto indica que no se encontró el elemento. (Días, 2006)

1.4.2.2. Búsqueda Dicotómica (Binaria)

Si los elementos sobre los que se realiza la búsqueda están ordenados, entonces podemos utilizar un algoritmo de búsqueda mucho más rápido que el secuencial, la búsqueda binaria. El algoritmo consiste en reducir paulatinamente el ámbito de búsqueda a la mitad de los elementos, basándose en comparar el elemento a buscar con el elemento que se encuentra en la mitad del intervalo y en base a esta comparación:

- Si el elemento buscado es menor que el elemento medio, entonces sabemos que el elemento está en la mitad inferior de la tabla.
- Si es mayor es porque el elemento está en la mitad superior.
- Si es igual se finaliza con éxito la búsqueda ya que se ha encontrado el elemento.

Se puede aplicar tanto a datos en listas lineales (Vectores, Matrices, etc.) como en árboles binarios de búsqueda. Los prerrequisitos principales para la búsqueda binaria son:

- La lista debe estar ordenada en un orden específico de acuerdo al valor de la llave.
- Debe conocerse el número de registros.

La búsqueda binaria consiste en dividir el array por su elemento medio en dos subarrays más pequeños, y comparar el elemento con el del centro. Si coinciden, la búsqueda se termina. Si el elemento es menor, debe estar (si está) en el primer subarray, y si es mayor está en el segundo. Por ejemplo, para buscar el elemento 3 en el array {1,2,3,4,5,6,7,8,9} se realizarían los siguientes pasos:

Se toma el elemento central y se divide el array en dos:

{1,2,3,4}-5-{6,7,8,9}

Como el elemento buscado (3) es menor que el central (5), debe estar en el primer subarray: {1,2,3,4}

Se vuelve a dividir el array en dos:

{1}-2-{3,4}

Como el elemento buscado es mayor que el central, debe estar en el segundo subarray: {3,4}
Se vuelve a dividir en dos:

{}-3-{4}

Como el elemento buscado coincide con el central, lo hemos encontrado.

Si al final de la búsqueda todavía no lo hemos encontrado, y el subarray a dividir está vacío {}, el elemento no se encuentra en el array. La implementación sería:

int desde, hasta, medio, elemento, posición; donde desde y hasta indican los límites del array que se está mirando.

1. Se le da valor a elemento. intarray[N];

2. Desde 0 hasta el valor del array menos 1 repetir, mientras el incremento sea menor o igual al valor del array.
   Si el array sólo tiene un elemento
3. Entonces
4. Se ha encontrado la solución
5. Se le da el valor
6. Caso contrario
7. No se encuentra en el array

En otro caso:

8. Se divide el array en dos. (desde+hasta)/2
9. Se escoge el elemento central
10. Si el elemento a buscar coincide con el central
11. Entonces se ha encontrado la solución
12. Caso contrario si es menor
13. Se elige el array de la izquierda.
14. Caso contrario si es mayor
15. Se elige el array de la derecha. (Días, 2006)

El tipo de algoritmo de búsqueda que se aplicó en el proyecto es el de búsqueda secuencial. Teniendo en cuenta que la información que contiene la base de datos no está lo suficientemente ordenada, este algoritmo fue el único que permitió que se realicen las búsquedas al recorrer cada uno de los datos de la base, hasta encontrar y destacar las coincidencias, para que de esta manera puedan ser mostrados al usuario los resultados solicitados.

1.5. DESARROLLO DE SOFTWARE

1.5.1. Herramienta de Desarrollo

1.5.1.1. Xampp

Xampp es un servidor independiente en base a software libre, con el cual podemos disponer de un servidor propio o simplemente usarlo para hacer pruebas de nuestras páginas web, bases de datos, para desarrollar aplicaciones en php, con conexión a base de datos sql (LAMPP= Linux + Apache + MySQL + PHP + Perl)

El programa está liberado bajo la licencia GNU y actúa como un servidor web libre, fácil de usar y capaz de interpretar páginas dinámicas. Actualmente Xampp está disponible para GNU/Linux, Microsoft Windows, Solaris y MacOS X.

Xampp solamente requiere descargar y ejecutar un archivo de tipo zip, tar o exe, con unas pequeñas configuraciones en alguno de sus componentes que el servidor Web necesitará. Xampp se actualiza regularmente para incorporar las últimas versiones de Apache/MySQL/PHP y Perl. También incluye otros módulos como OpenSSL y phpMyAdmin. Para instalar Xampp se requiere solamente una pequeña fracción del tiempo necesario para descargar y configurar los programas por separado. (UM, 2015)

1.5.1.2. PHP

Según Gonzales (2016), PHP es un lenguaje de código abierto muy popular, adecuado para desarrollo web y que puede ser incrustado en HTML. Es popular porque un gran número de páginas y portales web están creadas con PHP. Código abierto significa que es de uso libre y gratuito para todos los programadores que quieran usarlo. Incrustado en HTML significa que en un mismo archivo vamos a poder combinar código PHP con código HTML, siguiendo unas reglas.

PHP se utiliza para generar páginas web dinámicas. Recordar que llamamos página estática a aquella cuyos contenidos permanecen siempre igual, mientras que llamamos

páginas dinámicas a aquellas cuyo contenido no es el mismo siempre. Por ejemplo, los contenidos pueden cambiar en base a los cambios que haya en una base de datos, de búsquedas o aportaciones de los usuarios, etc.

1.5.1.3. JQuery

Para Duarte (2013), jQuery es una librería JavaScript open-source, que funciona en múltiples navegadores, y que es compatible con CSS3. Su objetivo principal es hacer la programación "scripting" mucho más fácil y rápida del lado del cliente. Con jQuery se pueden producir páginas dinámicas así como animaciones parecidas a Flash en relativamente corto tiempo.

Ventajas

La ventaja principal de jQuery es que es mucho más fácil que sus competidores. Se puede agregar plugins fácilmente, traduciéndose esto en un ahorro substancial de tiempo y esfuerzo. De hecho, una de las principales razones por la cual crearon jQuery fue para ganar tiempo (en el mundo de desarrollo web, el tiempo importa mucho).

La licencia open source de jQuery permite que la librería siempre cuente con soporte constante y rápido, publicándose y actualizaciones de manera constante. La comunidad jQuery es activa y sumamente trabajadora. (Duarte, 2013)

1.5.1.4. Sistema Gestor de Base de Datos PostgreSQL.

PostgreSQL es un sistema de gestión de bases de datos objeto-relacional, y con el código fuente disponible libremente. Es el sistema de gestión de bases de datos de código abierto más potente del mercado y en sus últimas versiones no tiene nada que envidiarle a otras bases de datos comerciales.

PostgreSQL utiliza un modelo cliente/servidor y usa multiprocesos en vez de multihilos para garantizar la estabilidad del sistema. Un fallo en uno de los procesos no afectará el resto y el sistema continuará funcionando. (PostgreSQL, 2010)

A continuación en la ilustración 3, se muestra de manera general los componentes más importantes en un sistema PostgreSQL.

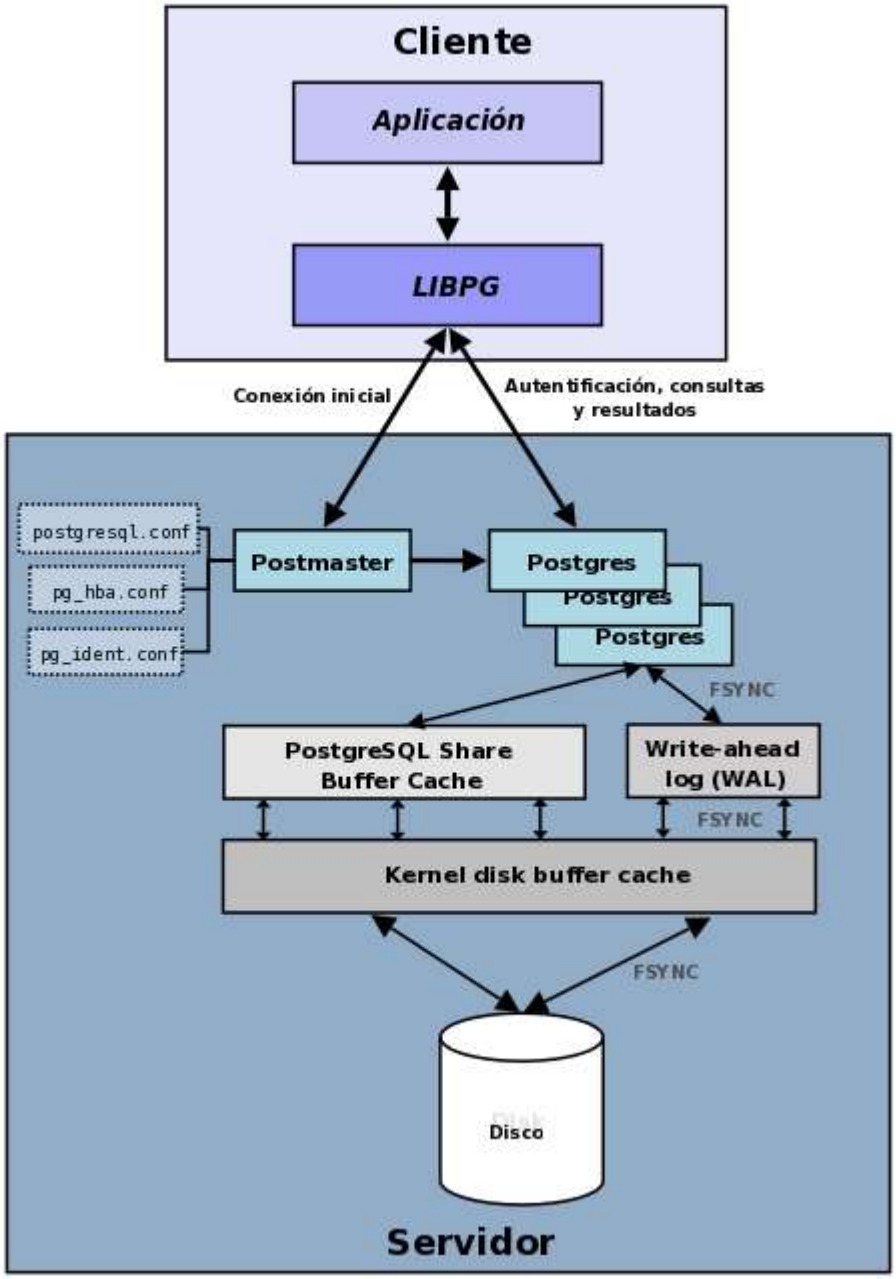


Ilustración 3: Componentes de un Sistema PostgreSQL

Fuente: (PostgreSQL, 2010)

- **Aplicación cliente:** Esta es la aplicación cliente que utiliza PostgreSQL como administrador de bases de datos. La conexión puede ocurrir via TCP/IP ó sockets locales.

- **Demonio postmaster:** Este es el proceso principal de PostgreSQL. Es el encargado de escuchar por un puerto/socket por conexiones entrantes de clientes. Tambien es el encargado de crear los procesos hijos que se encargaran de autentificar estas peticiones, gestionar las consultas y mandar los resultados a las aplicaciones clientes.
- **Ficheros de configuracion:** Los 3 ficheros principales de configuración utilizados por PostgreSQL, postgresql.conf, pg_hba.conf y pg_ident.conf.
- **Procesos hijos postgres:** Procesos hijos que se encargan de autentificar a los clientes, de gestionar las consultas y mandar los resultados a las aplicaciones clientes.
- **PostgreSQL share buffer cache:** Memoria compartida usada por POstgreSQL para almacenar datos en caché.
- **Write-Ahead Log (WAL):** Componente del sistema encargado de asegurar la integridad de los datos (recuperación de tipo REDO)
- **Kernel disk buffer cache:** Caché de disco del sistema operativo.
- **Disco:** Disco físico donde se almacenan los datos y toda la información necesaria para que PostgreSQL funcione. (PostgreSQL, 2010)

1.5.2. Metodologías de Desarrollo de Software

Las metodologías de desarrollo de software son un conjunto de procedimientos, técnicas y ayudas a la documentación para el desarrollo de productos software.

Es como un libro de recetas de cocina, en el que se van indicando paso a paso todas las actividades a realizar para lograr el producto informático deseado, indicando además qué personas deben participar en el desarrollo de las actividades y qué papel deben de tener. Además detallan la información que se debe producir como resultado de una actividad y la información necesaria para comenzarla.

Actualmente es imprescindible considerar los riesgos, aunque habitualmente las empresas, no han sido concienciadas de los riesgos inherentes al procesamiento de la información mediante ordenadores, a lo que han contribuido, a veces, los propios responsables de informática, que no han sabido explicar con la suficiente claridad las

consecuencias de una política de seguridad insuficiente o incluso inexistente. Por otro lado, debido a una cierta deformación profesional en la aplicación de los criterios de coste/beneficio, el directivo desconocedor de la informática no acostumbra a autorizar inversiones que no lleven implícito un beneficio demostrable, tangible y mensurable. (Universisas de Mursia, 2011)

1.5.2.1. Programación Extrema XP

Es una metodología ágil centrada en potenciar las relaciones interpersonales como clave para el éxito en desarrollo de software, promoviendo el trabajo en equipo, preocupándose por el aprendizaje de los desarrolladores, y propiciando un buen clima de trabajo. XP se basa en realimentación continua entre el cliente y el equipo de desarrollo, comunicación fluida entre todos los participantes, simplicidad en las soluciones implementadas y coraje para enfrentar los cambios. XP se define como especialmente adecuada para proyectos con requisitos imprecisos y muy cambiantes, y donde existe un alto riesgo técnico. (Universidad Uión Bolivariana, 2014)

A continuación se presentan doce prácticas para llevar a cabo esta metodología en el desarrollo de un software: (Universidad Uión Bolivariana, 2014)

- ✓ **Equipo completo:** Forman parte del equipo todas las personas que tienen algo que ver con el proyecto, incluido el cliente y el responsable del proyecto.
- ✓ **Planificación:** Se hacen las historias de usuario y se planifica en qué orden se van a hacer y las mini-versiones. La planificación se revisa continuamente.
- ✓ **Test del cliente:** El cliente, con la ayuda de los desarrolladores, propone sus propias pruebas para validar las mini-versiones.
- ✓ **Versiones pequeñas:** Las mini-versiones deben ser lo suficientemente pequeñas como para poder hacer una cada pocas semanas. Deben ser versiones que ofrezcan algo útil al usuario final y no trozos de código que no pueda ver funcionando.
- ✓ **Diseño simple:** Hacer siempre lo mínimo imprescindible de la forma más sencilla posible. Mantener siempre sencillo el código.

- ✓ **Pareja de programadores:** Los programadores trabajan por parejas (dos delante del mismo ordenador) y se intercambian las parejas con frecuencia (un cambio diario).
- ✓ **Desarrollo guiado por las pruebas automáticas:** Se deben realizar programas de prueba automática y deben ejecutarse con mucha frecuencia. Cuantas más pruebas se hagan, mejor.
- ✓ **Integración continua:** Deben tenerse siempre un ejecutable del proyecto que funcione y en cuanto se tenga una nueva pequeña funcionalidad, debe recompilarse y probarse. Es un error mantener una versión congelada dos meses mientras se hacen mejoras y luego integrarlas todas de golpe. Cuando falle algo, no se sabe qué es lo que falla de todo lo que hemos metido.
- ✓ **El código es de todos:** Cualquiera puede y debe tocar y conocer cualquier parte del código. Para eso se hacen las pruebas automáticas.
- ✓ **Normas de codificación:** Debe haber un estilo común de codificación (no importa cuál), de forma que parezca que ha sido realizado por una única persona.
- ✓ **Metáforas:** Hay que buscar unas frases o nombres que definan cómo funcionan las distintas partes del programa, de forma que sólo con los nombres se pueda uno hacer una idea de qué es lo que hace cada parte del programa. Un ejemplo claro es el "recolector de basura" de java. Ayuda a que todos los programadores (y el cliente) sepan de qué estamos hablando y que no haya mal entendidos.
- ✓ **Ritmo sostenible:** Se debe trabajar a un ritmo que se pueda mantener indefinidamente. Esto quiere decir que no debe haber días muertos en que no se sabe qué hacer y que no se deben hacer un exceso de horas otros días. Al tener claro semana a semana lo que debe hacerse, hay que trabajar duro en ello para conseguir el objetivo cercano de terminar una historia de usuario o mini-versión.

1.5.3. Modelos de Desarrollo de Software

1.5.3.1 Diagramas de Flujo de Datos (DFD)

Un diagrama de flujo de datos es una técnica muy apropiada para reflejar de una forma clara y precisa los procesos que conforman el sistema de información. Permite representar gráficamente los límites del sistema y la lógica de los procesos, estableciendo qué funciones hay que desarrollar. Además, muestra el flujo o movimiento de los datos a través del sistema y sus transformaciones como resultado de la ejecución de los procesos.

Esta técnica consiste en la descomposición sucesiva de los procesos, desde un nivel general, hasta llegar al nivel de detalle necesario para reflejar toda la semántica que debe soportar el sistema en estudio.

El objetivo del diagrama de flujo de datos es la obtención de un modelo lógico de procesos que represente el sistema, con independencia de las restricciones físicas del entorno. Así se facilita su comprensión por los usuarios y los miembros del equipo de desarrollo.

El sistema se divide en distintos niveles de detalle, con el objetivo de:

- Simplificar la complejidad del sistema, representando los diferentes procesos de que consta.
- Facilitar el mantenimiento del sistema. (Cillero, 2009)

1.5.3.2 Lenguaje Unificado de Modelado (UML)

UML no es un método de desarrollo. No va a decir cómo pasar del análisis al diseño y de este al código. No son una serie de pasos que llevan a producir código a partir de unas especificaciones.

UML al no ser un método de desarrollo es independiente del ciclo de desarrollo que vayas a seguir, puede encajar en un tradicional ciclo en cascada, o en un evolutivo ciclo en espiral o incluso en los métodos ágiles de desarrollo. (Garcia, 2005)

El UML está compuesto por diversos elementos gráficos que se combinan para conformar diagramas. Debido a que el UML es un lenguaje, cuenta con reglas para combinar tales elementos.

Shirley (2012), describe que la finalidad de los diagramas es presentar un sistema desde diversas perspectivas, a las cuales se les conoce como modelo. Cabe recalcar que un modelo es una representación simplificada de la realidad; el modelo UML describe lo que supuestamente hará un sistema, pero no dice cómo implementar dicho sistema.

1.5.3.2.1 Diagramas UML

- **Diagramas de Casos de Uso**

Un diagrama de caso de uso describe lo que hace un sistema desde el punto de vista de un observador externo, debido a esto, un diagrama de este tipo generalmente es de los más sencillos de interpretar en UML, ya que su razón de ser se concentra en un ¿Qué hace? el sistema, a diferencia de otros diagramas UML que intentan dar respuesta a un ¿Cómo logra su comportamiento el sistema?.

