

**Universidad de Cienfuegos, facultad de Ingeniería, carrera
de Ingeniería Informática**

**BMETRIC, SISTEMA INFORMÁTICO PARA VALORAR EL
PROGRESO ACTUAL DE UN CAMPO DE INVESTIGACIÓN**

Trabajo de diploma para optar por el título de Ingeniería en Informática

Autor: Eduardo Portal Geroy.

Tutores: Dr. Yuniol Álvarez Betancourt

Ing. Luis Enrique García Hernández.

Cienfuegos, Cuba, 2015

DECLARACIÓN DE AUTORÍA

Declaro que soy el único autor de este trabajo y autorizo al departamento de informática de la Universidad de Cienfuegos para que haga el uso que estime pertinente con este trabajo.

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

(Si procede)

Nombre completo del primer autor

Nombre completo del segundo autor

(Si procede)

Nombre completo del primer tutor

Nombre completo del segundo tutor

Agradecimientos

Agradezco en primer lugar a Dios por otorgarme esta vocación y haberme permitido estudiarla. *Ab imo pectore, Deo gratias.*

Agradezco a mi tutor Yuniol Álvarez, por haberme escogido para esta tarea en la que él siempre tuvo mucho empeño. Felicidades para Ud. Y su familia.

Agradezco a mis padres, por haberme soportado en todos los momentos de estrés y ayudado incondicionalmente.

Agradezco enormemente a mi archi-compañero-amigo-hermano Yinier Rojas Martínez, que el pobre, lleva cargando conmigo ya unos cuantos años. Para ti hermano lo mejor de este mundo.

Agradezco especialmente a mis hermanos de la *Inhospita trinchera*, David, Idi Bárbara, Lázaro, Darovis, Yinier y especialmente a María Amelia, gracias por mantener el catolicismo vivo en esta isla, por acompañarme y apoyarme, le agradezco a todos.

Agradezco profundamente a quien con su ejemplo ha sabido desensoberbecer mi espíritu desde que le conocí y demostrado con la humildad que le caracteriza, que el entendimiento le pertenece auténticamente a quien no especula con él. Un millón de gracias a Eduardo Concepción Morales.

Agradezco a mi tutor Luis Enrique García, ya que por ser casi el único profesor con experiencia en la industria de software, me enseñó con mansedumbre mucho en poco tiempo.

Agradecer a María Amelia por ser compañera y apoyo más importante, por darme la paz que tanto necesito en días como los que acontecen, gracias amor mío. En general a la familia Molina por acogerme con calidez.

Agradezco profundamente a todos los que me acompañaron desde el primer momento y hasta el final.

Resumen

El presente trabajo se desarrolló en la Universidad de Cienfuegos y aborda las tecnologías informáticas y de las comunicaciones en el campo de la bibliometría. Surge de la necesidad de valorar el estado actual en un área del conocimiento específica, de actualizarse en relación al comportamiento de esta o simplemente de analizar las tendencias de una disciplina de forma coherente y metodológicamente viable. “bMetric, Sistema informático para valorar el progreso actual de un campo de investigación” es un software que ofrece criterios objetivos a partir de registros bibliográficos contenidos en ficheros con formato RIS (Sistemas de Información de Investigación, en inglés, *Research Information Systems*) arrojados por gestores bibliográficos (e. g. *WoS*, *Scopus*, *Zotero*, *EndNote*, etc.), brindando diversos reportes y datos estadísticos que apoyan un estudio bibliométrico realizado en un área del conocimiento durante un periodo de tiempo determinado.

Actualmente las revisiones bibliográficas son llevadas a cabo muchas veces de forma manual y engorrosa, dando lugar a errores humanos propios de procesos de esta índole; la puesta en práctica de “bMetric” agiliza la tarea, sustituye los procesos manuales por automáticos o asistidos y proporciona una base de información para analizar a posteriori y generar conclusiones de alto valor agregado y aceptadas por la comunidad científica. Además, el sistema guía al usuario a través de una metodología propuesta en el 2013 que en gran medida asegura la calidad de los resultados.

Palabras claves: bibliometría, cienciometría, RIS, sistema web.

Abstract

The current work was developed at Universidad de Cienfuegos rounds about IT in the field of bibliometrics. Comes as a response to the need of evaluate the current state of science on a knowledge area, update in relation to its behavior or simply analyze trends in a right way and methodologically viable. “bMetric, Sistema informático para valorar el progreso actual de un campo de investigación” is a piece of software that offers objective criteria using bibliographic records contained in files with RIS (stands for Research Information System) format, used for many popular bibliographic managements software’s (e. g. *WoS*, *Scopus*, *Zotero*, *EndNote*, etc.). Offers several reports and statistic data that support a bibliometric study realized in a knowledge area in a defined period of time.

Currently bibliographic revisions are made many times in a handy way, giving place to human mistakes; the use of “bMetric” boost the task, avoid manual process using instead automatic or assisted operations, provides an information base for a analyzing a *posteriori* and generate conclusions with a high added value and accepted for the scientific community. Also, the system guides the user through a methodology proposed at 2013 that in great measure assures the quality of the results.

Keywords: bibliometric, scientometric, RIS, web system.

Índice

INTRODUCCIÓN	10
CAPÍTULO 1 FUNDAMENTOS TEÓRICOS	14
1.1 <i>INTRODUCCIÓN</i>	14
1.2 <i>OBJETIVOS ESTRATÉGICOS DE LA ACADEMIA</i>	14
1.2.1 <i>Investigación y producción científica</i>	14
1.2.2 <i>Análisis crítico de la valoración de un área del conocimiento</i>	14
1.3 <i>PROCESOS OBJETO DE AUTOMATIZACIÓN</i>	14
1.3.1 <i>Desambiguación de datos</i>	15
1.3.2 <i>Recuento de publicaciones</i>	15
1.3.3 <i>Recuento de publicaciones por tipo de publicación</i>	15
1.3.4 <i>Recuento de citas por autor</i>	15
1.3.5 <i>Recuento de citas por publicación</i>	16
1.3.6 <i>Recuento de colaboraciones</i>	16
1.3.7 <i>Recuento de publicación por año</i>	17
1.3.8 <i>Red de colaboración</i>	17
1.4 <i>SISTEMAS AUTOMATIZADOS EXISTENTES VINCULADOS AL CAMPO DE ACCIÓN</i>	18
1.4.1 <i>Web of Knowledge (Web del conocimiento, WOK por sus siglas en inglés)</i>	18
1.4.2 <i>SciVerse Scopus</i>	18
1.4.3 <i>Google Scholar (Google Académico)</i>	19
1.5 <i>METODOLOGÍAS DE DESARROLLO DE SOFTWARE</i>	19
1.5.1 <i>El Proceso Unificado de Desarrollo de Software (RUP)</i>	19
1.6 <i>PATRONES ARQUITECTÓNICOS</i>	19
1.6.1 <i>Modelo Vista Controlador (MVC)</i>	19
1.7 <i>SISTEMA GESTOR DE BASES DE DATOS</i>	20
1.7.1 <i>Base de Datos (BD)</i>	20
1.7.2 <i>Sistema Gestor de Bases de Datos (SGBD)</i>	20
1.7.3 <i>MySQL</i>	20
1.8 <i>LENGUAJES</i>	20
1.8.1 <i>Unified Modeling Language (UML)</i>	20
1.8.2 <i>Lenguaje de programación PHP</i>	20
1.8.3 <i>Lenguaje de programación JavaScript</i>	21
1.8.4 <i>HTML 5</i>	21
1.8.5 <i>CSS 3</i>	21
1.9 <i>FRAMEWORKS</i>	21
1.9.1 <i>Yii framework</i>	21