Un caso de uso está muy relacionado con lo que pudiera ser considerado un escenario en el sistema, esto es, lo que ocurre cuando alguien interactúa con el sistema: "Acude un mesero a colocar la orden, la orden es tomada por el cocinero, y posteriormente se abona a la cuenta del cliente el cargo". (Osmosis Latina, 2005)

Este diagrama es empleado con más frecuencia en alguna de las siguientes etapas: (Osmosis Latina, 2005)

- Determinación de Requerimientos: Por lo general nuevos requerimientos de sistema generan nuevos usos-casos, conforme es analizado y diseñado el sistema.

- Comunicación con el Cliente: Debido a la sencillez de este tipo de diagramas, son fáciles de emplear para comunicarse con el cliente final del proyecto.

- Generación de pruebas de Sistemas: A través de los diagramas uso-caso se pueden generar una serie de pruebas de sistema.

- **Diagrama de Clases**

Los diagramas de clases de UML pueden utilizarse para una gran variedad de propósitos:

- Para proporcionar una descripción de los tipos que se utilizan en un sistema y se pasan entre sus componentes que no tenga nada que ver con su implementación.

Por ejemplo, en el código .NET, el tipo Pedido de menú podría implementarse en la capa del negocio; en XML, en las interfaces entre los componentes; en SQL, en la base de datos, y en HTML, en la interfaz de usuario. Aunque estas implementaciones tengan un nivel de detalle diferente, la relación entre el tipo Pedido de menú y otros tipos, como Menú y Pago, es siempre la misma. El diagrama de clases de UML permite analizar estas relaciones con independencia de las implementaciones.

- Para clarificar el glosario de términos que se utiliza en la comunicación entre la aplicación y los usuarios y en las descripciones de las necesidades de los usuarios. (Microsoft, 2007)

1.5.4 Arquitecturas de Sistemas de Información

Una arquitectura es un diseño estructural integrado de un sistema, sus elementos y definiciones dependen de los requerimientos proporcionados. El concepto de "arquitectura" es ampliamente usado en el contexto de la construcción de computadoras.

Cuando se aplica a los sistemas de información, una arquitectura es un plano abstracto que incluye los diseños de procesos de un sistema, basado en principios de diseño y dentro de un marco metodológico. (Anzaldo, 2005)

1.5.5 Fases metodológicas

La fase metodológica es una fase de diseño, en la que la idea toma forma. En esta fase se dibuja el "traje" que se le ha confeccionado al estudio a partir de una idea original. Sin una conceptualización adecuada del problema de investigación, resulta muy difícil poder concretar las partes que forman parte del diseño: (Pedraz, 2003)

1. Elección del diseño de investigación: ¿Qué diseño se adapta mejor al objeto del estudio? ¿Describir la realidad o ponerla a prueba? ¿Qué metodología permitirá encontrar unos resultados más ricos y que se ajusten más al tema de investigación?

2. Definición de los sujetos del estudio: ¿Quién es la población de estudio? ¿Cómo muestrearla? ¿Quiénes deben resultar excluidos de la investigación?

3. Descripción de las variables de la investigación: Acercamiento conceptual y operativo al objeto de la investigación. ¿Qué se entiende por cada una de las partes del objeto de estudio? ¿Cómo se van a medirlas?

4. Elección de las herramientas de recogida y análisis de los datos: ¿Desde qué perspectiva se está abordando la investigación? ¿Qué herramientas son las más adecuadas para recoger los datos de la investigación? Este es el momento en el que se decide si resulta más conveniente pasar una encuesta o "hacer un grupo de discusión", si se debe construir una escala o realizar entrevistas en profundidad. Y debemos explicar además cómo vamos a analizar los datos que recojamos en nuestro estudio. (Pedraz, 2003)

A continuación la ilustración 4 presenta las fases metodológicas, el proceso lógico que se sigue para la elaboración de un proyecto exitoso.

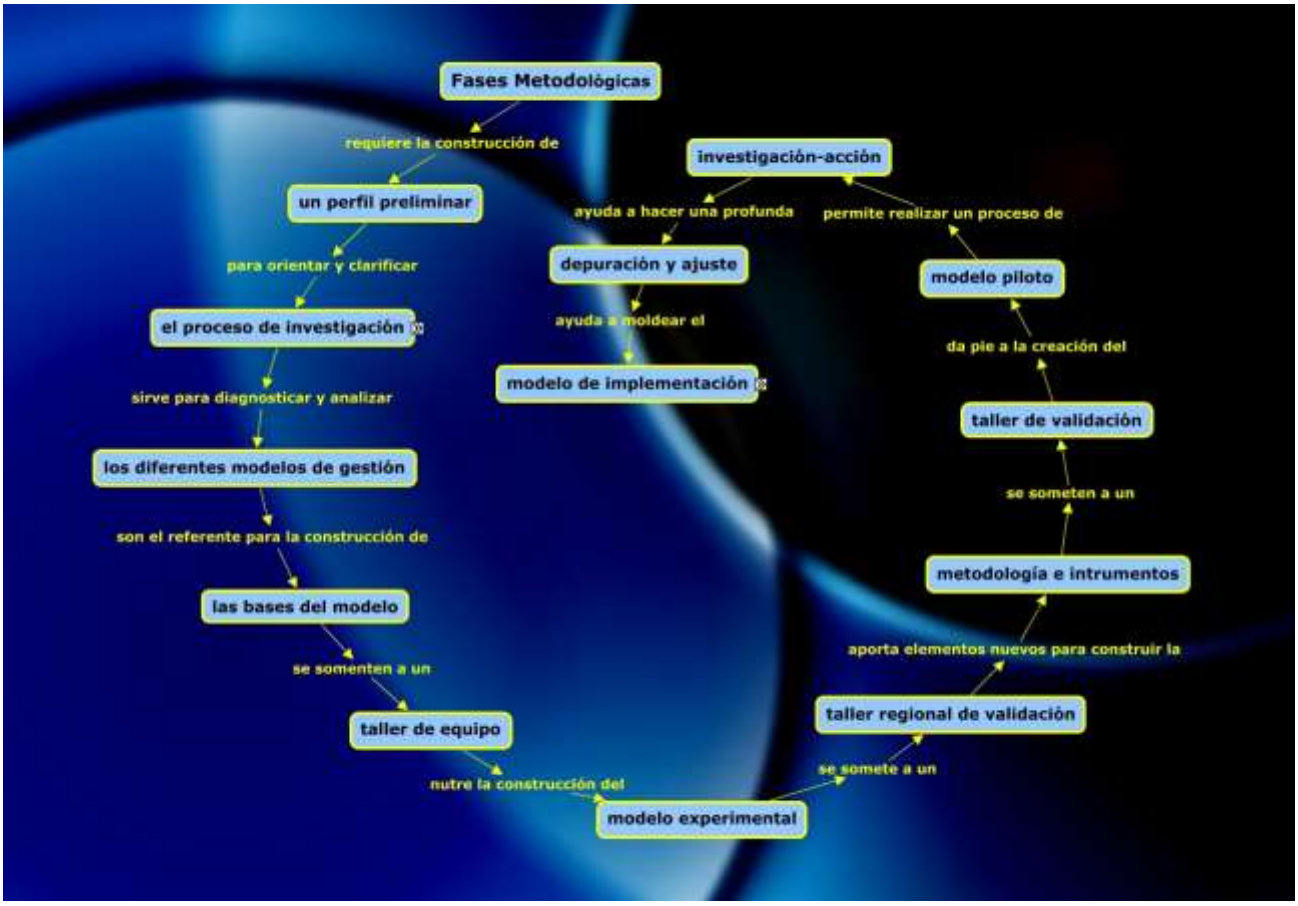


Ilustración 4: Fases Metodológicas

Fuente: (Pedraz, 2003)

CAPÍTULO II

DIAGNÓSTICO

2.1. ANTECEDENTES DIAGNÓSTICOS

En la Pontificia Universidad Católica del Ecuador Sede Esmeraldas aún se realizan procesos manualmente, como es el caso del proceso de consultas a los informes de tesis o proyectos de grado digitalizados, esto genera inconformidad en los usuarios (investigadores), al no tener total libertad para obtener información a estos documentos, debido a que realizan sus pedidos en biblioteca la cual tiene ciertas limitaciones en cuanto al acceso de los ejemplares. Es por esta situación que se llevó a determinar que es necesario el uso de un sistema informático que ayude a los investigadores, a agilizar el proceso de acceso a la información de estos documentos.

Para realizar un estudio sobre el proceso de préstamo y consulta que los investigadores realizan, se empleó el método de investigación Inductivo – Deductivo, el cual permitió obtener la información necesaria y captar los aspectos relevantes relacionados con el proyecto. Además mediante el uso de técnicas como la entrevista realizada al encargado de Biblioteca y al encargado del Departamento de Sistemas de la PUCESE, se pudo conocer de forma más precisa los procesos que se llevan a cabo, y los sistemas en donde alojan toda la información relacionada con los documentos de tesis o proyectos de grado digitalizados. También se empleó la técnica de la encuesta, realizada a la comunidad investigadora (estudiantes, docentes y miembros de CID), que dio a conocer cuáles son los inconvenientes al momento de dirigirse a realizar una consulta o al querer acceder a los documentos de tesis o proyectos de grado específicamente.

2.1.1. Objetivos Diagnósticos

- Conocer los procesos correspondientes a la administración de los informes de tesis o proyectos de grado de la PUCESE.
- Identificar los sistemas relacionados con el manejo y almacenamiento de la información correspondiente a los informes de proyectos o tesis de grado de la PUCESE.
- Determinar el grado de satisfacción de los estudiantes e investigadores con respecto al acceso de documentación de las tesis o proyectos de grado.

2.1.2. Variables Diagnósticas

- **Administración:** Los proceso llevados a cabo en la administración de los documentos de tesis o proyectos de grado
- **Tecnología:** Sistemas informáticos utilizados para el almacenamiento, manejo y repositorios de los informes de tesis o proyectos de grado.
- **Eficiencia:** Satisfacción de los investigadores (estudiantes, docentes y miembros del CID), acerca del acceso a los contenidos de los informes de tesis o proyectos de grado.

2.1.3. Indicadores

Administración

- ✓ Personal Involucrado.
- ✓ Políticas Departamentales.
- ✓ Registro y codificación de informes de tesis.
- ✓ Control de préstamos.
- ✓ Frecuencia en solicitudes de préstamo de informes de tesis.
- ✓ Tiempo empleado para la búsqueda
- ✓ Inconvenientes en el manejo de informes de tesis.
- ✓ Proyectos de grado digitalizados.

Tecnología

- ✓ Sistemas
- ✓ Bases de Datos
- ✓ Servidores
- ✓ Repositorios

Eficiencia

- ✓ Frecuencia de investigación en los documentos de tesis.
- ✓ Medios de consulta
- ✓ Motivos de consulta
- ✓ Inconvenientes al solicitar un informe de tesis.
- ✓ Grado de satisfacción.
- ✓ Factores que afectan el servicio de préstamo.
- ✓ Factores que impiden las consultas de contenidos.
- ✓ Parámetros de búsqueda.

2.2. MATRIZ DE RELACIÓN DIAGNÓSTICA

Con el propósito de visualizar y relacionar los objetivos, variables e indicadores con las técnicas que se utilizaron para captar la información y con las fuentes que nos proporcionaron dicha información, se muestra en la tabla 1 una matriz en la que se interrelacionan todos los componentes antes mencionados.

OBJETIVO	VARIABLES	INDICADORES	TÉCNICA	FUENTE
Conocer los procesos correspondientes a la administración de los informes de tesis o proyectos de grado de la PUCESE.	Administración	✓ Personal Involucrado. ✓ Políticas Departamentales. ✓ Registro y codificación de informes de tesis. ✓ Control de préstamos. ✓ Frecuencia en solicitudes de préstamo de informes de tesis. ✓ Tiempo empleado para la búsqueda ✓ Inconvenientes en el manejo de informes de tesis. ✓ Proyectos de grado digitalizados.	Entrevista	**Departamento de Biblioteca**
Identificar los sistemas relacionados con el manejo y almacenamiento de la información correspondiente a los informes de proyectos o tesis de grado de la PUCESE.	Tecnologías	✓ Sistemas ✓ Bases de Datos ✓ Servidores ✓ Repositorios	Entrevista	**Encargado del Departamento de Sistemas**

Determinar el grado de satisfacción de los estudiantes e investigadores con respecto al acceso a la documentación de tesis o proyectos de grado.	Eficacia	✓ Frecuencia de investigación en los documentos de tesis. ✓ Medios de consulta ✓ Motivos de consulta ✓ Inconvenientes al solicitar un informe de tesis. ✓ Grado de satisfacción. ✓ Factores que afectan el servicio de préstamo. ✓ Factores que impiden las consultas de contenidos. ✓ Paramentos de búsqueda.	Encuesta	**Investigadores (estudiantes, docentes y miembros del CID)**

Tabla 1: Matriz Diagnóstica

2.3. MECÁNICA OPERATIVA

2.3.1. Población

Según los datos obtenidos por la secretaría general de la PUCESE, en marzo del 2015, la población de los investigadores estaba conformada por: 1.600 estudiantes, 140 docentes y 10 miembros del CID. Con un total de: 1.750. Los mismos que fueron de vital importancia para determinar la situación actual del proceso de préstamos y consultas a los informes de tesis o proyectos de grado.

2.3.2. Muestra

Debido a la magnitud de la población es necesario aplicar la muestra. Para realizar la misma se aplicó el método del muestreo probabilístico estratificado, ya que todos los individuos de la población tienen las mismas posibilidades de ser seleccionados para ser parte de la muestra que se calculó con la fórmula 1.

$$n = \frac{N}{E^2(N-1)+1}$$

Fórmula 1: Muestreo Probabilístico

n =Tamaño de la muestra

N=Población

E=Error de muestreo

N= 1.750

E= 0,08

$$n = \frac{1.750}{0{,}08^2(1.750-1)+1}$$

$$n = \frac{1.750}{0{,}0064(1.749)+1}$$

$$n = \frac{1.750}{12,1936}$$

$n = 143,52 \rightarrow$ Tamaño de la muestra = 144

2.3.3. Información Primaria

Se aplicaron entrevistas al encargado de Biblioteca y al encargado del Departamento de Sistemas.

Además se realizaron encuestas a estudiantes, docentes y miembros del CID de la PUCESE, para determinar las necesidades y requerimientos de búsquedas de los informes de proyectos de grado.

Los cuestionarios de las correspondientes encuestas y entrevista se pueden observar en el anexo N°1, anexo N°2 y anexo N°3.

2.3.4. Información Secundaria

La información que se pudo recopilar, necesitó de consultas a distintas revistas, textos y artículos científicos.

2.3. ANÁLISIS Y TABULACIÓN DE DATOS

Encuesta realizada en marzo del 2015.

ENCUESTA DIRIGIDA A LOS ESTUDIANTES, DOCENTES Y MIENBROS DEL CENTRO DE INVESTIGACIÓN Y DESARROLLO (CID) DE LA PUCESE

OBJETIVO: *Identificar los requerimientos y necesidades de los investigadores con respecto a la búsqueda y recuperación de contenidos de los informes digitalizados de proyectos o tesis de grado de la PUCESE.*

1. **¿Utiliza la biblioteca para recopilar información de investigaciones desarrolladas en proyectos o tesis de grado de la PUCESE?**

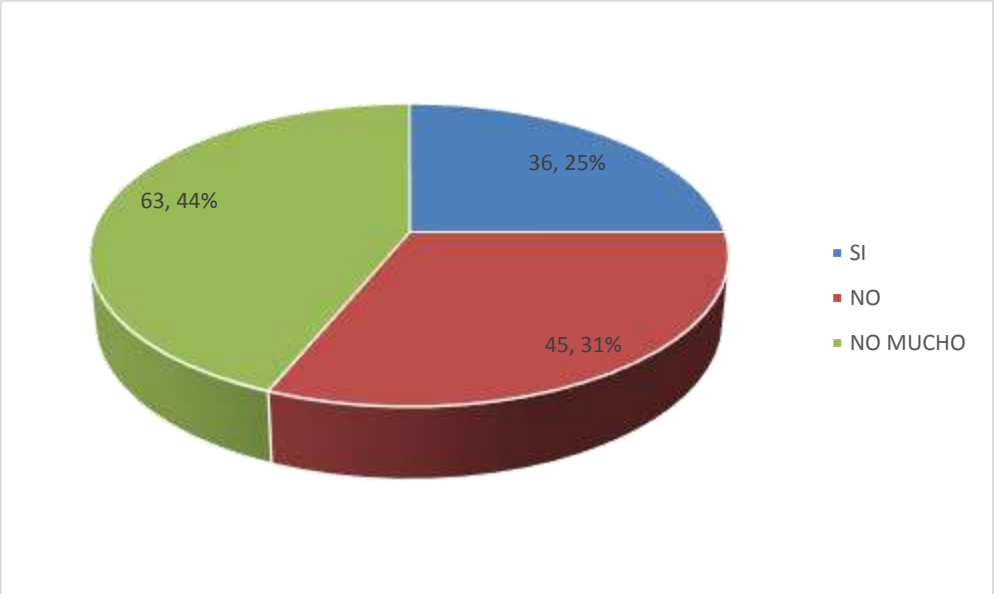


Ilustración 5: Pregunta 1-Investigadores

Interpretación

La mayoría de los investigadores encuestados expresaron que no acuden mucho al departamento de Biblioteca para recopilar información de investigaciones plasmadas en los documentos de tesis o proyectos de grado, mientras que muy pocos investigadores siempre realizan este tipo de investigaciones en la Biblioteca.

2. ¿Con que frecuencia realiza usted investigaciones en estos documentos?

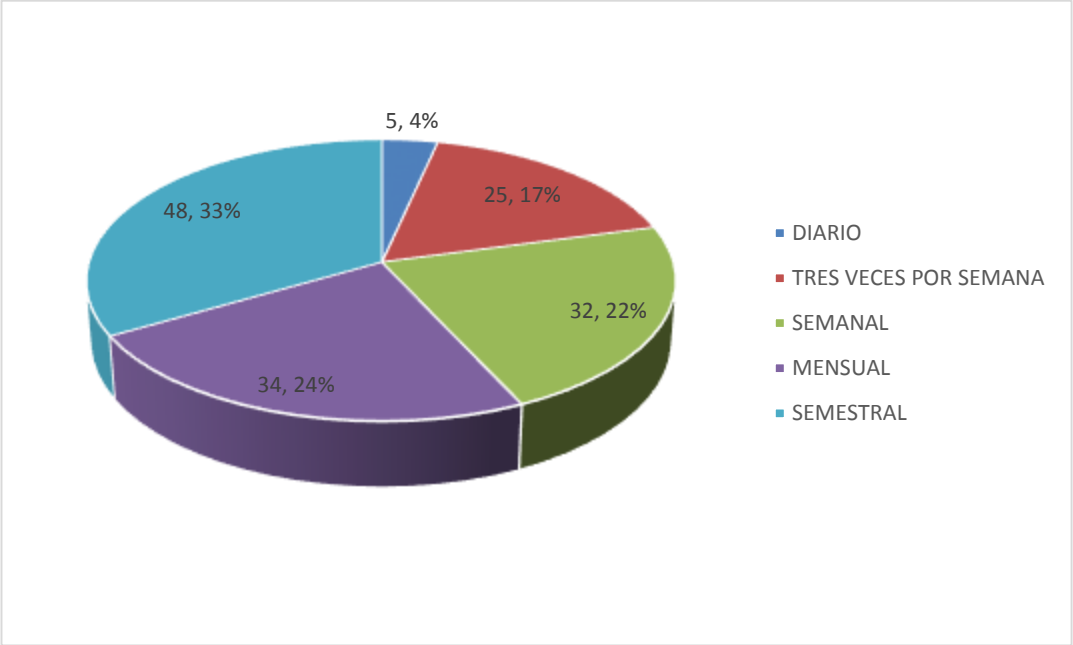


Ilustración 6: Pregunta 2-Investigadores

Interpretación

La frecuencia con la que los encuestados realizan investigaciones en los informes de tesis es semestral, debido a la necesidad que la mayoría de los investigadores tienen, en contar con información de diversas fuentes para la realización de proyectos e investigaciones que deben realizar en determinado momento.

3. ¿Cómo realiza usted la búsqueda de informes de tesis o proyectos de grado?

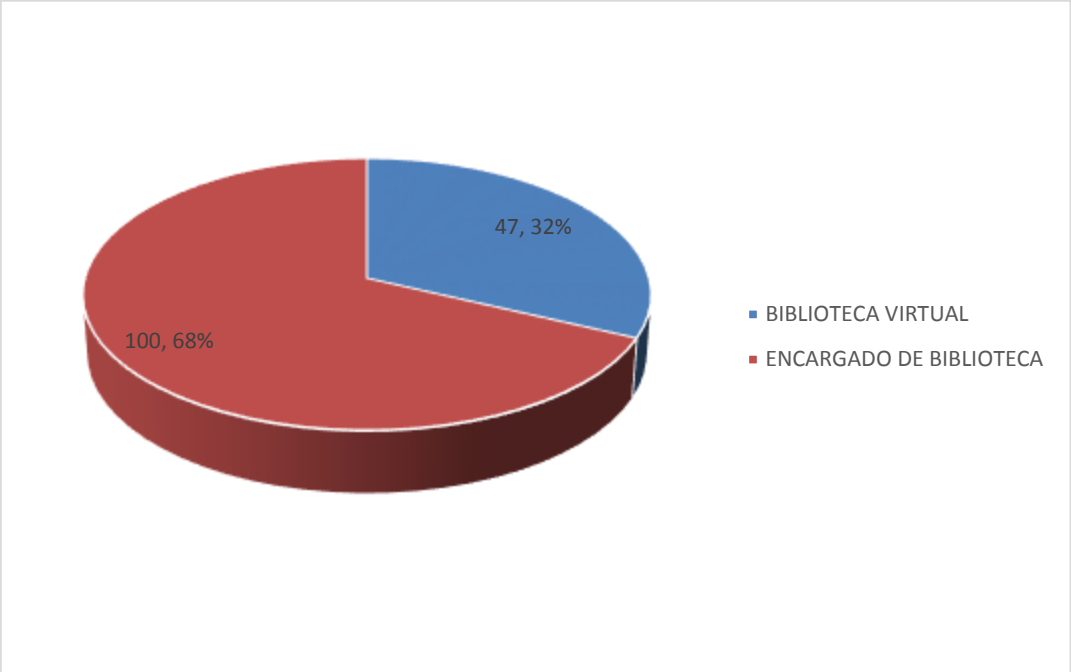


Ilustración 7: Pregunta 3-Investigadores

Interpretación

Al realizar investigaciones en este tipo de documentos las personas encuestadas prefieren acudir a la Biblioteca, esto se debe a que en la biblioteca virtual se encuentra información específica de los tomos disponibles, mas no pueden realizar las consultas y obtener información de los contenidos de los mismos.

4. ¿Cuál es el motivo de consulta de estos documentos?

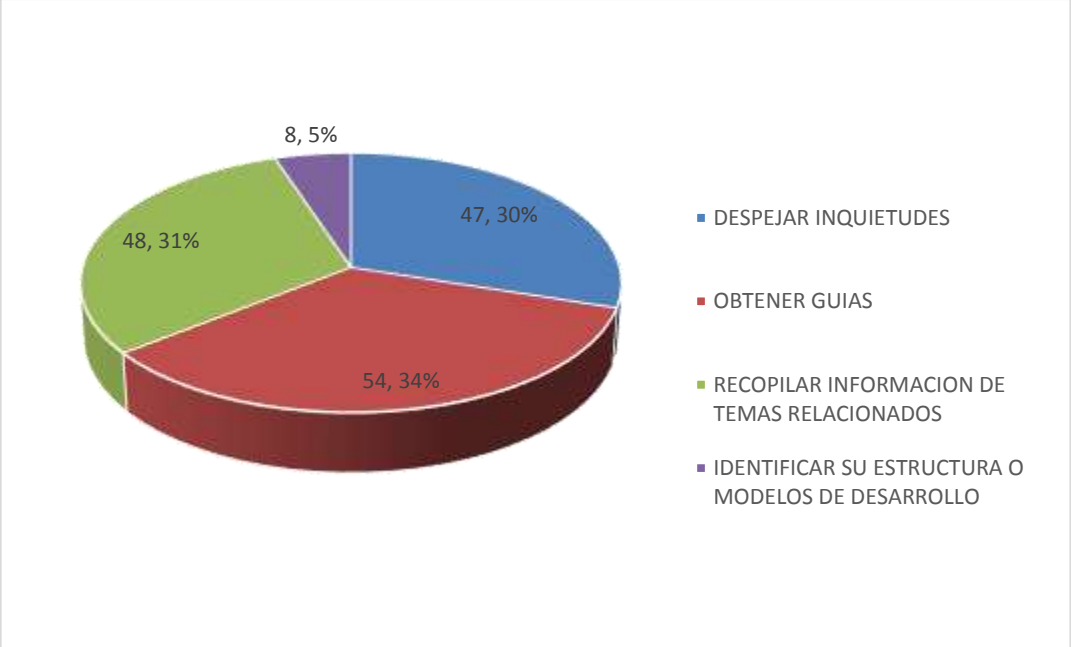


Ilustración 8: Pregunta 4-Investigadores

Interpretación

Observando la ilustración 8, se puede apreciar que tres opciones tienen porcentajes aproximados, lo que da a notar que las personas encuestadas realizan investigaciones en los documentos de tesis para despejar inquietudes sobre cualquier tema, obtener las guías necesarias para el encaminamiento en proyectos o investigaciones y recopilar toda la información necesaria que apoye al desarrollo de las mismas.

5. **¿Cuáles son los inconvenientes que tiene al momento de solicitar un informe de tesis de grado?**

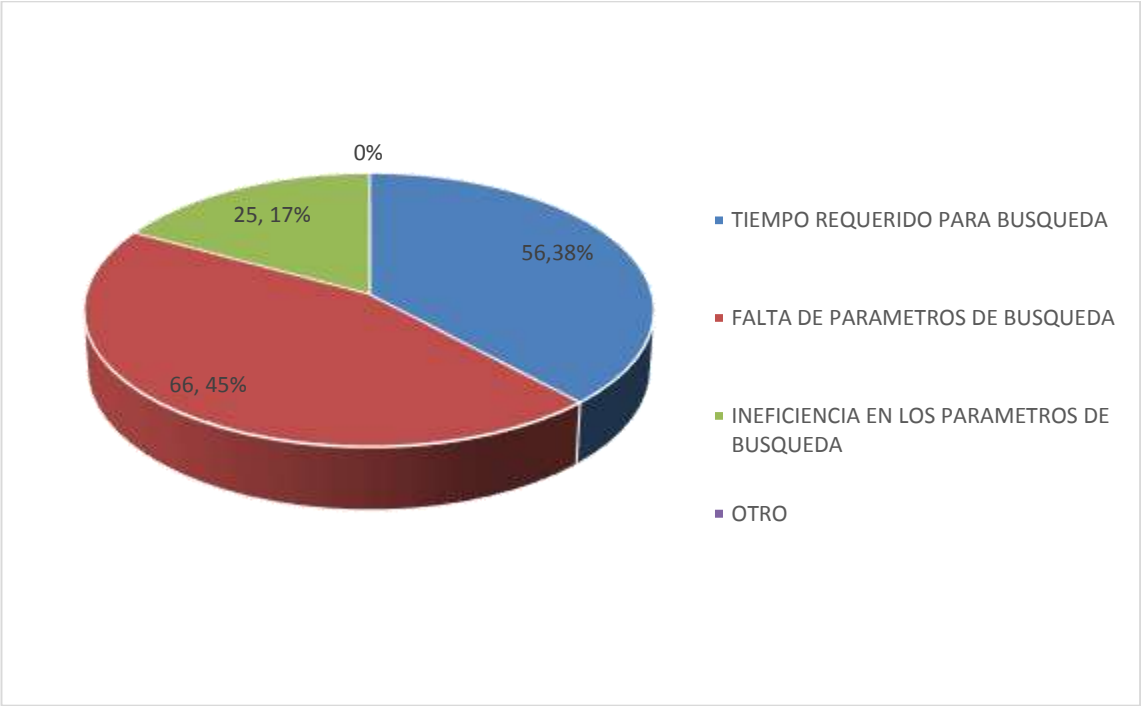


Ilustración 9: Pregunta 5-Investigadores

Interpretación

Casi la mitad de los encuestados expresaron que uno de los inconvenientes que tiene mayor relevancia al momento de solicitar un informe de tesis es la falta de parámetros en las búsquedas, ya que al existir este inconveniente se dificulta poder encontrar el informe que puede contener la información realmente necesaria para el investigador.

6. **¿Cuánto tiempo utiliza el encargado de biblioteca para realizar la búsqueda de los informes de tesis o proyectos de grado, en el proceso de préstamo de estos documentos?**

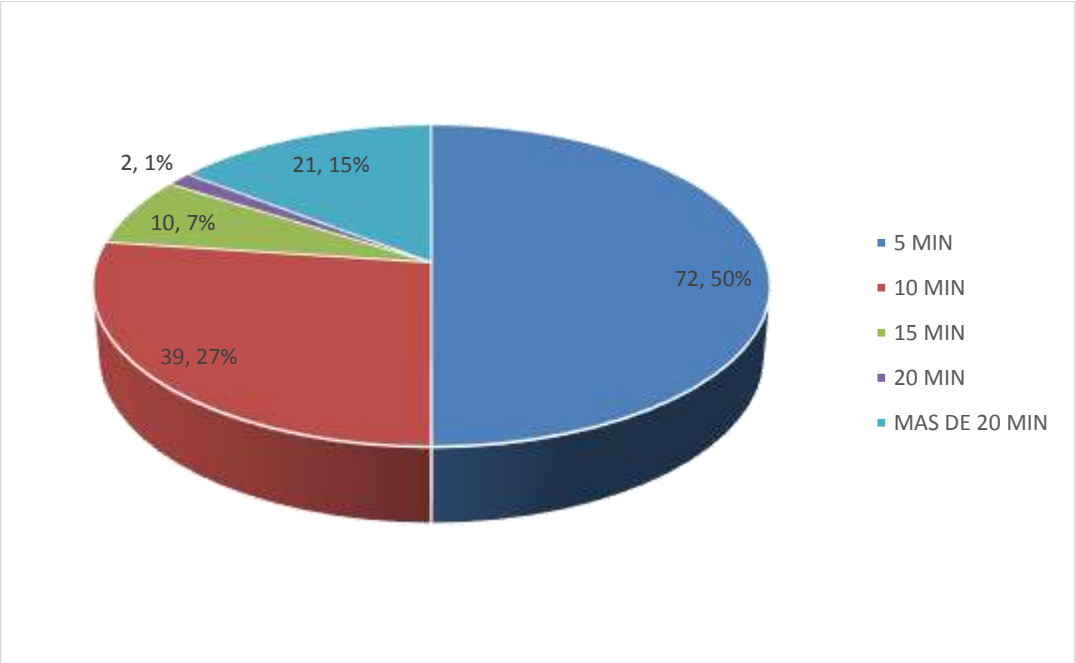


Ilustración 10: Pregunta 6-Investigadores

Interpretación

La mitad de los investigadores han sido atendidos en 5 minutos al momento de solicitar un informe de tesis, pero la otra mitad expresa que el encargado de biblioteca utiliza más de 10 minutos en encontrar el informe que contenga la información requerida. Esto podría estar sucediendo debido a que en ocasiones el encargado no pude encontrar el informe que el solicitante necesita por la falta de parámetros de búsqueda que permitan que se encuentre de forma más ágil el documento.

7. ¿Cómo califica el servicio de préstamo de informes de proyectos de grado en la PUCESE?

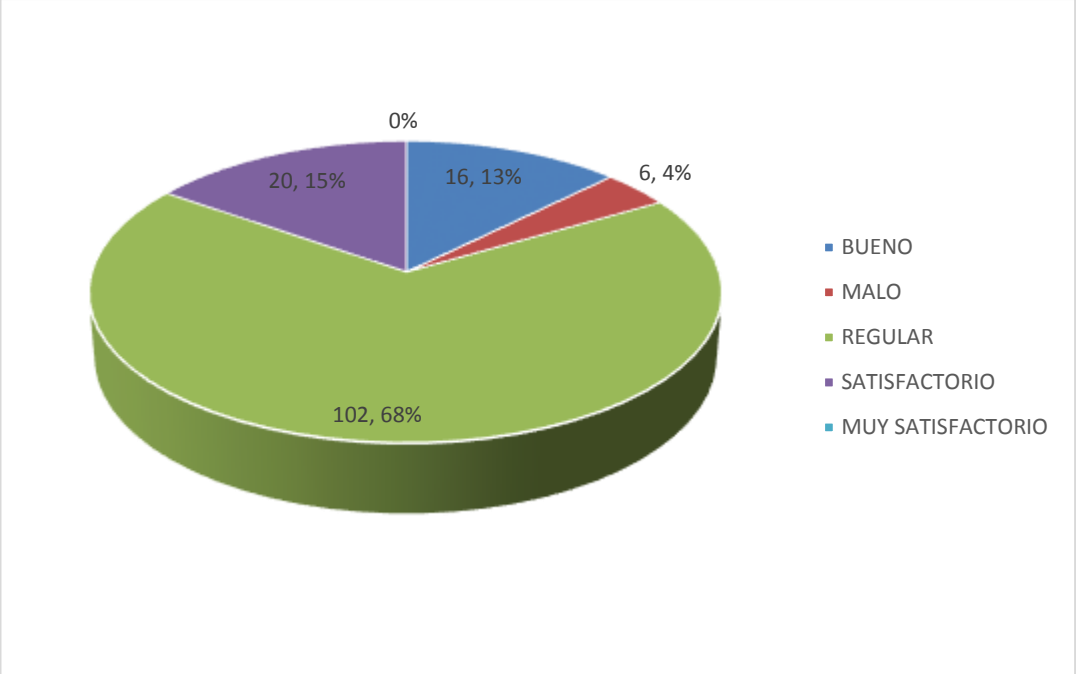


Ilustración 11: Pregunta 7-Investigadores

Interpretación

La mayor parte de los investigadores califican como regular el servicio que brinda el Departamento de Biblioteca, esto debido a que los tomos proporcionados no llenan las expectativas de los acudientes, en cuanto al contenido de información valiosa para sus investigaciones.

8. **Seleccione los factores que impiden la búsqueda de contenidos en tesis o proyectos de grado.**

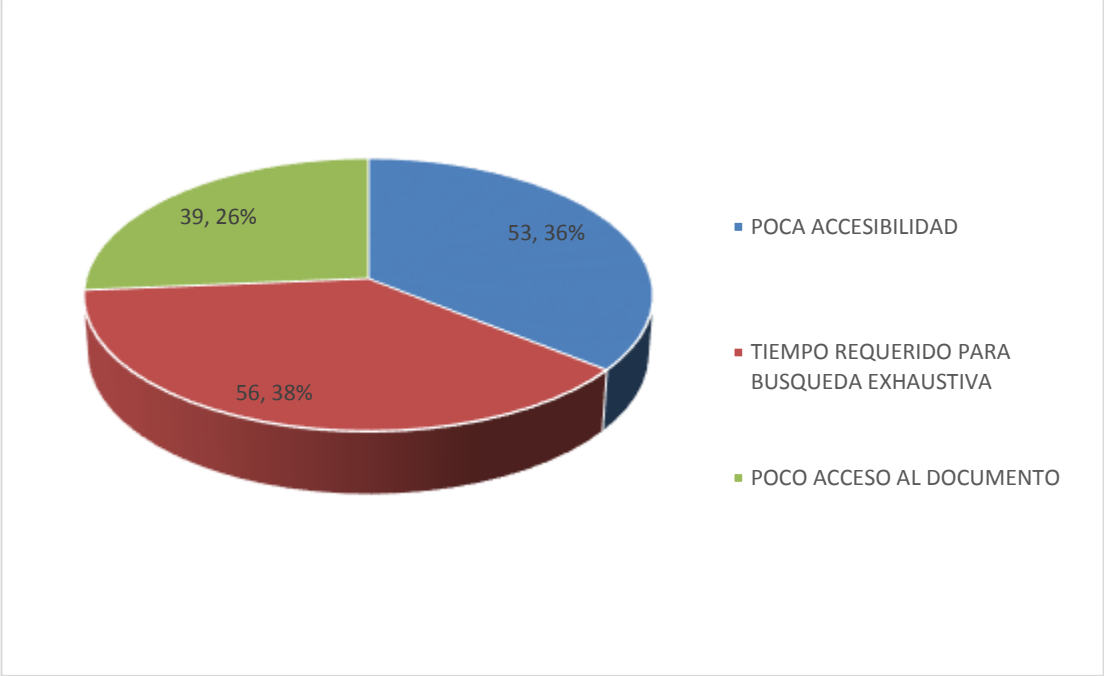


Ilustración 12: Pregunta 8-Investigadores

Interpretación

Los factores que le impiden a los investigadores realizar la búsqueda de información en los informes de tesis de grado es el tiempo que se requiere para realizar la respectiva revisión a los contenidos del documento para encontrar la información que pueda ser realmente necesaria y precisa para la investigación.

9. ¿Cuáles son los parámetros de búsqueda de contenidos a los que le gustaría acceder cuando solicite dicho documento?