1.9.2	<i>jQuery</i>	21
1.9.3	<i>Sigma</i>	21
1.9.4	<i>Bootstrap</i>	22
1.10	HERRAMIENTAS UTILIZADAS	22
1.10.1	<i>Formato de fichero RIS</i>	22
1.10.2	<i>NetBeans IDE</i>	22
1.10.3	<i>Firefox</i>	22
1.10.4	<i>Firebug</i>	22
1.10.5	<i>MySQL Workbench</i>	22
1.11	CONCLUSIONES	22
CAPÍTULO 2 MODELO DEL NEGOCIO.....		24
2.1	INTRODUCCIÓN	24
2.2	MODELO DEL NEGOCIO	24
2.2.1	<i>Adquisición de datos</i>	24
2.2.2	<i>Desambiguación de datos</i>	24
2.2.3	<i>Análisis a priori</i>	24
2.2.4	<i>Presentación</i>	24
2.2.5	<i>Análisis a posteriori</i>	25
2.3	REGLAS DEL NEGOCIO A CONSIDERAR	25
2.4	TRABAJADOR DEL NEGOCIO	25
2.5	DIAGRAMA DE CASOS DE USO DEL NEGOCIO	26
2.6	DIAGRAMAS DE ACTIVIDAD	26
2.7	MODELO DE OBJETOS DEL NEGOCIO	28
2.8	CONCLUSIONES.....	28
CAPÍTULO 3 REQUISITOS		29
3.1	ACTORES DEL SISTEMA A AUTOMATIZAR.....	29
3.2	DEFINICIÓN DE LOS REQUISITOS FUNCIONALES	29
3.3	DIAGRAMA DE CASOS DE USO DEL SISTEMA A AUTOMATIZAR.....	31
3.4	DEFINICIÓN DE LOS REQUISITOS NO FUNCIONALES	32
3.5	DESCRIPCIÓN DE LOS CASOS DE USO	32
3.6	CONCLUSIONES.....	37
CAPÍTULO 4 DESCRIPCIÓN DE LA SOLUCIÓN PROPUESTA		38
4.1	INTRODUCCIÓN	38
4.2	DIAGRAMA DE CLASES DEL DISEÑO	38
4.3	PRINCIPIOS DE DISEÑO	38

4.3.1	<i>Patrones de diseño</i>	38
4.3.2	<i>Interfaz de usuario</i>	39
4.3.3	<i>Formato de salida de los reportes</i>	39
4.4	<i>TRATAMIENTO DE ERRORES</i>	39
4.5	<i>DISEÑO DE LA BASE DE DATOS</i>	40
4.5.1	<i>Modelo físico de datos</i>	40
4.6	<i>DIAGRAMA DE DESPLIEGUE</i>	41
4.7	<i>CONCLUSIONES</i>	41
CAPÍTULO 5 ESTUDIO DE FACTIBILIDAD		42
5.1	<i>INTRODUCCIÓN</i>	42
5.2	<i>PLANIFICACIÓN BASADA EN CASOS DE USO</i>	42
5.3	<i>BENEFICIOS TANGIBLES E INTANGIBLES</i>	45
5.4	<i>CONCLUSIONES</i>	45
CONCLUSIONES		46
RECOMENDACIONES		47
REFERENCIAS BIBLIOGRÁFICAS		48
GLOSARIO DE TÉRMINOS		49

Índice de tablas

Tabla 1. Ejemplo de resultado de un recuento de publicaciones	15
Tabla 2. Ejemplo de resultado de un recuento de citas	16
Tabla 3. Ejemplo de resultado de un recuento de citas por publicación	16
Tabla 4. Ejemplo de resultado de un recuento de publicaciones.	17
Tabla 5. Ejemplo de recuento de publicaciones por año.....	17
Tabla 6. Clasificación de los casos de uso	42
Tabla 7. Conteo de las clasificaciones de los casos de uso.....	42
Tabla 8. Relación de clasificación y peso	43
Tabla 9. Relación de complejidad técnica.	44
Tabla 10. Relación de factor de ambiente.....	44

Índice de figuras

Ilustración 1. Ejemplo de red de colaboración entre autores	18
Ilustración 2. Diagrama representativo del patrón MVC.....	20
Ilustración 3. Diagrama de casos de uso del negocio.....	26
Ilustración 5 Diagrama de actividad del negocio, desambiguar datos.	26
Ilustración 6. Diagrama de actividad del negocio, analizar datos.....	27
Ilustración 7. Diagrama de actividad del negocio, reportar datos.....	27
Ilustración 8. Diagrama de actividad del negocio, analizar reportes.	28
Ilustración 9. Modelo de objetos del negocio.	28
Ilustración 10. Diagrama de casos de uso del sistema	31
Ilustración 11. Diagrama de clases del diseño	38

Introducción

Numerosos teóricos describen la sociedad actual como la sociedad de la información, no de forma arbitraria, pues el paradigma del conocimiento tecnológico es el imperante, ocasionando abismales cambios en la forma de observar y comprender la sociedad en toda su composición. De forma similar sucede con el conocimiento científico, tanto es que en un periodo de 10 años la literatura científica crece de manera que iguala a toda la existente, aunque este fenómeno depende en gran parte del área del conocimiento a la que se refiera[1]. En medio de este florecimiento para evitar lo que Cornella define como “infoxicación”[2] surge la bibliometría que definida por sus actividades comprende la:

- Aplicación de análisis estadísticos para estudiar las características del uso y creación de documentos.
- Estudio cuantitativo de la producción de documentos como se refleja en las bibliografías.
- Aplicación de métodos matemáticos y estadísticos al estudio del uso que se hace de los libros y otros soportes dentro y entre los sistemas de bibliotecas.
- Estudio cuantitativo de las unidades físicas publicadas, o de las unidades bibliográficas, o de sus sustitutos.[3]