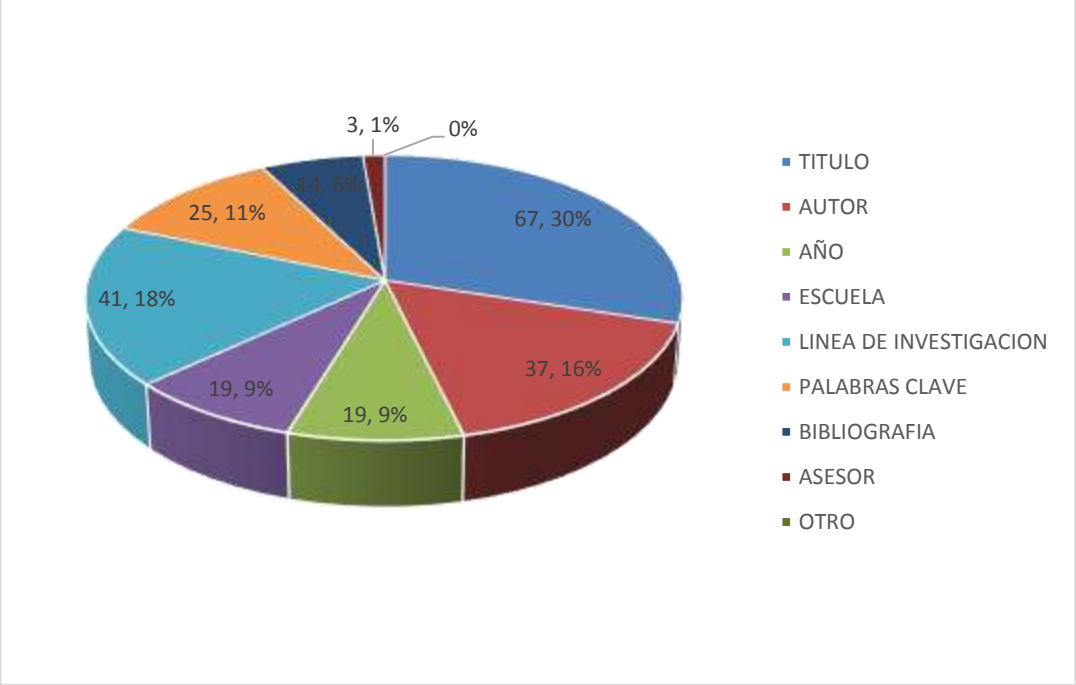


Ilustración 13: Pregunta 9-Investigadores

Interpretación

Sin duda el parámetro de búsqueda que los investigadores requieren y que les sirve de guía para buscar temas relacionados con sus investigaciones son los títulos de los informes, aunque hay que resaltar que no en todas ocasiones sea el punto clave para encontrar la información esperada. Se necesitan también de otros parámetros que sirvan de apoyo para ser más precisos en lo que se desea encontrar. Así es como se destacan también parámetros como: búsquedas por autor, líneas de investigación y palabras clave.

2.4. ANÁLISIS DE ENTREVISTAS

Entrevista al encargado de Biblioteca de la PUCESE – Marzo del 2016.

Al realizar la entrevista a la persona encargada de la biblioteca la Ing. Elina Morcillo, se pudo conocer las funciones que desempeña en la institución, lo cual es el apoyo a la docencia y atender a consultas bibliográficas de estudiantes y de la comunidad Esmeraldeña en general. En este departamento trabajan dos personas la bibliotecaria o encargada de la biblioteca y un auxiliar.

Con respecto a la codificación y archivo de los informes de tesis, se inicia el proceso luego de que Secretaría General hace la entrega del ejemplar mediante un memo, ingresando los datos respectivos del mismo en el inventario, luego se suben al repositorio digital y al SIABUC9 aquí se le asigna un código compuesto por letras que representa la especialidad y un número para determinar su ubicación, los usuarios deben ingresar a este sistema y obtener el código para luego dirigirse a biblioteca y solicitar el préstamo presentando la cédula de ciudadanía o carnet estudiantil.

Se manifestó que hay pocos inconvenientes en cuanto a los préstamos, no hay pérdida de documentos. En cuanto al tiempo de respuesta de las consultas a los investigadores se requieren de minutos, siempre y cuando el usuario sepa usar la búsqueda online y presente el código respectivo.

Para realizar los préstamos de los documentos de tesis o proyectos de grado los usuarios acuden a diario por ellos, desde luego existen restricciones que prohíben que los ejemplares salgan de la sala de la biblioteca o que sean fotografiados.

A la fecha existen alrededor de 500 documentos subidos al repositorio digital.

Entrevista al encargado del Departamento de Sistemas de la PUCESE – Marzo 2015.

En realizada al Mgt. Klever Posligua, se pudo conocer el software utilizado para el alojamiento de datos correspondientes a los documentos de tesis o proyectos de grado digitalizados.

Los datos se encuentran alojados en un servidor de tipo clon, en el Gestor de Base de Datos PostgreSQL de donde arrastra información el sistema Dspace que al momento no está en funcionamiento.

El proceso de digitalización de los documentos de tesis digitalizados está al 100%, lo que falta es la reestructuración de la base de datos del Departamento de Sistemas.

2.5. F.O.D.A.

Análisis aplicado a la Pontificia Universidad Católica del Ecuador Sede Esmeraldas, para identificar las fortalezas, oportunidades, debilidades y amenazas de la misma, que puedan afectar o apoyar de alguna u otra manera en el desarrollo del proyecto.

FORTALEZAS

F1. Personal capacitado.

F2. Alto porcentaje de egresados que culmina su tesis exitosamente.

F3. Disponibilidad de repositorios digitales.

DEBILIDADES

D1. Procesos aun sin automatizar.

D2. Demoras en la realización de algunos procesos.

D3. Desconocimiento por parte de los estudiantes sobre cómo se llevan los procesos para gestiones académicas.

OPORTUNIDADES

O1. Lograr la acreditación de todas las carreras.

O2. Compartir resultados de investigaciones con otros centros de educación superior.

O3. Cierre de la competencia.

AMENAZAS

A1. Evaluación del CEASES

A2. Apertura de nuevas universidades en la ciudad.

FACTORES INTERNOS / FACTORES EXTERNOS	FORTALEZAS	DEBILIDADES
	F1. Personal capacitado. F2. Alto porcentaje de egresados que culmina su tesis exitosamente. F3. Disponibilidad de repositorios digitales.	D1. Procesos aun sin automatizar. D2. Demoras en la realización de algunos procesos. D3. Desconocimiento por parte de los estudiantes sobre cómo se llevan los procesos para gestiones académicas.
OPORTUNIDADES O1. Lograr la acreditación de todas las carreras. O2. Compartir resultados de investigaciones con otros centros de educación superior. O3. Cierre de la competencia.	**FO** (F1,F2,O1) MANTENER LA POTENCIALIDAD DE DOCENTES Y ESTUDIANTES PARA LOGRAR LA ACREDITACIÓN.	**DO** (D1,D2,O1) EVALUACIÓN DE PROCESOS PARA BUSCAR LA MANERA DE AUTOMATIZARLOS, Y AYUDE A QUE DICHOS PROCESOS SE REALICEN MÁS ÁGILMENTE.
AMENAZAS A1. Evaluación del CEASES A2. Apertura de nuevas universidades en la ciudad.	**FA** (F1,F2,A1) MANTENER LOS PROCEDIMIENTOS QUE SE ESTAN REALIZANDO PARA APROBAR LA EVALUACIÓN.	**DA** (D1,D2,A1) ACELERAR LOS PROCESOS ESTANCADOS PARA SU PRONTA CULMINACIÓN.

Tabla 2: F.O.D.A.

2.6. DETERMINACIÓN DEL PROBLEMA DIAGNÓSTICO

Al haber culminado con la realización de las respectivas encuestas y entrevistas se pudo detectar las siguientes situaciones:

Los investigadores que desean obtener información de estos informes, acuden al Departamento de Biblioteca, en donde en ocasiones no quedan satisfechos con la información encontrada, debido a la falta de parámetros de búsqueda que les permita obtener la información correcta. Además para obtener el documento con la información precisa demanda mucho tiempo en la búsqueda, ya que el investigador debe realizar sus peticiones al bibliotecario para que pueda llevarles los documentos que él cree convenientes de acuerdo a la información proporcionada por el investigador.

De esta manera, con toda la información adquirida, se puede llegar al hecho que el proceso que se lleva a cabo actualmente, referente a la gestión de consultas en los informes de tesis o proyectos de grado es ineficiente, por lo que es factible el desarrollo de un sistema que automatice este proceso de modo que este pueda realizarse ágilmente.

CAPÍTULO III

PROPUESTA

3.1 JUSTIFICACIÓN

A nivel organizacional la automatización de la información denota comodidad tanto para las entidades como para los clientes, agilizando procesos que en otras instancias eran demasiado engorrosos, manifestando desventaja para la empresa por la insatisfacción de los clientes.

Los sistemas de información, permiten brindar soluciones a problemáticas de índole organizacional, administrativo y de seguridad, a través del registro, búsqueda, clasificación, y actualización de la información.

La implementación de un sistema de búsqueda y recuperación de documentos para los informes de tesis o proyectos digitalizados de la PUCESE, beneficiará de manera directa a la institución y a los funcionarios de la Biblioteca. De manera indirecta, a docentes, estudiantes, investigadores y público en general que requieran el servicio del acceso a esta documentación.

La institución será beneficiada, con el cumplimiento de un requerimiento para la acreditación, impuesta por el Consejo de Evaluación, Acreditación y Aseguramiento de la Calidad de la Educación Superior (CEAACES), que obliga a las universidades a digitalizar los documentos de tesis o proyectos de grado, y garantizar la administración óptima de los mismos.

Además, el sistema permitirá al encargado de biblioteca, reducir los tiempos de búsqueda física de esta documentación, agilitando el proceso de préstamo.

Los estudiantes, docentes, investigadores y público en general, obtendrán beneficio, al acceder de manera eficiente a los contenidos, e información relacionada con temas, líneas de investigación, autor, asesor, escuela, palabras claves, entre otros, de los proyectos de grado desarrollados por los diferentes graduados de la institución, a través de búsquedas que relacionan dichos parámetros.

De esta manera, el proyecto contribuirá de forma positiva a la comunidad universitaria, optimizando la gestión de esta documentación y complementando el proceso de digitalización de los informes de proyectos de tesis de la institución.

3.2 OBJETIVOS

3.2.1 General

- Optimizar la gestión de consultas a los informes de proyectos o tesis de grado de la Pontificia Universidad Católica del Ecuador Sede Esmeraldas (PUCESE), mediante la implementación de un sistema de información para la búsqueda y recuperación de contenidos.

3.2.2 Específicos

- Aplicar técnicas de búsqueda como es la de recuperación difusa, para permitir que las búsquedas sean eficientes dejando que el usuario coloque la palabra o frase que considere relevante, obviando acentos, mayúsculas, etc. evitando así que los resultados se vean afectados.

- Emplear el algoritmo de búsqueda secuencial, que permita al sistema recorrer los datos, destacando las coincidencias dependiendo de la palabra o palabras claves que sean ingresadas.

- Implementar un sistema de búsqueda y recuperación de documentación, orientado a los informes de tesis de grado de la PUCESE.

- Elaborar un manual, para que los usuarios del sistema tengan noción del funcionamiento del mismo.

3.3 SOLUCIÓN INFORMÁTICA.

3.3.1 Metodología de Software.

Para poder desarrollar el presente proyecto se dispuso del uso de una metodología de software que permitiese garantizar la presencia de calidad, de poder identificar los riesgos que pudieran presentarse y que pudiera dar seguridad de que el sistema pueda estar sujeto a cambios en el origen de datos, en caso de que estos presentasen modificaciones.

De esta manera se determinó que la metodología a implementa fuera Programación Extrema o XP, por permitir que la programación se realice ágilmente, tratando de poder resolver problemas ante imprevistos que se presenten durante el desarrollo del proyecto. Esta metodología garantiza la simplicidad para enfocarse en el desarrollo de la programación antes que en el diseño del sistema, en esta parte se documenta la programación, poniendo los nombres adecuados a funciones y clases y agregando comentarios que hagan que la programación sea más organizada y no se llegue a crearse dudas o a desorientarse durante el proceso de codificación.

3.3.2 Diseño

Este sistema está dirigido a los investigadores, estudiantes, docentes y público en general, los cuales a través de una conexión a internet podrán acceder al sistema ingresando a la página web de la PUCESE.

Dentro del sistema podrán acceder a información relacionada a los documentos de tesis o proyectos de grado digitalizados disponibles en el repositorio de la institución, dependiendo del tipo de búsqueda que seleccionen ya sea búsqueda por categoría o búsqueda por contenido, deberán ingresar en cualquiera de estas búsquedas los parámetros adecuados para obtener los resultados correspondientes a su petición.