Una vez definida la bibliometría y aceptando ciencia como todo conocimiento racional, la cienciometría no es más que la aplicación de técnicas bibliométricas a la literatura de carácter científico[3]. Supliendo la necesidad de datos que genera esta rama del saber (cienciometría) nacen bases de datos extensas constituidas por un conjunto de registros con información bibliográfica (autor, título de la contribución, de la publicación, fecha de la publicación, editorial, etc.), almacenado y gestionado mediante sistemas informáticos. Muchas de estas bases de datos contienen descriptores, palabras claves y resúmenes. Y unas pocas contienen citas (índices de citas)[4]. Durante muchos años los índices de citas disponibles fueron:

1. *Science Citation Index (SCI, Índice de Citas Científicas)*
2. *Social Sciences Citation Index (SSCI, Índice de Citas de Ciencias Sociales)*
3. *Arts and Humanities Citation Index (AHCI, Índice de Citas de Artes y Humanidades)*

Elaborados por el Instituto para la Información Científica (*Institute for Scientific Information, ISI* por sus siglas en inglés) fundado por Eugene Garfield en 1960. En la actualidad estos índices están integrados en el *Web of Science (WOS)*, suministrado por Thomson Reuters a través de la plataforma *Web of Knowledge (WOK)*. El WOS indexa unas 9.300 de las revistas de investigación más prestigiosas. De hecho, que una publicación periódica esté indexada en esta base de datos se considera internacionalmente un criterio significativo de calidad. Otro de los productos ofrecidos por el WOK es el Reporte de Citas a Revistas (*Journal Citation Reports*) que evalúa el

impacto de las revistas a partir de las citas. Existen dos ediciones, la Edición Científica (*Science Edition*), que ofrece datos sobre el impacto de 5.900 publicaciones periódicas en ciencias; y el Edición Ciencias Sociales (*Social Sciences Edition*) que ofrece datos sobre 1.700 publicaciones en ciencias sociales y algunas de humanidades[5]. El año 2002 *Elsevier* comenzó a publicar *Scopus* oficialmente *SciVerse Scopus*, una base de datos bibliográfica de citas con voluntad de ser mucho más exhaustiva que el WOS. En la actualidad indexa unos 19 500 títulos. Sin embargo mientras que el SCI recoge documentación que se remonta a 1900 y el SSCI desde 1956, Scopus solamente recoge resúmenes desde 1960 y citas desde 1996[6]. A finales de 2004 *Google Inc* lanzó *Google Scholar* o *Google Académico*. Fiel a la filosofía de *Google*, a diferencia de las bases de datos tradicionales, no vacía los contenidos de revistas sino que explora sistemáticamente la Web. Pero en este producto convergen dos servicios, un buscador de publicaciones científicas y un índice de citas que permita conocer el impacto de los trabajos publicados. Sin embargo, el uso del rastreo automático indiscriminado conlleva una importante dosis de errores. Por otra parte se indiza una amplia variedad de tipos de documentos, algunos de los cuales distan de lo que se considera académico (guías académicas, documentos administrativos, bibliografía de asignaturas, libros de divulgación, etc.). Este punto es importante ya que no es lo mismo ser citado en un documento científico que en otro tipo de documento. Más recientemente *Google Inc* ha incorporado dos nuevos productos *Google Scholar Citations (GSC)* y *Google Scholar Metrics (GSM)*. El primero recopila la producción científica de un investigador y la ofrece agregada en una página web, añadiendo información sobre el número de citas de cada referencia. En cambio *GSM* ofrece el impacto de las revistas científicas a partir de los recuentos de citas[7][4].

Esta “abundancia” de información lleva a varios científicos a realizar manuales y engorrosas investigaciones, dando lugar a errores humanos propios de procesos *ad libitum*, al no seguir una metodología que garantice la calidad y veracidad de los resultados obtenidos.

Es por ello que teniendo en cuenta todo lo planteado anteriormente se define como:

Problema: ¿Cómo valorar el progreso actual de un campo de investigación determinado a partir de registros bibliográficos?

Objeto de estudio: Las tecnologías de la información y la comunicación en el campo de la bibliometría.

Campo de acción: El proceso de valoración del progreso actual de un campo de investigación determinado utilizando bibliometría.

Idea a defender: Con la utilización de sistema informático basado en una metodología bibliométrica, se facilitará y mejorará el proceso de valoración del progreso actual de un campo de investigación determinado.

Objetivo general: Desarrollar un sistema informático basado en una metodología bibliométrica para valorar el progreso actual de un campo de investigación determinado a partir de registros bibliográficos.

Objetivos específicos:

1. Analizar las tecnologías y herramientas para el desarrollo de sistemas web.
2. Diseñar e implementar un sistema informático para valorar el progreso actual de un campo de investigación determinado, basado en formatos estándares de referencias bibliográficas.
3. Validar el sistema.

Tareas a desarrollar:

1. Realizar una revisión bibliográfica sobre cienciometría, bibliometría, sus leyes e índices más utilizados.
2. Realizar una revisión bibliográfica y en internet sobre desarrollo de sistemas web utilizando el lenguaje de programación PHP.
3. Realizar una revisión bibliográfica y en internet sobre librerías de representación de datos en forma gráfica, especialmente en formatos estándares.
4. Diseñar un sistema web capaz de ofrecer los servicios regulares que sirven este tipo de sistemas y permita valorar el progreso actual de un campo de investigación determinado a partir de ficheros de formato estándar.
5. Implementar el sistema informático propuesto.
6. Analizar la metodología propuesta en *An overview of Iris recognition* por Yuniol Álvarez Betancourt en 2013 y orientar el sistema a que se ajuste a esta.
7. Realizar pruebas con distintas fuentes de datos para comprobar la correcta implementación del sistema.

Aporte práctico: Un sistema informático basado en una metodología bibliométrica para valorar el progreso actual de un campo de investigación determinado a partir de registros bibliográficos.

Resumen capitular

Para la estructura de este trabajo se proponen 5 capítulos. Cada uno contiene un epígrafe introductorio que comenta de forma general lo que se estará tratando en el capítulo. En cada capítulo, existe un epígrafe de conclusiones parciales, donde se sintetizan los aspectos más puntuales de este.

Capítulo I Fundamentos teóricos: tratara sobre el estado del arte de la bibliometría y la cienciometría, ofrece la base de conocimiento necesaria para el desarrollo y/o comprensión de este trabajo, por tanto se tocan temas como las leyes bibliométricas, los indicadores bibliométricos y los autores principales que los proponen, también expone las tecnologías, lenguajes y metodologías utilizadas para su desarrollo teniendo en cuenta las tendencias actuales.

Capítulo II Modelo del negocio: explica el problema a resolver y describe el campo de acción donde se desarrolla, se especifican las características y conceptos fundamentales que conforman sus entidades definiéndose como reglas. Se realizan los diagramas, descripciones y modelos correspondientes a la metodología.