3.3.2.1 Esquema Funcional del Sistema.

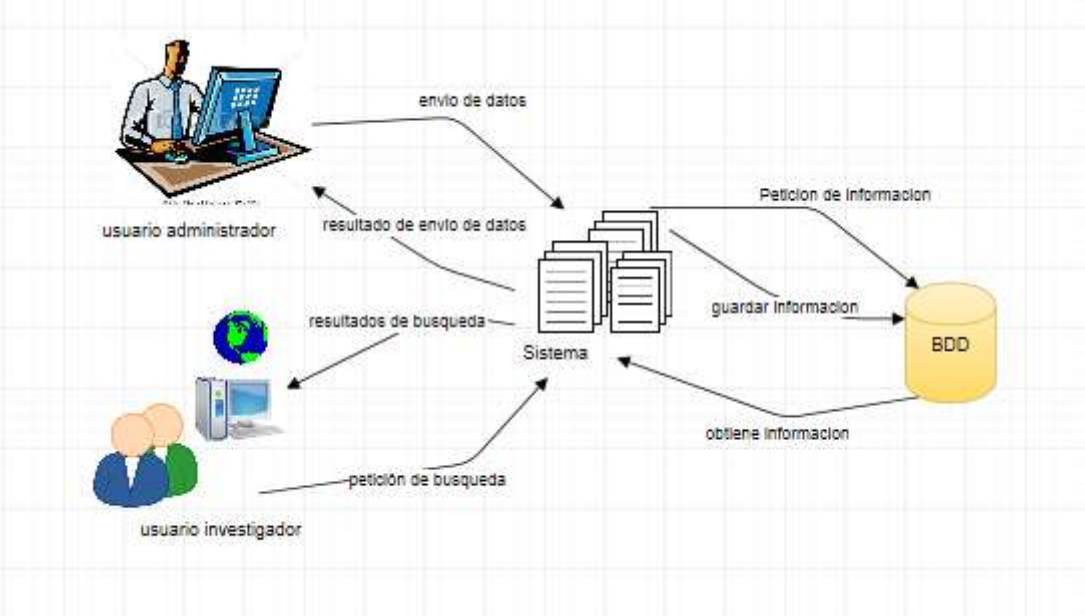


Ilustración 14: Diagrama Funcional del Sistema

3.3.2.2 Diagrama de Flujo de Datos (DFD)

3.3.2.2.1 Nivel cero – Diagrama de contexto.

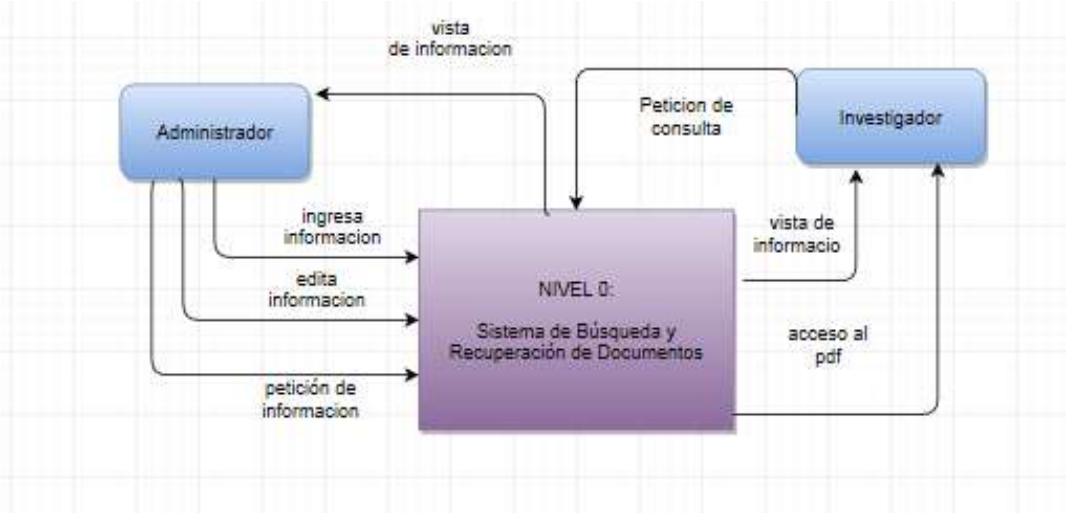


Ilustración 15: DFD - Nivel 0

3.3.2.2.2 Nivel uno – Diagrama de nivel superior.

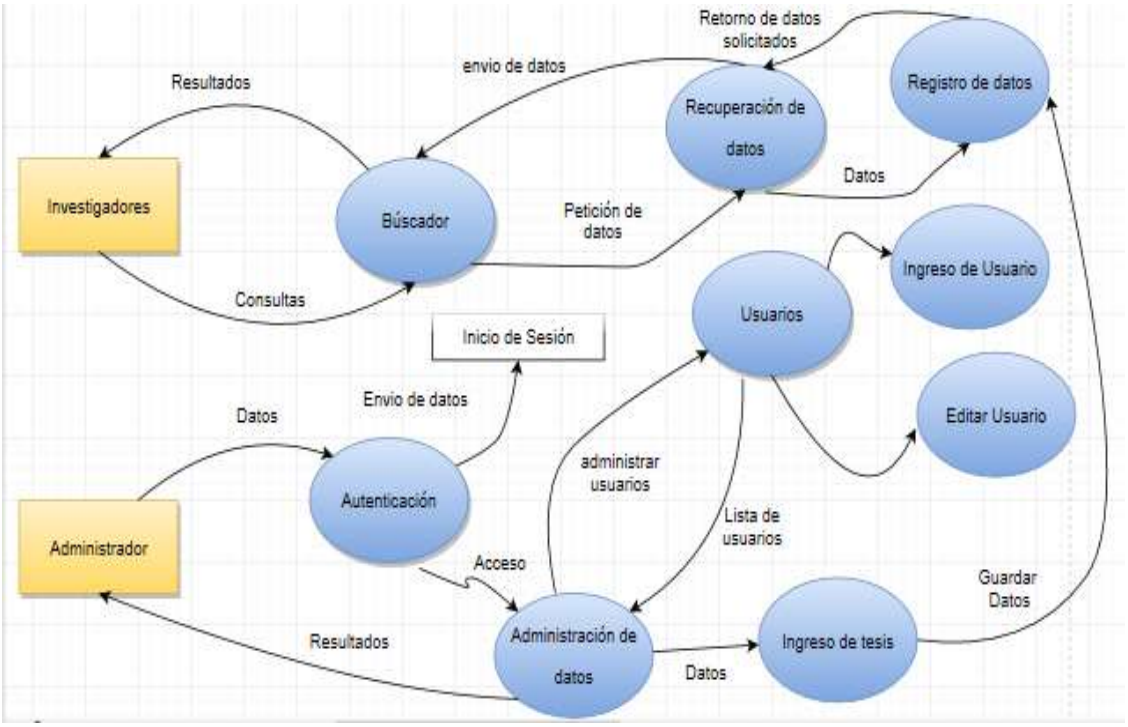


Ilustración 16: DFD Nivel 1

3.3.2.2.3 Nivel dos: Investigadores

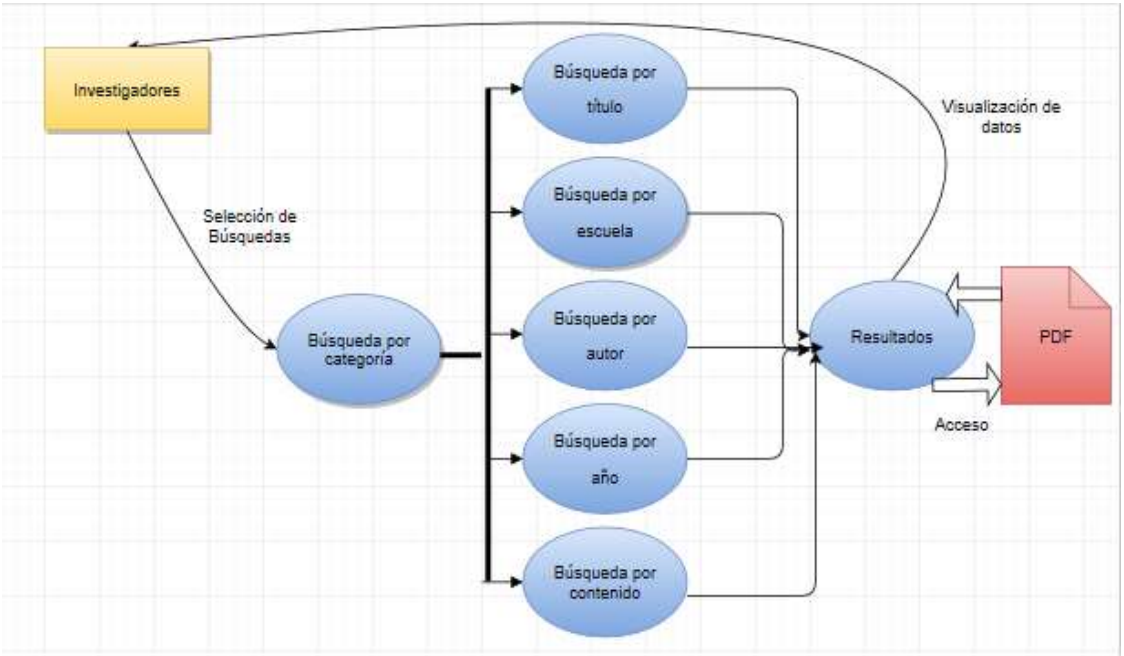
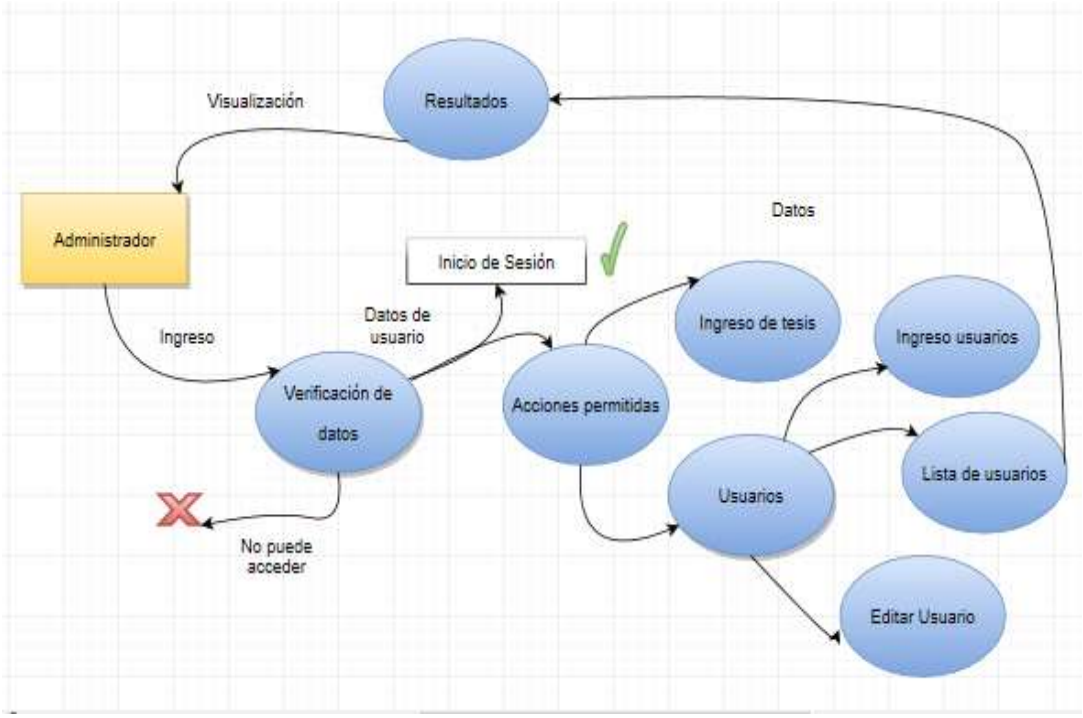


Ilustración 17: DFD Nivel 2 - Investigadores



3.3.2.2.4 Nivel dos: Administrador

Ilustración 18: DFD Nivel 2 – Administrador

3.3.2.3 Diagramas de Caso de Uso.

A través de los diagramas de UML o Lenguaje de Modelado Unificado, refiriéndose a los diagramas de caso de uso específicamente, los cuales reflejan la interacción de los usuarios con la aplicación. Se muestra a continuación en los siguientes diagramas la representación del funcionamiento del sistema y como los investigadores interactuaran con el mismo.

3.3.2.3.1 Diagrama de Caso de Uso. Etapa 1: Ingreso al sistema.

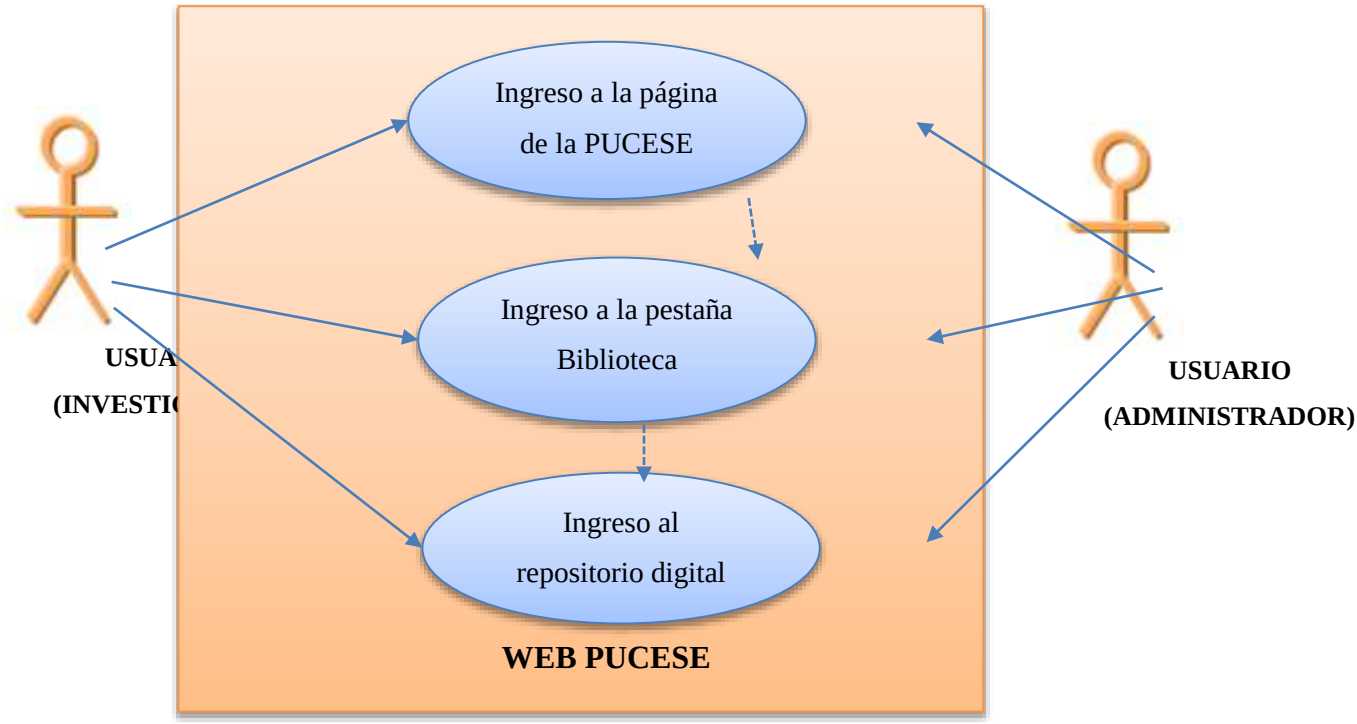


Ilustración 19: Diagrama de caso de uso. Primera etapa - ingreso al sistema

3.3.2.3.2 Diagrama de Caso de Uso. Etapa 2: Elección de búsqueda y autenticación de usuario.

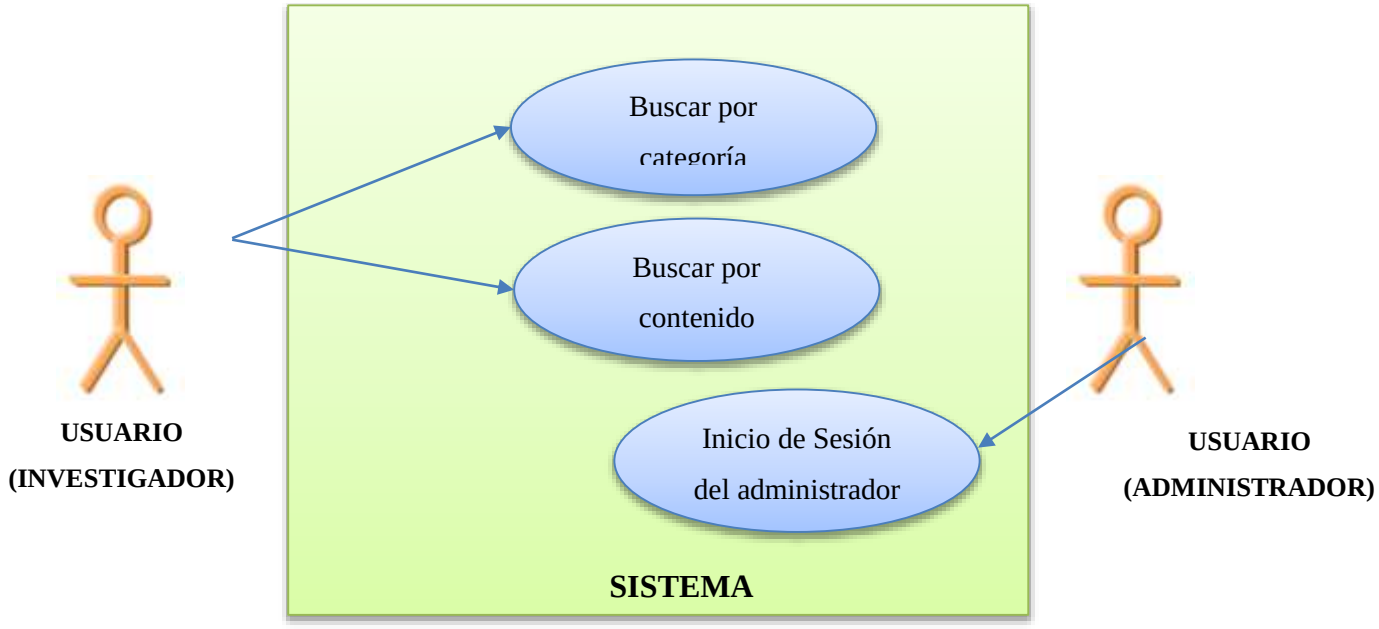


Ilustración 20: Diagrama de caso de uso. Segunda etapa – Selección del tipo de búsqueda.

3.3.2.3.3 Diagrama de Caso de Uso: Tipo de búsqueda por categoría.

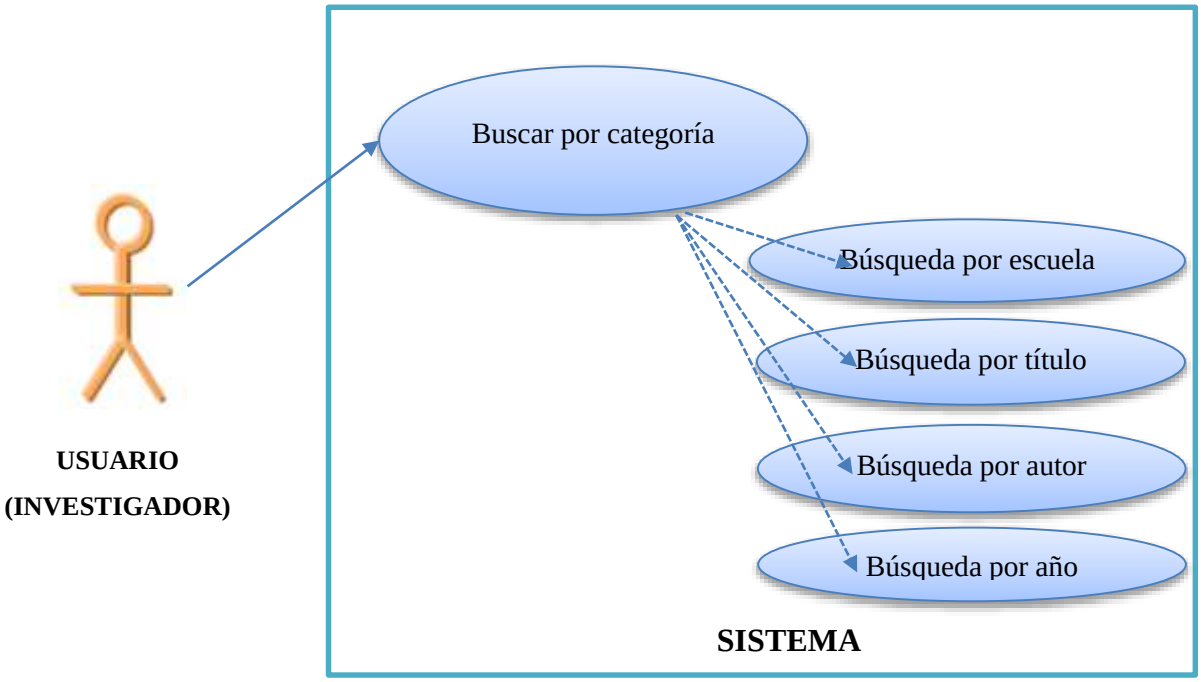


Ilustración 21: Diagrama de caso de uso. Búsqueda por categoría.

3.3.2.3.4 Diagrama de Caso de Uso: Tipo de búsqueda por contenido.

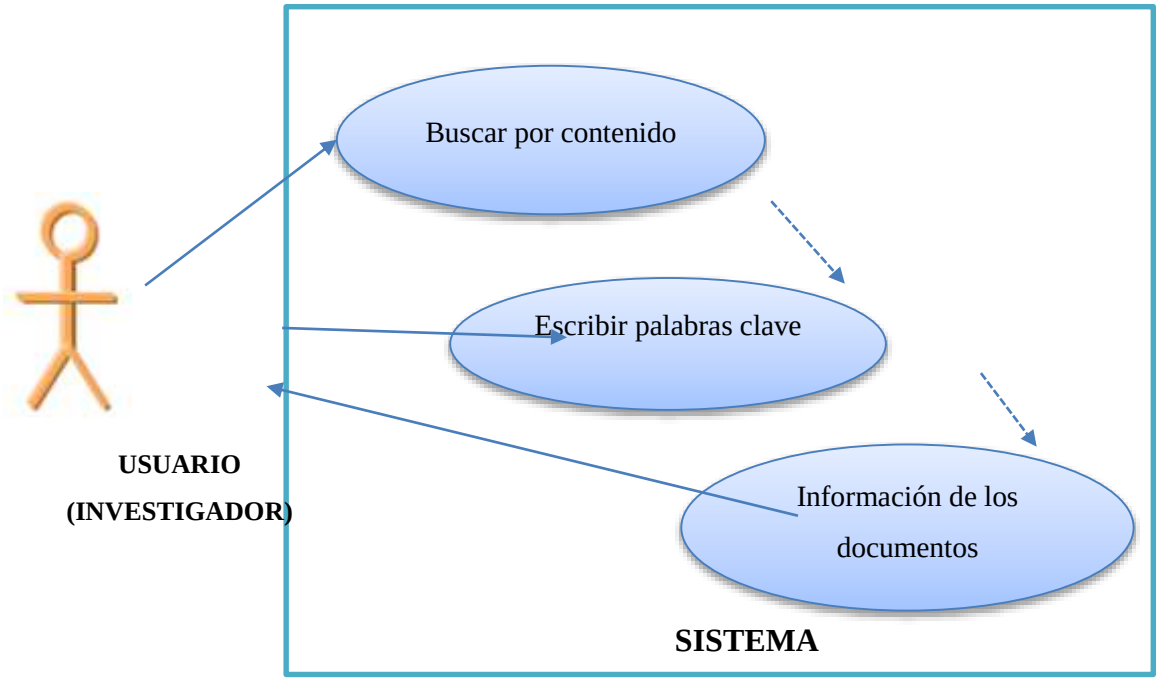


Ilustración 22: Diagrama de caso de uso. Búsqueda por contenido.