Capítulo III Requisitos: Describe profundamente los actores y casos de uso a automatizar y define los requisitos funcionales y no funcionales.

Capítulo IV Descripción de la solución propuesta: se enfoca en describir detalladamente el proceso llevado a cabo para la construcción de la solución de la problemática y describe de forma general el funcionamiento del sistema.

Capítulo V Estudio de factibilidad: describe un estudio de la factibilidad del sistema propuesto, analizando los costos y beneficios para su desarrollo.

Capítulo 1 Fundamentos teóricos

1.1 Introducción

En este capítulo se ofrece una base teórica y conceptual para el desarrollo del sistema en cuestión. Expone esencialmente los conceptos básicos relacionados con bibliometría y cienciometría sus leyes e índices. De la misma forma menciona las metodologías para el desarrollo de software y las herramientas para su correcta implementación.

1.2 Objetivos estratégicos de la academia

Academia, definida por el *DRAE* (Diccionario de la Real Academia Española) sería “*sociedad científica, literaria o artística establecida con autoridad pública y como establecimiento docente, público o privado, de carácter profesional, artístico, técnico, o simplemente práctico*”[8], dicha sociedad en un contexto práctico tiene como misión la creación y transmisión del conocimiento y para mantener esta autoridad, con la que ha sido dotada debe seguir la línea de la razón y la objetividad.

1.2.1 Investigación y producción científica

En este contexto Price (1963) definió la ciencia como aquello que se edita en las publicaciones científicas y al científico como la persona que ha colaborado escribiendo alguna de esas publicaciones[9]. Aunque es una definición muy abierta lo cierto es que la investigación científica queda incompleta sin su publicación, puesto que ésta es la que proporciona el proceso de conexión con la academia que puede evaluarla y dar o no su asentimiento[4].

1.2.2 Análisis crítico de la valoración de un área del conocimiento

Actualmente resulta necesario por varias razones valorar un área del conocimiento. Regularmente este tipo de investigación es llevada a cabo por expertos en la propia área y/o investigadores que se inician en ella, ambos están forzados a realizar múltiples tareas de forma manual o semiautomática que conduce a errores como ambigüedad de datos u otros comunes de los procesos engorrosos.

1.3 Procesos objeto de automatización

En el proceso de valoración de un área del conocimiento se llevan a cabo varios subprocesos, a continuación se describen solo aquellos subprocesos objetos de automatización que guardan una diferencia sustancial con los subprocesos actuales.

1.3.1 Desambiguación de datos

En contexto, la desambiguación de datos sería el proceso de deshacerse de publicaciones repetidas, con las que se pudiera cometer el error de desvalorar la investigación. [10]

1. Exportar las referencias a una hoja de cálculo o similar (e. g. *Microsoft Office Excel, Libre Office Calc*)
2. Ordenar alfabéticamente por autor, año, número de serie y título
3. Encontrar similitudes entre adyacentes y comprobar de que se trate de datos distintos en esencia
4. Eliminar o combinar datos que no sean distintos en esencia
5. Guardar la base de datos desambiguada

Es recomendable realizar este proceso más de una vez para asegurar la calidad de la nueva base de datos, ya que este debe ser el primero y su calidad pesa sobre el resto de los procesos.[10]

1.3.2 Recuento de publicaciones

A posteriori de la desambiguación, para medir la productividad por autores y conformar un escalafón debe realizarse un recuento de publicaciones por autor como se muestra en Tabla 1.

1. Exportar las referencias a una hoja de cálculo o similar
2. Ordenar alfabéticamente por el nombre del autor y contar los adyacentes iguales
3. Ordenar por cantidad de publicaciones[10]

NOMBRE	CANTIDAD DE PUBLICACIONES
DE Solla PRICE, D. J.	5
SPINAK, E.	4

Tabla 1. Ejemplo de resultado de un recuento de publicaciones

1.3.3 Recuento de publicaciones por tipo de publicación

A posteriori del recuento de publicaciones, para medir con mayor detalle el medio en donde se publica se realiza un recuento por tipo de publicación, siendo los más populares: revista, libro, capítulo (de libro), tesis, conferencia, encuesta, etc.

1. Exportar las referencias a una hoja de cálculo o similar
2. Ordenar alfabéticamente por el nombre del autor y contar los adyacentes iguales desechando las publicaciones que no sean en el medio determinado
3. Ordenar por cantidad de publicaciones[10]

1.3.4 Recuento de citas por autor

A posteriori del recuento de publicaciones, para medir los índices de impacto y así conformar un escalafón, debe realizarse un recuento de citas por autor como se muestra en Tabla 2.

1. Exportar las referencias a una hoja de cálculo o similar
2. Ordenar alfabéticamente por el nombre del autor

3. Para cada autor tomar cada publicación y buscar en sus notas (específicamente donde se indican las referencias), el nombre del autor y contar las coincidencias
4. Conformar un escalafón ordenado por cantidad de citas y otro ordenado por impacto[10]

Teniendo en cuenta que:

$$Im = \frac{Cc}{Cp}$$

Donde Im es el índice de impacto, Cc es la cantidad de citas y Cp es la cantidad de publicaciones.

AUTOR	CANTIDAD DE CITAS (CC)	CANTIDAD DE PUBLICACIONES (CP)	ÍNDICE DE IMPACTO (IM)
SPINAK, E.	8	4	2
DE SOLLA	6	5	1.2
PRICE, D. J.			

Tabla 2. Ejemplo de resultado de un recuento de citas

1.3.5 Recuento de citas por publicación

Para determinar el impacto de una publicación específica es necesario el recuento de citas, pero esta vez por publicación, como se muestra en Tabla 3.

1. Exportar las referencias a una hoja de cálculo o similar
2. Para cada publicación a medir tomar cada publicación y buscar en sus notas (específicamente donde se indican las referencias), el título de la publicación a medir y contar las coincidencias
3. Conformar un escalafón ordenado por cantidad de citas[10]

TÍTULO DE LA PUBLICACIÓN	CANTIDAD DE CITAS
INDICADORES CIENCIOMETRICOS	6
A GENERAL THEORY OF BIBLIOMETRIC AND OTHER CUMULATIVE ADVANTAGE PROCESSES	4

Tabla 3. Ejemplo de resultado de un recuento de citas por publicación

1.3.6 Recuento de colaboraciones

Para determinar los índices de colaboración entre autores es necesario un recuento de autoría, pero esta vez solo donde figuran los autores no principales como se muestra en Tabla 4.

1. Exportar las referencias a una hoja de cálculo o similar
2. Ordenar alfabéticamente por el nombre del autor
3. Para cada autor tomar cada publicación y buscar en sus autores secundarios, el nombre del autor a medir y contar las coincidencias

4. Conformar un escalafón con la cantidad de publicaciones total y ordenado por cantidad de colaboraciones[10]

Teniendo en cuenta que:

$$Ca = \frac{Co}{Ct}$$

$$Ct = Cp + Co$$

Donde Ca es el índice de colaboración entre autores, Co es la cantidad de colaboraciones y Ct es la cantidad de publicaciones en total.