3.3.2.3.5 Diagrama de Caso de Uso: Vista del Administrador.

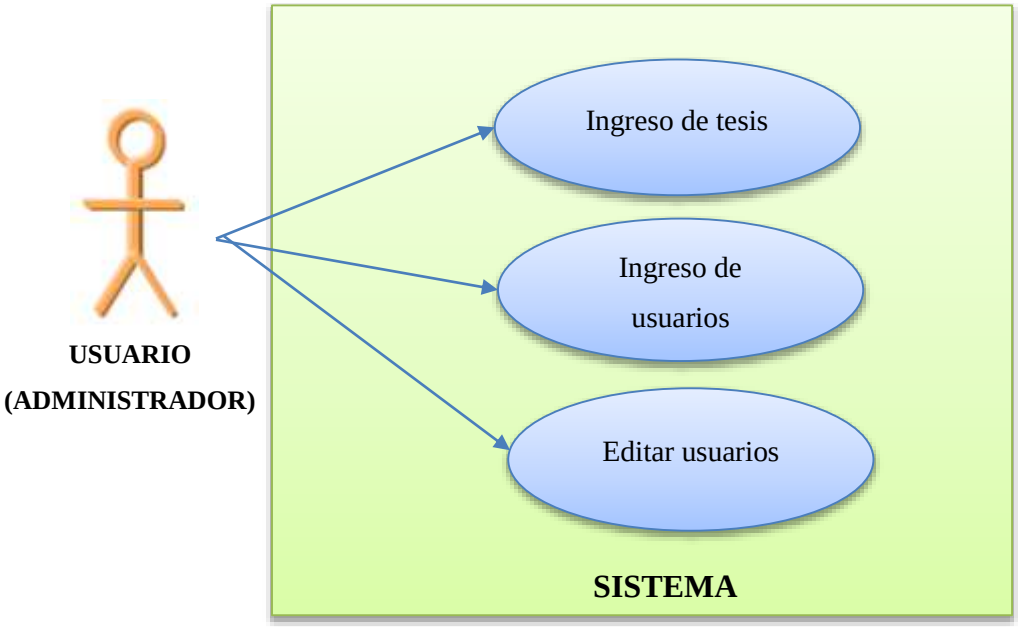


Ilustración 23: Diagrama de Caso de Uso - Vista del administrador

3.3.3 BASE DE DATOS PostgreSQL.

Para poder empezar con el desarrollo del sistema, fue totalmente indispensable toda la información que la Pontificia Universidad Católica del Ecuador Sede Esmeraldas (PUCESE), almacena sobre los documentos digitalizados de las tesis o proyectos de grado. Para ello el departamento de sistemas de la institución autorizó el uso de una copia de la misma, con los datos contenidos en el sistema gestor de base de datos PostgreSQL

De esta manera después de realizar un análisis y revisión exhaustiva de todas las tablas de la base de datos para identificar las tablas a utilizar, las relaciones entre ellas e identificar los campos que contienen la información que se va a necesitar para el sistema.

Se destacaron las tablas que a continuación se describen de forma más detallada en la siguiente tabla.

Tabla	Descripción	Campos	Tipo de campo	Campos destacados
bi_2_dis	Nombre de los **autores** de los documentos.	Id (PK)	Integer	**Value (nombre)**
		authority	Character varying (100)	
		value	Text	
		sort_value	Text	
bi_item	**Temas** y **fechas** de los documentos.	Id (PK)	Integer	**sort_1 (título)** **sort_2 (año)**
		item_id	Integer	
		sort_1	Text	
		sort_2	Text	
		sort_3	Text	
	Relación entre bi_2_dis y	map_id (PK)	Integer	

bi_2_dmap	bi_item.	item_id	Integer	
		distinct_id	integer	
community	**Escuelas** de la PUCESE.	community_id (PK)	Integer	Name (escuelas)
		name	Character varying (128)	
		short_description	Character varying (512)	
		introductory	Text	
		logo_bitstream_id	Integer	
		copyright_text	Text	
		side_bar_text	Text	
		admin	integer	
communities2item	Relación entre community y bi_item	Id (PK)	Integer	
		community_id	Integer	
		item_id	integer	

bitstream	**Ruta** del servidor donde se almacenan los documentos digitales	bitstream_id (PK)	Integer	**Sourse (ruta/dirección)**
		bitstream_format_id	Integer	
		name	Character varying (256)	
		side_bytes	Bigint	
		checksum	Character varging (64)	
		checksum_algorithm	Character varying (32)	
		description	Text	
		user_format_description	Text	
		sourse	Character varyin (256)	
		internal_id	Character varying (256)	
		deleted	Boolean	
		store_number	Integer	
		sequence_id	integer	

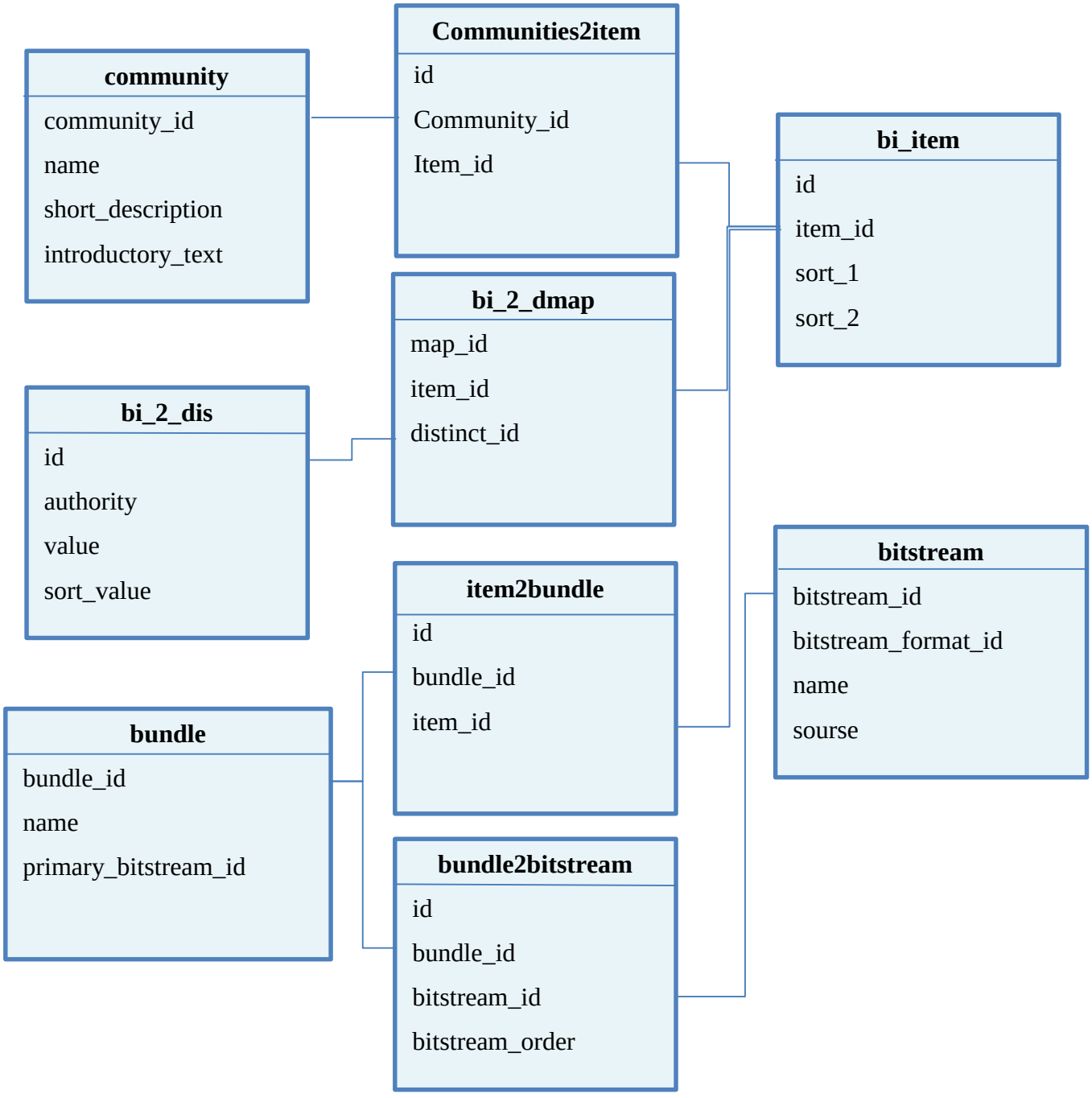
bundle	Puente para llegar a relacionar bitstream con bi_item	bundle_id (PK)	Integer	
		name	Character varying (16)	
		primary_bitstream_id	integer	
bunde2bitstream	Relación entre bitstream y bundle	Id (PK)	Integer	
		bundle_id	Integer	
		bitstream_id	Integer	
		bitstream_order	integer	
item2bundle	Relación entre bundle y bi_item	Id (PK)	Integer	
		item_id	Integer	
		bundle_id	Integer	

Tabla 3: Descripción de tablas en Postgresql

3.3.4 MODELADO DE DATOS

3.3.4.1 MODELO CONCEPTUAL

En el siguiente esquema conceptual se puede apreciar de modo más amigable las tablas que fueron tomadas de la base de datos para realizar el proyecto y las relaciones identificadas entre ellas, las cuales fueron imprescindibles para hacer coincidir los datos que van a ser mostrados al usuario.



3.3.4.2 MODELO ENTIDAD RELACIÓN

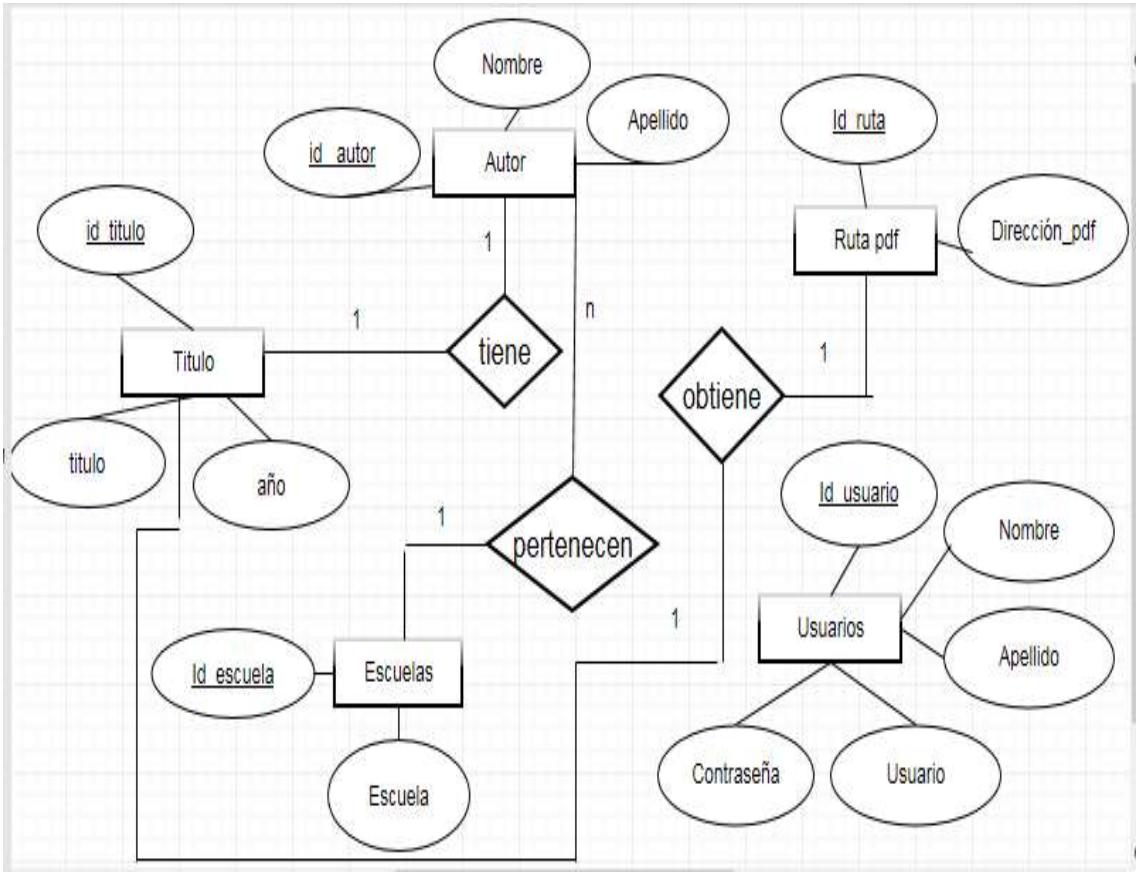


Ilustración 24: Modelo Entidad Relación

3.3.4.3 MODELO RELACIONAL

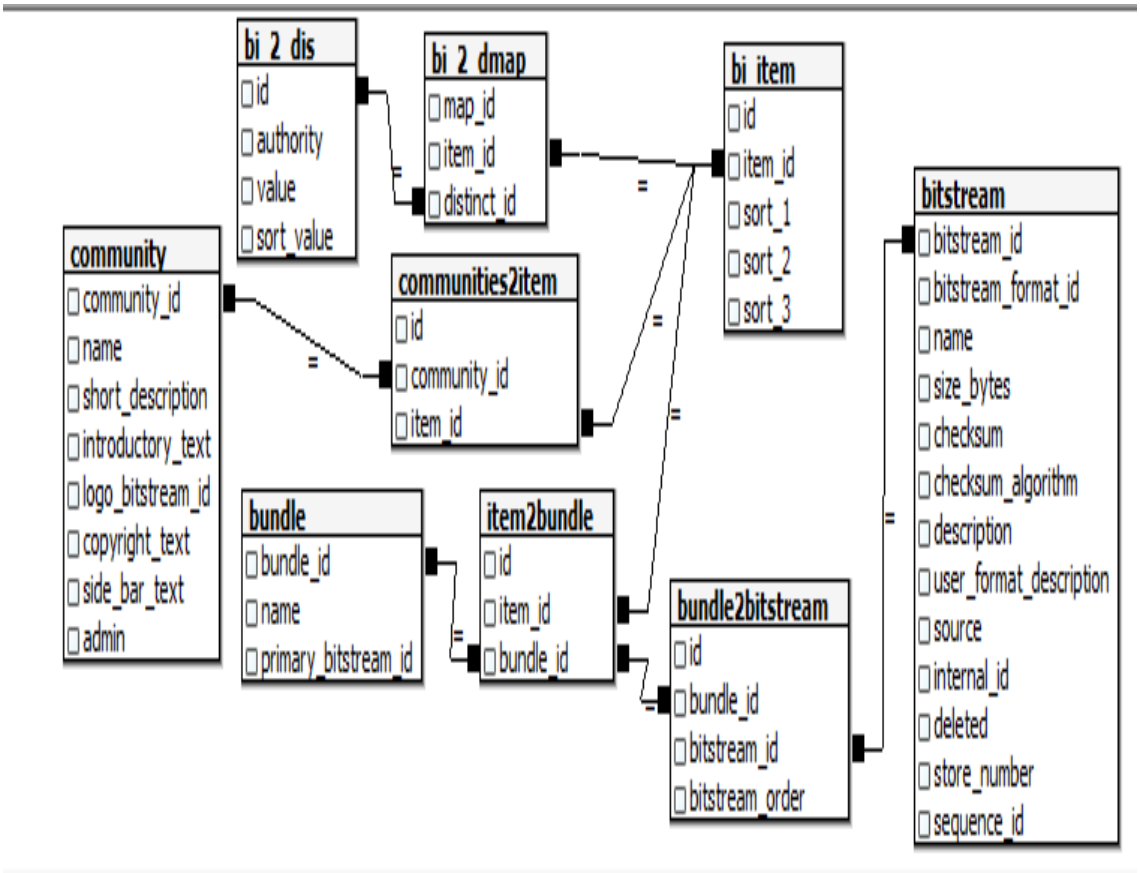


Ilustración 25: Modelo Relacional

3.3.5 PLATAFORMAS DE DESARROLLO.

El software utilizado en el sistema, fue escogido por su versatilidad y flexibilidad en cuanto a desarrollo de proyectos web se refiere, proporcionando al desarrollador las herramientas óptimas para agregar las funcionalidades que necesite, es por este motivo que para el desarrollo del proyecto se tomaron en cuenta las siguientes plataformas:

- Xampp 3.2.1: servidor local, el cual permitió realizar las pruebas del sistema durante todo el proceso de desarrollo.

- PHP: lenguaje de programación que proporcionó la interfaz para poder establecer la conexión con la base de datos y por la cual a través de funciones se logró la interacción con la misma.

- Jquery: Permite proporcionarle dinamismo a la página del sistema.

- CSS: hoja de estilo que permitió dotar de personalidad al sistema al darle un toque amigable y vistoso para el usuario.

- PostgreSQL: gestor de la base de datos proporcionada, del cual fue extraída toda la información del sistema.

3.3.6 MÓDULOS DEL SISTEMA

El sistema cuenta con el módulo de búsqueda por categoría el cual contiene:

- Búsqueda por título.
- Búsqueda por escuela.
- Búsqueda por autor.
- Búsqueda por año.
- Búsqueda por contenido.

El módulo del administrador:

- Ingreso de tesis.
- Ingreso de usuarios.
- Editar usuarios.

3.3.7 MANUAL DE USUARIO

Página principal

La siguiente ilustración muestra la página principal del sistema.



Ilustración 26: Página Principal del Sistema

En esta sección puede seleccionar el tipo de búsqueda que se desee realizar. Búsqueda por título, escuela, año, autor y búsqueda por contenido.

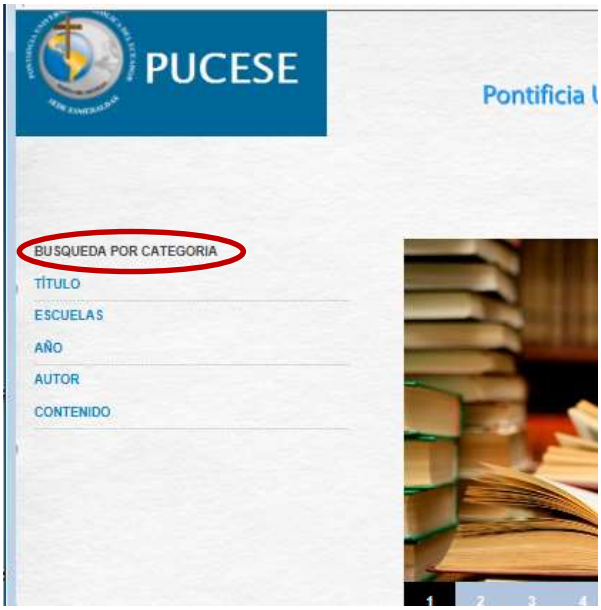


Ilustración 27: Opciones de búsqueda

Al darle click en la opción título, se tendrá que colocar una palabra referente al título que está buscando. De esta manera aplastando en el botón buscar se desplegaran todas las coincidencias encontradas, y en donde aparecerá marcada la palabra de referencia.



Ilustración 28: Búsqueda por título

En la opción escuela aparecerá el listado de todas las escuelas vigentes en la institución. Se deberá seleccionar la escuela de la cual se requiere información de sus documentos de tesis o proyectos de grado.



Ilustración 29: Búsqueda por Escuela

En la opción año se desplegará una lista de los años en que fueron creados los documentos.



Ilustración 30: Búsqueda por año

Para buscar por autor, se debe ingresar el nombre o apellido o ambos, para poder obtener la información referente al proyecto de una persona en específico.



Ilustración 31: Búsqueda por autor

Para buscar por contenido se deberá ingresar la palabra o frase que se tiene de referencia, a continuación el sistema buscará en el contenido del documento extrayendo la información de forma textual evaluando las coincidencias entre el dato ingresado y el texto del documento.