AUTOR	CANTIDAD DE PUBLICACIONES EN TOTAL (CT)	CANTIDAD DE COLABORACIONES (CO)	ÍNDICE DE COLABORACIÓN ENTRE AUTORES (CA)
SPINAK, E.	10	6	0.6
DE SOLLA	8	3	0.375
PRICE, D. J.			

Tabla 4. Ejemplo de resultado de un recuento de publicaciones.

1.3.7 Recuento de publicación por año

Para clasificar un área del conocimiento en *emergente*, *decadente* o *conservada* en un periodo de tiempo es necesaria la revisión de su evolución temporal como se muestra en

1. Exportar las referencias a una hoja de cálculo o similar
2. Ordenar numéricamente por año
3. Para cada año contar la cantidad de publicaciones que lo conforman
4. Clasificar el periodo entre el año actual y el anterior según convenido[10]

Teniendo en cuenta que:

$$Ec = Y_n - Y_{n-1}$$

Donde Ec es la clasificación según la emergencia: *Si $Ec < 0$: decadente; Si $Ec > 0$: emergente; Si $Ec = 0$: conservada.* Y_n la cantidad de publicaciones en el año enésimo.

AÑO	CANTIDAD DE PUBLICACIONES	CLASIFICACIÓN
2013	68	-
2014	116	emergente
2015	92	decadente

Tabla 5. Ejemplo de recuento de publicaciones por año

1.3.8 Red de colaboración

Para obtener una visión mejor representada de la colaboración entre autores se procede a realizar la red de colaboración, que no es más que un grafo social donde el

nodo es un autor y los arcos apuntan a los otros autores con los que colabora, como se muestra en Ilustración 1.

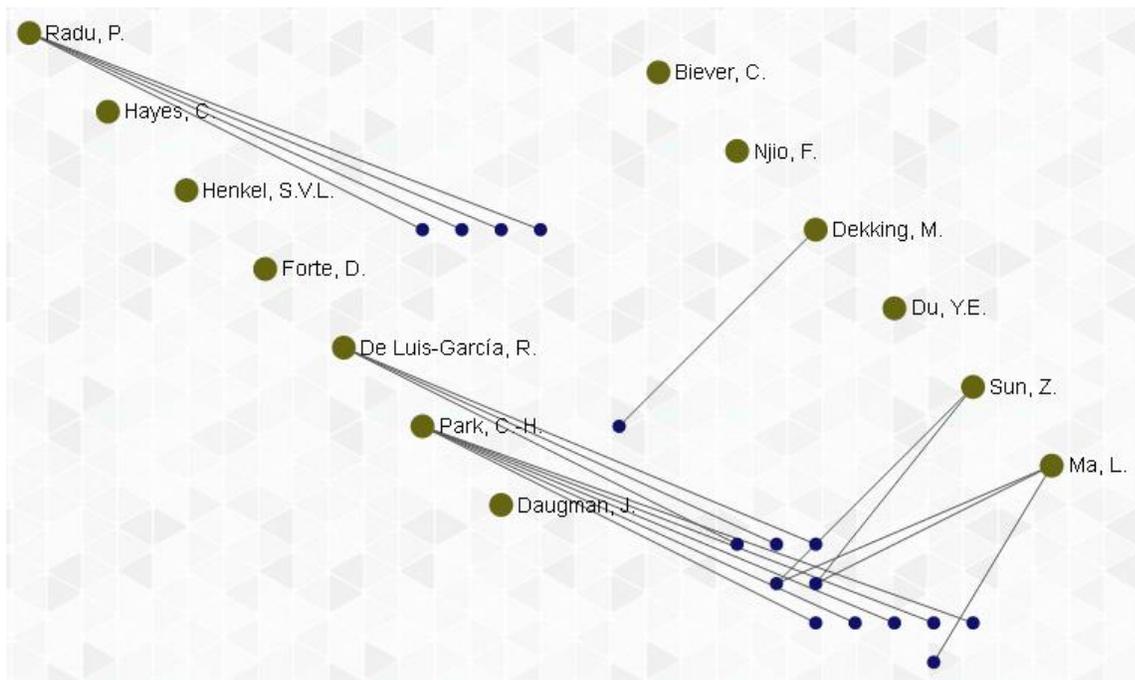


Ilustración 1. Ejemplo de red de colaboración entre autores

1.4 Sistemas automatizados existentes vinculados al campo de acción

1.4.1 Web of Knowledge (Web del conocimiento, WOK por sus siglas en inglés)

Durante muchos años los únicos índices de citas disponibles

Science Citation Index (SCI), *Social Sciences Citation Index (SSCI)* y *Arts and Humanities Citation Index (AHCI)* fueron los elaborados por el *Institute for Scientific Information (ISI)* fundado por Eugene Garfield en 1960. En la actualidad estos índices están integrados en el *Web of Science (WOS)*, suministrado por Thomson Reuters a través de la plataforma *Web of Knowledge (WOK)*. El *WOS* indexa unas 9.300 de las revistas de investigación más prestigiosas. De hecho, que una publicación periódica esté indexada en esta base de datos se considera internacionalmente un criterio significativo de calidad. Otro de los productos ofrecidos por el *WOK* es el *Journal Citation Reports* que evalúa el impacto de las revistas a partir de las citas. Existen dos ediciones, la *Science Edition*, que ofrece datos sobre el impacto de 5.900 publicaciones periódicas en ciencias; y el *Social Sciences Edition* que ofrece datos sobre 1.700 publicaciones en ciencias sociales y algunas de humanidades.[4]

1.4.2 SciVerse Scopus

El año 2002 Elsevier comenzó a publicar *Scopus*, oficialmente *SciVerse Scopus*, una base de datos bibliográfica con citas con voluntad de ser mucho más exhaustiva que el *WOS*. En la actualidad indexa unos 19.500 títulos. Sin embargo mientras que el *SCI*

recogen documentación que se remonta a 1900 y el *SSCI* desde 1956, *Scopus* solamente recoge resúmenes desde de 1960 y citas desde de 1996[4].

1.4.3 *Google Scholar (Google Académico)*

A finales de 2004 *Google Inc* lanzó *Google Scholar* o Google Académico. Fiel a la filosofía de *Google*, a diferencia de las bases de datos tradicionales, no vacía los contenidos de revistas sino que explora sistemáticamente la Web. Pero en este producto convergen dos servicios, un buscador de publicaciones científicas y un índice de citas que permita conocer el impacto de los trabajos publicados. Sin embargo, el uso del rastreo automático indiscriminado conlleva una importante dosis de errores. Por otra parte se indiza una amplia variedad de tipos de documentos, algunos de los cuales distan de lo que se considera académico (guías académicas, documentos administrativos, bibliografía de asignaturas, libros de divulgación, etc.). Este punto es importante ya que no es lo mismo ser citado en un documento científico que en otro tipo de documento. Más recientemente *Google Inc* ha incorporado dos nuevos productos *Google Scholar Citations (GSC)* y *Google Scholar Metrics (GSM)*. El primero recopila la producción científica de un investigador y la ofrece agregada en una página web, añadiendo información sobre el número de citas de cada referencia. En cambio *GSM* ofrece el impacto de las revistas científicas a partir de los recuentos de citas.

1.5 *Metodologías de desarrollo de software*

Un proceso de desarrollo de software minucioso, detallado y completo suele denominarse “Metodología”. Las metodologías se basan en una combinación de los modelos de proceso genéricos (cascada, evolutivo, incremental, espiral, entre otros). Adicionalmente una metodología debería definir con precisión los artefactos, roles y actividades involucrados, junto con prácticas y técnicas recomendadas, guías de adaptación de la metodología al proyecto, guías para uso de herramientas de apoyo, etc. Habitualmente se utiliza el término “método” para referirse a técnicas, notaciones y guías asociadas, que son aplicables a una (o algunas) actividades del proceso de desarrollo, por ejemplo, suele hablarse de métodos de análisis y/o diseño[11].

1.5.1 *El Proceso Unificado de Desarrollo de Software (RUP)*

Entre las más completas y populares de las metodologías de desarrollo de software, *RUP* obtiene el primer puesto indiscutible. Debido a la detallada y exhaustiva documentación que genera *RUP* es seleccionada como la metodología a seguir.

1.6 Patrones arquitectónicos

La arquitectura de software, trata con el diseño y la implementación de estructuras de software. Resultado de ensamblar varios elementos arquitectónicos de forma adecuada para lograr una funcionalidad óptima.

1.6.1 *Modelo Vista Controlador (MVC)*

El Modelo Vista Controlador (MVC) es un patrón arquitectónico de software que separa los datos y la lógica de negocio de una aplicación de la interfaz de usuario y el módulo encargado de gestionar los eventos y las comunicaciones (Ilustración 2). Para ello tres componentes distintos que son el modelo, la vista y el controlador, es decir, por un lado define componentes para la representación de la información, y por otro lado para la interacción del usuario. Este patrón de arquitectura de software se basa en las ideas de reutilización de código y la separación de conceptos, características que buscan facilitar la tarea de desarrollo de aplicaciones y su posterior mantenimiento.

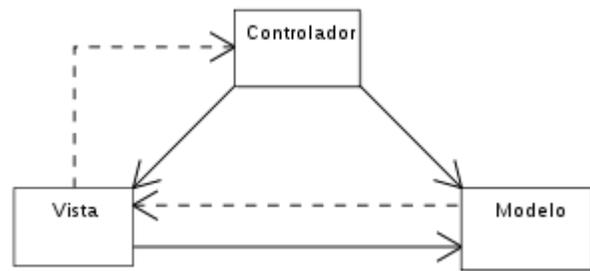


Ilustración 2. Diagrama representativo del patrón MVC.

1.7 Sistema Gestor de Bases de Datos

1.7.1 Base de Datos (BD)

Una Base de Datos (BD) es una colección estructurada de datos. Puede ser desde una simple lista de compras a una imagen de una galería o las inmensas cantidades de información de una red de corporaciones. Para agregar, acceder y procesar los datos almacenados en una base de datos es necesario un Sistema Gestor de Bases de Datos (SGBD)[12].

1.7.2 Sistema Gestor de Bases de Datos (SGBD)

Un Sistema Gestor de Bases de Datos (SGBD) es un conjunto de programas que sirven de capa de administración entre el usuario y la base de datos, son capaces de garantizar: integridad, seguridad, redundancia mínima, y consistencia.

1.7.3 MySQL

MySQL es un SGDB relacional, de licencia libre creado por David Axmark, Allan Larsson, and Michael “Monty” Widenius. Multihilo y multiusuario con más de seis millones de instalaciones es el SGBD más utilizado en el mundo, lo que le proporciona una amplia documentación[12].

1.8 Lenguajes

1.8.1 Unified Modeling Language (UML)

Unified Modeling Language (UML) es el lenguaje de modelado de sistemas de software más conocido y utilizado en la actualidad; está respaldado por el Object Management Group (OMG). Figura a ser un estándar aprobado por la Organización Internacional de Normalización (ISO por sus siglas en inglés) desde el año 2005.[11]

1.8.2 Lenguaje de programación PHP

PHP es un lenguaje de programación interpretado, diseñado originalmente para la creación de páginas web dinámicas. Se usa principalmente para la interpretación del

lado del servidor (*server-side scripting*) pero actualmente puede ser utilizado desde una interfaz de línea de comandos o en la creación de otros tipos de programas incluyendo aplicaciones con interfaz gráfica.[13]

1.8.3 Lenguaje de programación JavaScript

JavaScript es un lenguaje de programación interpretado, dialecto del estándar ECMAScript. Se define como orientado a objetos, basado en prototipos, imperativo, débilmente tipado y dinámico. Se utiliza principalmente en su forma del lado del cliente (*client-side*), implementado como parte de un navegador web permitiendo mejoras en la interfaz de usuario y páginas web dinámicas, en bases de datos locales al navegador, aunque existe una forma de JavaScript del lado del servidor (*Server-side JavaScript* o *SSJS*).[14]

1.8.4 HTML 5

HTML5 (*HyperText Markup Language*, versión 5) es la quinta revisión importante del lenguaje básico de la *World Wide Web*, HTML. HTML5 especifica dos variantes de sintaxis para HTML: un "clásico" HTML (*text/html*), la variante conocida como HTML5 y una variante XHTML conocida como sintaxis XHTML5 que deberá ser servida como XML (XHTML) (*application/xhtml+xml*).[15]

1.8.5 CSS 3

El nombre hojas de estilo en cascada viene del inglés *Cascading Style Sheets*, del que toma sus siglas. CSS es un lenguaje usado para definir la presentación de un documento estructurado escrito en HTML o XML (y por extensión en XHTML). El *W3C* (*World Wide Web Consortium*) es el encargado de formular la especificación de las hojas de estilo que servirán de estándar para los agentes de usuario o navegadores.

1.9 Frameworks

Los *frameworks* o marcos de trabajo son piezas de código, metodologías o prácticas que apoyan y guían el desarrollo de un software.

1.9.1 Yii framework

Yii es un *framework* orientado a objetos, software libre, de alto rendimiento basado en componentes, PHP. Yii se pronuncia en español como se escribe y es un acrónimo para "Yes It Is!" (En español: ¡Sí lo es!).