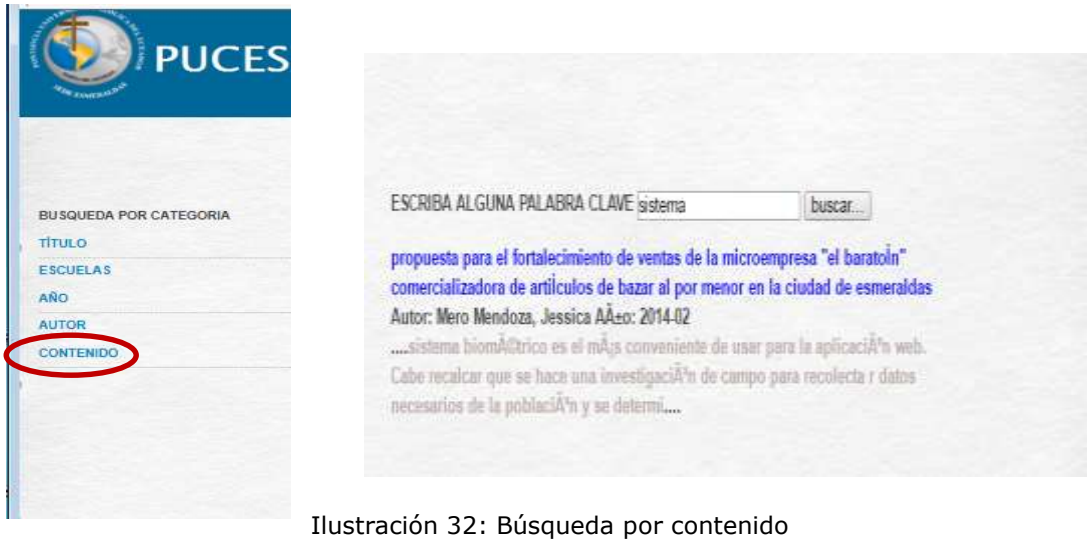


Ilustración 32: Búsqueda por contenido

En esta sección el usuario administrador podrá iniciar sesión, colocando el respectivo usuario y contraseña.



Ilustración 33: Inicio de sesión

El súper usuario puede realizar varias acciones, entre ellas agregar un nuevo tema de tesis con los datos correspondientes al autor, fecha, escuela y el documento en formato pdf.



Ilustración 34: Ingreso de Tesis

También puede agregar otros usuarios administradores, con sus respectivos datos de identificación en el sistema.



Ilustración 35: Ingreso de usuarios

Este súper usuario además puede obtener una lista de los demás administradores del sistema editarlos y eliminarlos.

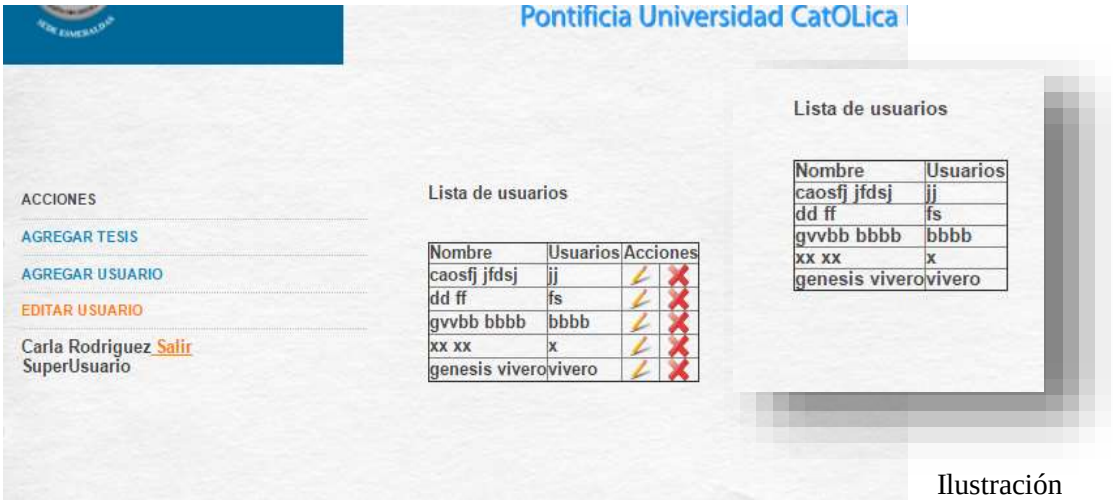


Ilustración 36: Editar usuario

CAPÍTULO IV
IMPACTOS

4.1 ANÁLISIS DE IMPACTOS

Al dar por culminado el sistema de búsqueda aplicado a las tesis o proyectos de grado digitalizados de la Pontificia Universidad Católica del Ecuador Sede Esmeraldas (PUCESE), se ha determinado evaluar los impactos que el presente proyecto causara. Tomando en cuenta que la incidencia de los impactos puede ser tanto positiva como negativa, se analizará el impacto tecnológico, económico y social, dependiendo de los valores que indica la siguiente matriz de impacto.

VALOR IMPACTO	DESCRIPCIÓN
-3	Impacto alto negativo
-2	Impacto medio negativo.
-1	Impacto bajo negativo.
0	No hay Impacto.
1	Impacto bajo positivo.
2	Impacto medio positivo.
3	Impacto alto positivo.

Tabla 4: Matriz de Impacto

4.1.1 IMPACTO TECNOLÓGICO

IMPACTO TECNOLÓGICO							
NIVELES DE IMPACTO / INDICADOR	**-3**	**-2**	**-1**	**0**	**1**	**2**	3
• Aceptación del sistema							X
• Uso de información útil							X
• Manejo de herramientas digitales.						X	
• Mejora de servicios						X	
TOTAL						4	6
							∑ = 10

$$\textbf{NI} = \frac{\mathbf{10}}{\mathbf{4}} = \mathbf{2.5} \approx \mathbf{3}$$

Nivel de impacto tecnológico = Alto positivo.

Tabla 5: Análisis de Impacto Tecnológico

Análisis:

- En cuanto a la aceptación del sistema se considera tendrá un nivel de impacto alto positivo, debido a que cubre las necesidades del usuario.

- Se estima que el nivel de uso de información útil sea alto positivo, porque el sistema facilitará el acceso a documentos específicos que contribuirán en la realización de investigaciones de acuerdo a las consultas efectuadas por los usuarios que ingresen al sistema.

- Para el manejo de herramientas digitales se prevé que tendrá un impacto medio positivo, porque ya no se realizará la búsqueda en estos documentos de la forma tradicional, si no que se podrá tener acceso a ellos desde donde se encuentre, tan solo ingresando la página de la PUCESE.

- El nivel de mejora de servicios tendrá un impacto medio positivo, debido a que el proceso de búsqueda se realizará de manera más ágil, lo que garantiza que los usuarios en este caso los investigadores tengan un mayor grado de satisfacción.

4.1.2 IMPACTO ECONÓMICO

IMPACTO ECONÓMICO							
NIVELES DE IMPACTO / INDICADOR	**-3**	**-2**	**-1**	**0**	**1**	**2**	3
• Software gratuito							X
• Reducción de tiempos							X
• Recurso tecnológico				X			
TOTAL				0			6

∑ = 6

$$\textbf{NI} = \frac{6}{3} = \textbf{2}$$

Nivel de impacto económico = Medio positivo

Tabla 6: Análisis de Impacto Económico

Análisis:

- En lo que se refiere al software gratuito, se estima tendrá un impacto alto positivo, ya que por motivos de la utilización de software libre la reducción en costos es inminente al no tener que invertir dinero en licencias.

- Al hacer más ágil el proceso de consulta de los documentos, implica considerablemente la reducción del tiempo empleado en ello, es por esto que se considera tenga un nivel de impacto alto positivo.

- En lo que respecta al recurso tecnológico tendrá un nivel de impacto nulo ya que para implementar el sistema no se requiere de la adquisición de algún equipo tecnológico, porque se puede utilizar los recursos tecnológicos ya existentes en la institución.

4.1.3 IMPACTO EDUCATIVO

IMPACTO EDUCATIVO							
NIVELES DE IMPACTO / INDICADOR	**-3**	**-2**	**-1**	**0**	**1**	**2**	3
• Investigación eficaz • Fácil uso del sistema						X X	
TOTAL						4	
							∑ = 4

$$\textbf{NI} = \frac{4}{2} = 2$$

Nivel de impacto educativo = Medio positivo

Tabla 7: Análisis de Impacto Educativo

Análisis:

- Para el nivel investigación eficaz se considera tenga un impacto medio positivo, porque los investigadores podrán consultar en documentos con información sustentada verídica que les serán de mucha utilidad y guía en el desarrollo su investigación.

- El sistema cuenta con una interfaz sencilla y fácil de manejar para los usuarios que ingresen a realizar una búsqueda, es por esto que se considera que el fácil uso del sistema tenga un nivel de impacto alto positivo.

4.1.4 IMPACTO GENERAL

IMPACTO GENERAL							
NIVELES DE IMPACTO / INDICADOR	**-3**	**-2**	**-1**	**0**	**1**	**2**	3
• Impacto Tecnológico							X
• Impacto Económico						X	
• Impacto Educativo						X	
TOTAL						4	3
							∑ = 7

$$\textbf{NI} = \frac{7}{3} = 2.33 \approx 2$$

Nivel de impacto general = Medio positivo

Tabla 8: Análisis de Impacto General

Análisis:

- Como se puede observar el ámbito tecnológico es el que más impacto tendrá, por lo que permitirá el fácil acceso a los documentos haciendo la búsqueda de manera más ágil, dejando así al usuario con todas sus necesidades cubiertas y satisfecho con el servicio.

- El ámbito económico obtuvo un impacto medio positivo porque se prevé no se invertirá ni en licencias de software ni en equipos tecnológicos, y además reducirá tiempo de búsqueda para los investigadores.

- En lo educativo el impacto será medio positivo, debido a que los usuarios sabrán cómo manejar el sistema y obtendrán información valiosa, útil y confiable, para enriquecer sus conocimientos en la investigación o proyectos que estén emprendiendo.

CAPÍTULO V
CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

- La aplicación de técnicas y algoritmos de búsqueda, permiten que el sistema obtenga eficiencia y pueda solucionar problemas en consultas, peticiones y en situaciones donde los datos que se deben mostrar no estén ordenados.

- El uso del sistema de búsqueda es de gran beneficio, debido a que proporciona los resultados deseados a los investigadores que buscan dar solución a problemáticas o simplemente despejar interrogantes, permitiendo el acceso a la información contenida en los documentos de tesis o proyectos de grado, con mayor libertad.

- Al realizar el sistema con herramientas de software libre, permite que la implementación del sistema no implique costos en cuanto a licencias, además el software podrá adaptarse a las necesidades del usuario, en caso de que ocurran modificaciones en los datos.

- Los investigadores tendrán acceso total a la información que proporciona la base de datos y a los documentos en formato .pdf, podrán hacer consultas con sosiego y teniendo en cuenta que la información es verídica.

5.2 RECOMENDACIONES

- Tratar de esclarecer la estructura de la base de datos de la PUCESE. Los nombres de las tablas y los campos son poco entendibles, lo que dificultó el trabajo de identificación de las mismas. Para ello se recomienda realizar un completo análisis y tratar de que los datos sean entendibles, para evitar futuras complicaciones.

- Proporcionar al usuario administrador la capacitación adecuada, para que el control y manejo de la información del sistema. de igual manera capacitar a los investigadores sobre las funcionalidades y ventajas del sistema, para que de esta manera ellos puedan replicar esta información.

- Inducir a los estudiantes a realizar trabajos investigativos, realizando consultas en fuentes verídicas de información, para que los resultados que alcancen sean los más óptimos en el análisis y recopilación de datos del proyecto que vayan a emprender.

GLOSARIO DE TÉRMINOS

A.

Aplicación Web.-

Herramientas que los usuarios pueden utilizar accediendo a un servidor web a través de Internet o de una intranet mediante un navegador. En otras palabras, es una aplicación software que se codifica en un lenguaje soportado por los navegadores web en la que se confía la ejecución al navegador.

B.

Base de Datos.-

Serie de datos organizados y relacionados entre sí, los cuales son recolectados y explotados por los sistemas de información de una empresa o negocio en particular.

Browser.-

Navegador. Aplicación para visualizar documentos WWW y navegar por Internet.

C.

Código.-

Es un sistema de signos y reglas para combinarlos, que por un lado es arbitrario y por otra parte debe de estar organizado de antemano.

H.

Hipertexto.-

Es el nombre que recibe el texto que en la pantalla de un dispositivo electrónico conduce a otro texto relacionado.

Host.-

Ordenador conectado a Internet. Ordenador en general. Literalmente anfitrión.

Html.-

(HyperTextMarkupLanguage) Lenguaje de Marcado de Hipertexto es el lenguaje de marcado predominante para la elaboración de páginas web.

I.

Ip.-

Internet Protocol – Protocolo Internet

Internet.-

Es un conjunto descentralizado de redes de comunicación interconectadas que utilizan la familia de protocolos TCP/IP, garantizando que las redes físicas heterogéneas que la componen funcionen como una red lógica única, de alcance mundial.

L.

Licencias.-

Contrato mediante el cual una persona recibe de otra el derecho de uso, de copia, de distribución, de estudio y de modificación, en el caso de Software Libre.

O.

Open Source.-

Es un término que describe partes de la licencia del movimiento por el software libre.

P.

Php.-

Lenguaje de programación interpretado, diseñado originalmente para la creación de páginas web dinámicas.

S.

Servidor.-

Una aplicación informática o programa que realiza algunas tareas en beneficio de otras aplicaciones. Algunos servicios habituales son los servicios de archivos, que permiten a los usuarios almacenar y acceder a los archivos de una computadora y los servicios de aplicaciones, que realizan tareas en beneficio directo del usuario final.

Software.-

Son los sistemas compuestos de aspectos intangibles y abstractos, como filosofías, políticas, directivas, programas, procedimientos, reglas y reglamentos, etc.

W.

World Wide Web. –

Comúnmente llamada wwww, La gran telaraña mundial

Referencias Bibliográficas

- Altova. (2014). *Altova*. Obtenido de Diagramas de objetos UML: http://www.altova.com/es/umodel/object-diagrams.html

- Amo, F. A., Martínez, L., & Segovia, F. (2005). *Introducción a la ingeniería del software.* Delta Publicaciones.

- Anzaldo, J. (15 de Noviembre de 2005). *Arquitectura de Información*. Obtenido de Arquitectura de Información: http://janzaldo.wordpress.com/2005/11/15/%C2%BFque-es-una-arquitectura-de-sistemas-de-informacion/

- Cebedo, A. (2010). Recursos informáticos para la detección del plagio académico. *Tejuelo, nº 8*, 8-26.

- Cedeño, A., & Rosso, P. (2009). Importancia de la reducción del espacio de búsqueda. *Procesamiento del Lenguaje Natural, núm. 43*.

- Cillero, M. (2009). *Diagrama de Flujo de Datos (DFD)*. Obtenido de Diagrama de Flujo de Datos (DFD): https://manuel.cillero.es/doc/metrica-3/tecnicas/diagrama-de-flujo-de-datos/

- Cobos, C., León, E., & Mendoza, M. (2010). Un algoritmo de búsqueda armónica para. *Rev. Fac. Ing. Univ. Antioquia N.° 55*.

- Días, N. (2006). *Tipos de Algoritmos de Búsqueda.* Obtenido de Tipos de Algoritmos de Búsqueda: http://artemisa.unicauca.edu.co/~nediaz/EDDI/cap02.htm

- Dorado, S. (2 de Noviembre de 2012). *Tecnólogo en Gestión Bibliográfica*. Obtenido de Tecnólogo en Gestión Bibliográfica: http://sandratecnologobiblioteca.blogspot.com/2012/11/ensayo-inventario-de-material.html

- Duarte, E. (16 de Marzo de 2013). *capacityacademy.com*. Obtenido de capacityacademy.com: http://blog.capacityacademy.com/2013/03/16/jquery-que-es-origenes-ventajas-desventajas/

- Ecuador Universitario. (13 de Noviembre de 2013). *EcuadorUniversitario.com*. Obtenido de SENESCYT anuncia nuevas políticas para el desarrollo de la educación superior: http://ecuadoruniversitario.com/noticias_destacadas/senescyt-anuncia-nuevas-politicas-para-el-desarrollo-de-la-educacion-superior/

- Fernández, H. (2011). *MOTORES DE BÚSQUEDA.* Obtenido de MOTORES DE BÚSQUEDA: http://especializacion.una.edu.ve/Internet/paginas/Lecturas/Fernandez.pdf

- Ferran, N., & Mario, P. (2011). *Búsqueda y recuperación de información.* Editorial UOC.
- Garcia, J. (2005). *Análisis y Modelado*. Obtenido de Análisis y Modelado: http://www.ingenierosoftware.com/analisisydiseno/uml.php
- Gonzáles, E. (2016). *aprenderaprogramar.com*. Obtenido de aprenderaprogramar.com: http://aprenderaprogramar.com/index.php?option=com_content&view=article&id=492:ique-es-php-y-ipara-que-sirve-un-potente-lenguaje-de-programacion-para-crear-paginas-web-cu00803b&catid=70:tutorial-basico-programador-web-php-desde-cero&Itemid=193
- Lamarca, M. (8 de Diciembre de 2013). *Hipertexto, el nuevo concepto de documento en la cultura de la imagen*. Obtenido de Hipertexto, el nuevo concepto de documento en la cultura de la imagen: http://www.hipertexto.info/documentos/robot_agent.htm
- Marín, A. (2008). Aplicación de dos nuevos algoritmos para agrupar. *Rev. Interam. Bibliot. Medellín (Colombia) Vol. 31 No. 1*.
- Márquez, J. (15 de Enero de 2009). *uaeh.edu.mx.* Recuperado el 9 de Febreo de 2015, de uaeh.edu.mx: http://www.uaeh.edu.mx/docencia/VI_Lectura/maestria/documentos/LECT43.pdf
- Martínez, F. (2004). *RECUPERACIÓN DE INFORMACIÓN: MODELOS, SISTEMAS Y EVALUACIÓN.* EL KIOSKO JMC.
- Microsoft. (2007). *Diagramas de clase UML: instrucciones*. Obtenido de Diagramas de clase UML: instrucciones: http://msdn.microsoft.com/es-ec/library/dd409416.aspx
- Ministerio del poder popular de cultura de Venezuela. (2008). *Inventario de Material Bibliográfico en la Bibliotecas Publicas.* Obtenido de Inventario de Material Bibliográfico en la Bibliotecas Publicas: http://www.bnv.gob.ve/pdf/bp0422008.pdf
- Navarra, P. L., & Martínez, J. Á. (2006). *Agentes inteligentes en la búsqueda y recuperación de información.* Barcelona: Planeta- UOC, S.L.
- Osmosis Latina. (7 de Septiembre de 2005). *Diagramas Uso-Caso.* Obtenido de Diagramas Uso-Caso: http://www.osmosislatina.com/lenguajes/uml/casos.htm
- Pedraz, A. (2003). *Etapas del proceso de investigación.* Obtenido de Etapas del proceso de investigación: http://webpersonal.uma.es/~jmpaez/websci/BLOQUEI/DocuI/Etapas.pdf
- Pérez, J. &. (2014). *Definición de repositorio* . Obtenido de Definición de repositorio : http://definicion.de/repositorio/
- Pinto, M. (13 de Abril de 2011). *Búsqueda y Recuperación de Información*. Obtenido de Búsqueda y Recuperación de Información: http://www.mariapinto.es/e-coms/recu_infor.htm

- PostgreSQL. (2 de Octubre de 2010). *Sobre PostgreSQL*. Obtenido de Sobre PostgreSQL: http://www.postgresql.org.es/sobre_postgresql
- PUCESE. (2014). *www.pucese.edu.ec*. Recuperado el 28 de Mayo de 2014, de www.pucese.edu.ec: http://www.pucese.edu.ec/
- Ril, Y., Toll, Y. d., & Fonseca, E. (2014). *Determinación de estilos de escritura para la detección de similitudes entre documentos digitales.* Barcelona: Universitat Oberta de Catalunya y University of New England.
- Romero, C., Castellá, J., & Viejo, A. (2014). *Distributed system for private web search with untrusted.* Tarragona (España): Elsevier B.V. All rights reserved.
- Salvador, M. F., Gupta, P., & Rosso, P. (2013). Análisis de similitud basado en grafos: Una nueva aproximación a la detección de plagio translingue. *Procesamiento del Lenguaje Natural, Revista nº 50*.
- Shirley. (Mayo de 2012). *Tipos de Diagramas*. Obtenido de http://ingenieriadesistemas-shirley.blogspot.com/2012/05/tipos-de-diagramas-uml.html
- Soluciones Integrales de Compras y Adquisiciones. (2014). *Importacia de hacer inventario*. Obtenido de http://www.califica.com.ec/noticias/la-importancia-de-hacer-inventario.html
- Storti, M., D'Elía, J., Paz, R., Lizandro, D., & Martín, P. (25 de Febrero de 2014). *Algoritmos y Estructuras de Datos.* Obtenido de Algoritmos de ordenamiento y busqueda: http://www.cimec.org.ar/~mstorti/aed/aednotes.pdf
- Tolasa, G., & Bordignon, F. (2008). *Introduccion a la Recuperación de Información.* Obtenido de http://eprints.rclis.org: http://eprints.rclis.org/12243/1/Introduccion-RI-v9f.pdf
- UM. (18 de Marzo de 2015). *Desarrollo de Aplicaciones Web*. Obtenido de Desarrollo de Aplicaciones Web: http://www.um.es/docencia/barzana/DAWEB/Desarrollo-de-aplicaciones-web-Xampp.html
- Universidad Interamericana para el Desarrollo (UNID). (2014). *Repositorios Digitales*. Obtenido de Repositorios Digitales: http://www.unid.edu.mx/tecnologiaeinnovacion/967-repositorios-digitales.html
- Universidad Uión Bolivariana. (2014). *Ingeniería de Software*. Obtenido de Ingeniería de Software: http://ingenieriadesoftware.mex.tl/52753_XP---Extreme-Programing.html
- Universisas de Mursia. (2011). *Metodologías de Desarrollo de Software.* Obtenido de Metodologías de Desarrollo de Software: http://www.um.es/docencia/barzana/IAGP/Iagp2.html

- Verri, L. (2012). *Capítulo 2: Exito en proyectos de capital.* Obtenido de Exito en proyectos de capital: http://www.mantenimientomundial.com/sites/mm/notas/Concepto%20de%20IPA%20y%20PMI.pdf

- Visual Studio. (2014). *Visual Studio*. Obtenido de Visual Studio: http://www.visualstudio.com/es-es/visual-studio-homepage-vs.aspx

• Trabajos citados

- Altova. (2014). *Altova*. Obtenido de Diagramas de objetos UML: http://www.altova.com/es/umodel/object-diagrams.html

- Amo, F. A., Martínez, L., & Segovia, F. (2005). *Introducción a la ingeniería del software.* Delta Publicaciones.

- Anzaldo, J. (15 de Noviembre de 2005). *Arquitectura de Información*. Obtenido de Arquitectura de Información: http://janzaldo.wordpress.com/2005/11/15/%C2%BFque-es-una-arquitectura-de-sistemas-de-informacion/

- Cebedo, A. (2010). Recursos informáticos para la detección del plagio académico. *Tejuelo, nº 8*, 8-26.

- Cedeño, A., & Rosso, P. (2009). Importancia de la reducción del espacio de búsqueda. *Procesamiento del Lenguaje Natural, núm. 43*.

- Cillero, M. (2009). *Diagrama de Flujo de Datos (DFD)*. Obtenido de Diagrama de Flujo de Datos (DFD): https://manuel.cillero.es/doc/metrica-3/tecnicas/diagrama-de-flujo-de-datos/

- Cobos, C., León, E., & Mendoza, M. (2010). Un algoritmo de búsqueda armónica para. *Rev. Fac. Ing. Univ. Antioquia N.° 55*.

- Días, N. (2006). *Tipos de Algoritmos de Búsqueda.* Obtenido de Tipos de Algoritmos de Búsqueda: http://artemisa.unicauca.edu.co/~nediaz/EDDI/cap02.htm

- Dorado, S. (2 de Noviembre de 2012). *Tecnólogo en Gestión Bibliográfica*. Obtenido de Tecnólogo en Gestión Bibliográfica: http://sandratecnologobiblioteca.blogspot.com/2012/11/ensayo-inventario-de-material.html

- Duarte, E. (16 de Marzo de 2013). *capacityacademy.com*. Obtenido de capacityacademy.com: http://blog.capacityacademy.com/2013/03/16/jquery-que-es-origenes-ventajas-desventajas/

- Ecuador Universitario. (13 de Noviembre de 2013). *EcuadorUniversitario.com*. Obtenido de SENESCYT anuncia nuevas políticas para el desarrollo de la educación superior: http://ecuadoruniversitario.com/noticias_destacadas/senescyt-anuncia-nuevas-politicas-para-el-desarrollo-de-la-educacion-superior/

- Fernández, H. (2011). *MOTORES DE BÚSQUEDA.* Obtenido de MOTORES DE BÚSQUEDA: http://especializacion.una.edu.ve/Internet/paginas/Lecturas/Fernandez.pdf

- Ferran, N., & Mario, P. (2011). *Búsqueda y recuperación de información.* Editorial UOC.

- Garcia, J. (2005). *Análisis y Modelado*. Obtenido de Análisis y Modelado: http://www.ingenierosoftware.com/analisisydiseno/uml.php

- Gonzáles, E. (2016). *aprenderaprogramar.com*. Obtenido de aprenderaprogramar.com: http://aprenderaprogramar.com/index.php?option=com_content&view=article&id=492:ique-es-php-y-ipara-que-sirve-un-potente-lenguaje-de-programacion-para-crear-paginas-web-cu00803b&catid=70:tutorial-basico-programador-web-php-desde-cero&Itemid=193

- Lamarca, M. (8 de Diciembre de 2013). *Hipertexto, el nuevo concepto de documento en la cultura de la imagen*. Obtenido de Hipertexto, el nuevo concepto de documento en la cultura de la imagen: http://www.hipertexto.info/documentos/robot_agent.htm

- Marín, A. (2008). Aplicación de dos nuevos algoritmos para agrupar. *Rev. Interam. Bibliot. Medellín (Colombia) Vol. 31 No. 1*.

- Márquez, J. (15 de Enero de 2009). *uaeh.edu.mx.* Recuperado el 9 de Febreo de 2015, de uaeh.edu.mx: http://www.uaeh.edu.mx/docencia/VI_Lectura/maestria/documentos/LECT43.pdf

- Martínez, F. (2004). *RECUPERACIÓN DE INFORMACIÓN: MODELOS, SISTEMAS Y EVALUACIÓN.* EL KIOSKO JMC.

- Microsoft. (2007). *Diagramas de clase UML: instrucciones*. Obtenido de Diagramas de clase UML: instrucciones: http://msdn.microsoft.com/es-ec/library/dd409416.aspx

- Ministerio del poder popular de cultura de Venezuela. (2008). *Inventario de Material Bibliográfico en la Bibliotecas Publicas.* Obtenido de Inventario de Material Bibliográfico en la Bibliotecas Publicas: http://www.bnv.gob.ve/pdf/bp0422008.pdf

- Navarra, P. L., & Martínez, J. Á. (2006). *Agentes inteligentes en la búsqueda y recuperación de información.* Barcelona: Planeta- UOC, S.L.

- Osmosis Latina. (7 de Septiembre de 2005). *Diagramas Uso-Caso.* Obtenido de Diagramas Uso-Caso: http://www.osmosislatina.com/lenguajes/uml/casos.htm

- Pedraz, A. (2003). *Etapas del proceso de investigación.* Obtenido de Etapas del proceso de investigación: http://webpersonal.uma.es/~jmpaez/websci/BLOQUEI/DocuI/Etapas.pdf

- Pérez, J. &. (2014). *Definición de repositorio* . Obtenido de Definición de repositorio : http://definicion.de/repositorio/

- Pinto, M. (13 de Abril de 2011). *Búsqueda y Recuperación de Información*. Obtenido de Búsqueda y Recuperación de Información: http://www.mariapinto.es/e-coms/recu_infor.htm

- PostgreSQL. (2 de Octubre de 2010). *Sobre PostgreSQL*. Obtenido de Sobre PostgreSQL: http://www.postgresql.org.es/sobre_postgresql

- PUCESE. (2014). *www.pucese.edu.ec*. Recuperado el 28 de Mayo de 2014, de www.pucese.edu.ec: http://www.pucese.edu.ec/

- Ril, Y., Toll, Y. d., & Fonseca, E. (2014). *Determinación de estilos de escritura para la detección de similitudes entre documentos digitales.* Barcelona: Universitat Oberta de Catalunya y University of New England.

- Romero, C., Castellá, J., & Viejo, A. (2014). *Distributed system for private web search with untrusted.* Tarragona (España): Elsevier B.V. All rights reserved.

- Salvador, M. F., Gupta, P., & Rosso, P. (2013). Análisis de similitud basado en grafos: Una nueva aproximación a la detección de plagio translingue. *Procesamiento del Lenguaje Natural, Revista nº 50*.

- Shirley. (Mayo de 2012). *Tipos de Diagramas*. Obtenido de http://ingenieriadesistemas-shirley.blogspot.com/2012/05/tipos-de-diagramas-uml.html

- Soluciones Integrales de Compras y Adquisiciones. (2014). *Importacia de hacer inventario*. Obtenido de http://www.califica.com.ec/noticias/la-importancia-de-hacer-inventario.html

- Storti, M., D'Elía, J., Paz, R., Lizandro, D., & Martín, P. (25 de Febrero de 2014). *Algoritmos y Estructuras de Datos.* Obtenido de Algoritmos de ordenamiento y busqueda: http://www.cimec.org.ar/~mstorti/aed/aednotes.pdf

- Tolasa, G., & Bordignon, F. (2008). *Introduccion a la Recuperación de Información.* Obtenido de http://eprints.rclis.org: http://eprints.rclis.org/12243/1/Introduccion-RI-v9f.pdf

- UM. (18 de Marzo de 2015). *Desarrollo de Aplicaciones Web*. Obtenido de Desarrollo de Aplicaciones Web: http://www.um.es/docencia/barzana/DAWEB/Desarrollo-de-aplicaciones-web-Xampp.html

- Universidad Interamericana para el Desarrollo (UNID). (2014). *Repositorios Digitales*. Obtenido de Repositorios Digitales: http://www.unid.edu.mx/tecnologiaeinnovacion/967-repositorios-digitales.html

- Universidad Uión Bolivariana. (2014). *Ingeniería de Software*. Obtenido de Ingeniería de Software: http://ingenieriadesoftware.mex.tl/52753_XP---Extreme-Programing.html

- Universisas de Mursia. (2011). *Metodologías de Desarrollo de Software.* Obtenido de Metodologías de Desarrollo de Software: http://www.um.es/docencia/barzana/IAGP/Iagp2.html

- Verri, L. (2012). *Capítulo 2: Exito en proyectos de capital.* Obtenido de Exito en proyectos de capital: http://www.mantenimientomundial.com/sites/mm/notas/Concepto%20de%20IPA%20y%20PMI.pdf

- Visual Studio. (2014). *Visual Studio*. Obtenido de Visual Studio: http://www.visualstudio.com/es-es/visual-studio-homepage-vs.aspx

ANEXOS

ANEXO 1: Guía de Encuesta

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR SEDE ESMERALDAS
ESCUELA DE INGENIERIA EN SISTEMAS Y COMPUTACIÓN

ENCUESTA DIRIGIDA A LOS ESTUDIANTES, DOCENTES Y MIENBROS DEL CID DE LA PUCESE

OBJETIVO: *Identificar los requerimientos y necesidades de los investigadores con respecto a la búsqueda y recuperación de contenidos de los informes digitalizados de proyectos o tesis de grado de la PUCESE.*

INSTRUCCIONES: *Conteste las preguntas marcando con una x según corresponda.*

1. **¿Utiliza la biblioteca para recopilar información de investigaciones desarrolladas en proyectos o tesis de grado de la PUCESE?**

   Si ( ) No ( )

2. **¿Con que frecuencia realiza usted investigaciones en estos documentos?**

   Diario ( )
   Tres veces por semana ( )
   Semanal ( )
   Mensual ( )
   Semestral ( )

3. **¿Cómo realiza usted la búsqueda de informes de tesis o proyectos de grado?**

   A través del sistema de biblioteca virtual ( )
   A través del bibliotecario encargado ( )

4. **¿Cuál es el motivo de consulta de estos documentos?**

   Despejar inquietudes ( )
   Obtener guías ( )
   Recopilar información de temas relacionados ( )
   Identificar su estructura o modelos de desarrollo ( )

5. **¿Cuáles son los inconvenientes que tiene al momento de solicitar un informe de tesis de grado?**

   Tiempo requerido para la búsqueda ( )
   Falta de parámetros de búsqueda ( )

Ineficiencia en los resultados de búsqueda ( )

6. **¿Cuánto tiempo utiliza el encargado de biblioteca para realizar la búsqueda de los informes de tesis o proyectos de grado, en el proceso de préstamo de estos documentos?**

5 minutos ( )
10 minutos ( )
15 minutos ( )
20 minutos ( )
Más de 20 minutos ( )

7. **¿Cómo califica el servicio de préstamo de informes de proyectos de grado en la PUCESE?**

Bueno ( )
Malo ( )
Regular ( )
Satisfactorio ( )
Muy satisfactorio ( )

8. **Seleccione los factores que impiden la búsqueda de contenidos en tesis o proyectos de grado.**

Poca accesibilidad ( )
Tiempo requerido para búsqueda exhaustiva ( )
Poco acceso al documento ( )

9. **¿Cuáles son los parámetros de búsqueda de contenidos a los que le gustaría acceder cuando solicite dicho documento?**

Titulo ( )
Autor ( )
Año ( )
Escuela ( )
Líneas de investigación ( )
Palabras claves ( )
Bibliografía ( )
Asesor ( )
Otro _____

Gracias por su colaboración

ANEXO 2: Guía de Entrevista al encargado del Departamento de Sistemas.

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR SEDE ESMERALDAS
ESCUELA DE INGENIERIA EN SISTEMAS Y COMPUTACIÓN

GUÍA PARA ENTREVISTA AL ENCARGADO DEL DEPARTAMENTO DE SISTEMAS DE LA PUCESE

Tema de Disertación de Grado: "*Sistema de búsqueda y recuperación de documentos, aplicado a los informes de tesis y proyectos de grado digitalizados de la Pontificia Universidad Católica del Ecuador Sede Esmeraldas".*

Objetivo: Identificarlos sistemas relacionados con el manejo y almacenamiento de la información correspondiente a los informes de proyectos o tesis de grado de la PUCESE.

1. **¿Qué sistema o sistemas utiliza el departamento de Biblioteca y quien es el encargado de darle mantenimiento?, ¿cuál es su funcionalidad?, ¿quién es el responsable de darle mantenimiento?**

   _____
   _____
   _____
   _____

2. **¿En qué Base de Datos se guarda la información el departamento de Biblioteca?**

   _____
   _____
   _____
   _____

3. **¿De qué Base de Datos arrastra información el sistema web de Biblioteca Virtual?**

   _____
   _____
   _____
   _____

4. **¿En qué Base de Datos se encuentran guardados los informes de proyectos o tesis de grado digitalizados y quien es el responsable de registrarlos?**

   _____
   _____

_____
_____

5. **¿En qué servidor se alojan todos estos datos?**

_____
_____
_____
_____

6. **¿Qué sistema o sistemas utiliza el departamento de Secretaría General y quién es el encargado de darle mantenimiento?, ¿cuál es su funcionalidad?, ¿quién es el responsable de darle mantenimiento?**

_____
_____
_____
_____

7. **¿En qué Base de Datos se guarda la información el departamento de Secretaría General?**

_____
_____
_____
_____

8. **¿La universidad está en un proceso de unificación de sus bases de datos, en qué Gestor de Base de Datos se está desarrollando?**

_____
_____
_____
_____

9. **¿En qué porcentaje se encuentra este proceso y quién está a cargo?**

_____
_____
_____
_____

ANEXO 3: Guía de Entrevista al encargado de Biblioteca de la PUCESE.

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR SEDE ESMERALDAS
ESCUELA DE INGENIERIA EN SISTEMAS Y COMPUTACIÓN

GUÍA PARA ENTREVISTA AL ENCARGADO DE LA BIBLIOTECA DE LA PUCESE

Tema de Disertación de Grado: "*Sistema de búsqueda y recuperación de documentos, aplicado a los informes de tesis y proyectos de grado digitalizados de la Pontificia Universidad Católica del Ecuador Sede Esmeraldas".*

Objetivo: Identificarlos procesos relacionados con la administración de los informes de proyectos o tesis de grado de la PUCESE.

1. **¿Qué funciones desempeña la biblioteca de la PUCESE?**
   _____
   _____
   _____
   _____

2. **¿Cuántas personas trabajan en la biblioteca y sus respectivos cargos?**
   _____
   _____
   _____
   _____

3. **¿Cuáles son las políticas con respecto a la codificación, registro, archivo y préstamo de los informes de tesis?**
   _____
   _____
   _____
   _____

4. **¿Las políticas en cuanto al préstamo de los informes de tesis o proyectos de grado son cumplidas y respetadas por los solicitantes?¿Cuales son inconvenientes y con qué frecuencia se presentan?**
   _____
   _____
   _____
   _____

5. **¿Cómo se codifican y se registran los informes de tesis?**

_____

_____

_____

_____

6. **¿Cómo se lleva el control del préstamo de los informes de tesis?**

_____

_____

_____

7. **¿Con que frecuencia acuden investigadores a solicitar los informes de tesis o proyectos de grado?**

_____

_____

_____

8. **¿Cuánto tiempo se requiere para realizar la búsqueda de los informes de tesis o proyectos de grado, en el proceso de préstamos?**

_____

_____

_____

_____

9. **¿Qué inconvenientes se presentan con respecto al manejo de los informes de tesis?**

_____

_____

_____

_____

10. **¿Está en marcha el proceso de digitalización de los informes de tesis, en qué nivel de desarrollo se encuentra?, ¿cuál es el tiempo estimado para su finalización?, ¿quién está a cargo?**

_____

_____

_____

_____