1.9.2 jQuery

jQuery es una biblioteca de JavaScript, creada inicialmente por John Resig, que permite simplificar la manera de interactuar con los documentos HTML, manipular el árbol DOM, manejar eventos, desarrollar animaciones y agregar interacción con la técnica AJAX a páginas web.[14]

1.9.3 Sigma

Sigma es una biblioteca de JavaScript dedicada al dibujo de grafos. Facilita la posibilidad de publicar redes en páginas web y brinda a los desarrolladores una forma sólida y limpia de programarla.

1.9.4 Bootstrap

Twitter Bootstrap es un *framework* o conjunto de herramientas de software libre para diseño de sitios y aplicación web, orientados a mejorar y facilitar el desarrollo de sitios web.[16]

1.10 Herramientas utilizadas

1.10.1 Formato de fichero RIS

El formato RIS (Sistemas de Información de Investigación, en inglés, *Research Information Systems*) es un formato de fichero estándar arrojado por gestores bibliográficos (e. g. *WoS*, *Scopus*, *Zotero*, *EndNote*, etc.)

1.10.2 NetBeans IDE

NetBeans es un entorno de desarrollo integrado libre, hecho principalmente para el lenguaje de programación Java. Existe además un número importante de módulos para extenderlo. NetBeans IDE es un producto libre y gratuito sin restricciones de uso.[17]

1.10.3 Firefox

Mozilla Firefox es un navegador web libre y de código abierto desarrollado para *Microsoft Windows*, *Mac OS X* y *GNU/Linux* coordinado por la Corporación Mozilla y la Fundación Mozilla. Usa el motor *Gecko* para renderizar páginas webs, el cual implementa actuales y futuros estándares web.

1.10.4 Firebug

Firebug es una extensión de Firefox creada y diseñada especialmente para desarrolladores y programadores web. Es un paquete de utilidades con el que se puede analizar (revisar velocidad de carga, estructura DOM), editar, monitorizar y depurar el código fuente, CSS, HTML y JavaScript de una página web de manera instantánea e inline. Firebug no es un simple inspector como *DOM Inspector*, además edita y permite guardar los cambios, un paso por delante del conocido *Web Developer*. Su atractiva e intuitiva interfaz, con solapas específicas para el análisis de cada tipo de elemento (consola, HTML, CSS, Script, DOM y Red), permite al usuario un manejo fácil y rápido. Firebug está encapsulado en forma de *plug-in* o complemento de Mozilla, es *Open Source*, libre y de distribución gratuita.

1.10.5 MySQL Workbench

MySQL Workbench es una herramienta visual de diseño de bases de datos que integra desarrollo de software, administración de bases de datos, diseño de bases de datos, creación y mantenimiento para el sistema de base de datos MySQL. Es el sucesor de *DBDesigner 4* de *fabFORCE.net*, y reemplaza el anterior conjunto de software, *MySQL GUI Tools Bundle*.

1.11 Conclusiones

Durante el capítulo fueron mencionados los fundamentos de la bibliometría y sus índices más importantes, se abordó sobre el *WOK*, *Scopus* y *Google Scholar* como los

sistemas proveedores de datos más extensos, prestigiosos o confiables. Fueron detallados los procesos objeto de automatización junto con ejemplos de salidas de datos necesarias. Se selecciona una metodología de desarrollo de software (RUP) y se menciona el patrón arquitectónico MVC. El resto del capítulo está dedicado a detallar los lenguajes, *frameworks* y herramientas utilizadas durante el desarrollo.

Capítulo 2 Modelo del negocio

2.1 Introducción

Durante el proceso de desarrollo de software, como la metodología (RUP) indica es de vital importancia identificar cada proceso del negocio y documentarlo. Los análisis bibliométricos regularmente se realizan *ad libitum* lo cual no garantiza en ninguna medida la calidad del resultado de la investigación. En 2013 fue publicado *An overview of iris recognition* donde se realiza un detallado acercamiento metodológico que nos facilita una lista de procesos ordenados a seguir para realizar un correcto estudio bibliométrico.

2.2 Modelo del negocio

Extraído y sintetizado de *An overview of iris recognition* la metodología propuesta en forma algorítmica es como sigue:

1. Adquisición de datos
2. Desambiguación de datos
3. Análisis *a priori*
4. Presentación
5. Análisis *a posteriori*

2.2.1 Adquisición de datos

Es necesaria una fuente de datos con un formato estándar, específicamente el formato RIS. Este formato es ampliamente usado por los gestores bibliográficos y describe las referencias bibliográficas extensamente. Es necesario que la fuente este lo menos contaminada posible, se recomienda que los datos sean extraídos de bases de datos oficiales como las ofrecidas por los productos del *WOK* o de *Scopus*.

2.2.2 Desambiguación de datos

Luego de adquirir los datos es prudente someterlos a un proceso de desambiguación para evitar duplicidad de datos, esta ambigüedad pudiera introducirse al insertar los datos por errores de escritura u otros.

2.2.3 Análisis *a priori*

Si bien en este momento es posible llegar a una conclusión, en este momento es donde se realiza lo descrito en *Procesos objeto de automatización*.

2.2.4 Presentación

Luego de culminar el análisis anterior, junto con los procesos que acarrea, queda tabular o representar gráficamente los resultados obtenidos en el Análisis *a priori*.

2.2.5 *Análisis a posteriori*

Una vez obtenidos los datos ofrecidos por los procesos anteriores y seguros de que la fuente de información es confiable, el investigador se puede disponer a generar las conclusiones más completas de su investigación bibliográfica.

2.3 *Reglas del negocio a considerar*

Las reglas del negocio describen políticas de cumplimiento mandatorio para un resultado coherente. Extraídas de la metodología antes mencionada e incluidas las necesarias para un servicio de calidad, las reglas del negocio son:

1. Cada investigador es el único autorizado de gestionar sus propios ejes de investigación.
2. El investigador solo puede gestionar sus referencias bibliográficas a través del eje de investigación correspondiente, solamente si la referencia pertinente pertenece al eje de investigación en cuestión y dicho eje pertenece al investigador.
3. No puede existir un investigador con más de un usuario.
4. El investigador es el único capaz de gestionar su perfil.
5. El administrador del sistema no puede acceder a los datos de los investigadores ni gestionarlos.

2.4 *Trabajador del negocio*

TRABAJADOR	DESCRIPCION
INVESTIGADOR	Investigador que pretende valorar un área del conocimiento determinada.

2.5 Diagrama de casos de uso del negocio

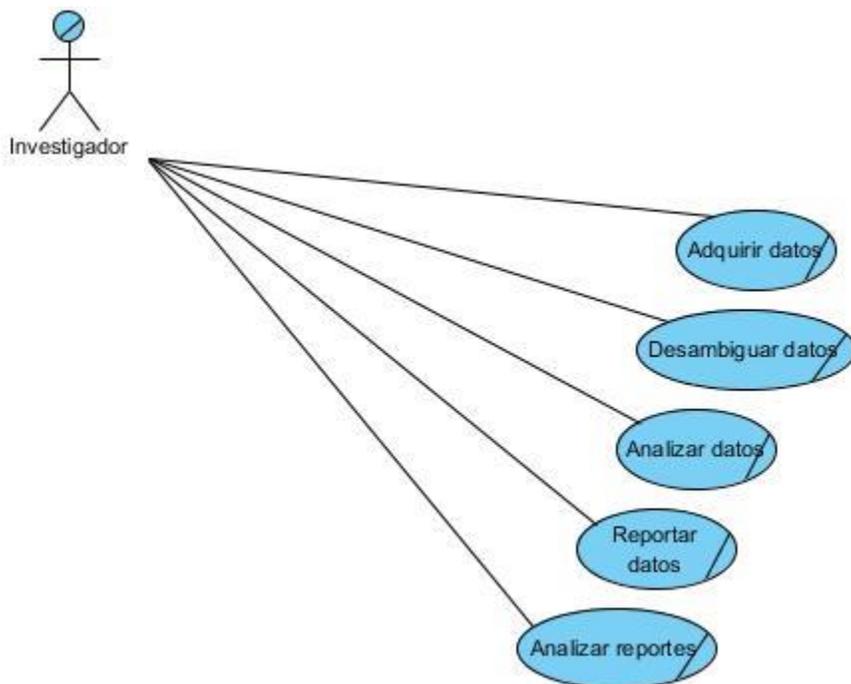


Ilustración 3. Diagrama de casos de uso del negocio

2.6 Diagramas de actividad

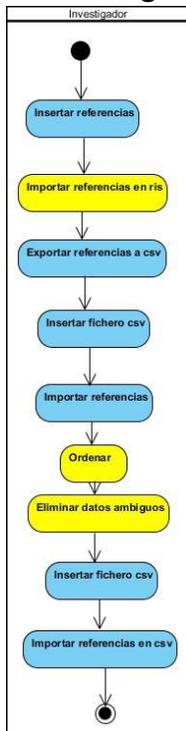


Ilustración 4 Diagrama de actividad del negocio, desambiguar datos.

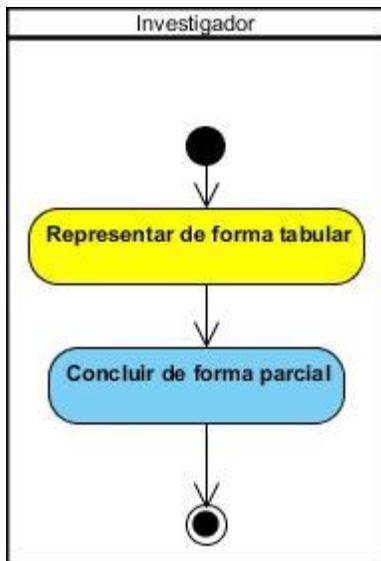


Ilustración 5. Diagrama de actividad del negocio, analizar datos.

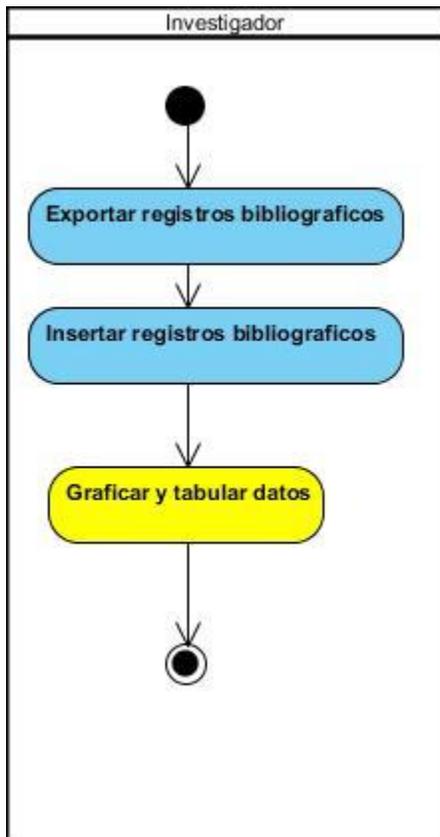


Ilustración 6. Diagrama de actividad del negocio, reportar datos.



Ilustración 7. Diagrama de actividad del negocio, analizar reportes.

2.7 Modelo de objetos del negocio

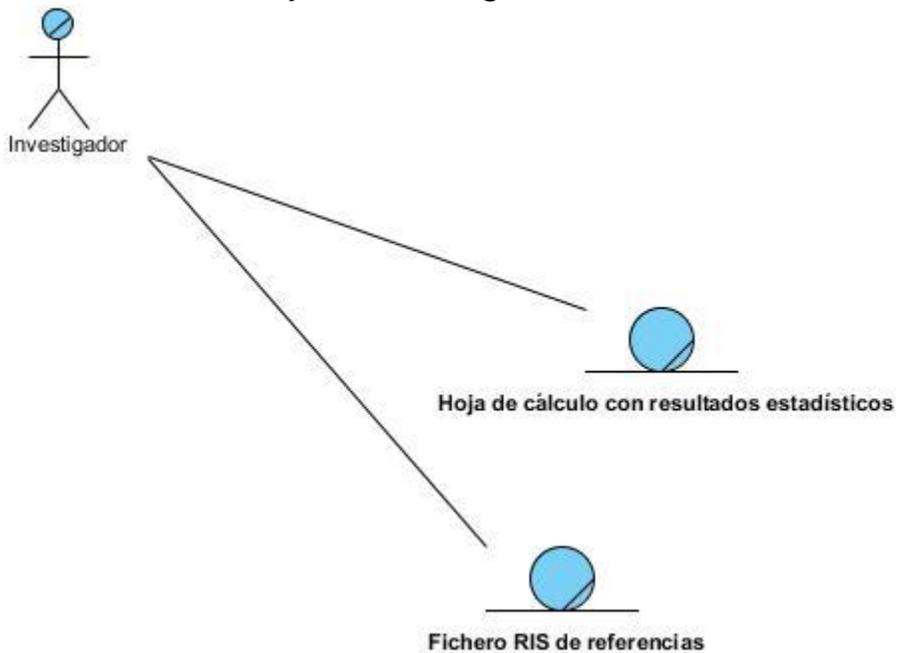


Ilustración 8. Modelo de objetos del negocio.

2.8 Conclusiones

Durante el capítulo fueron definidos como parte del flujo de trabajo de la modelación del negocio los siguientes artefactos: reglas del negocio, trabajador del negocio, diagrama de casos de usos del negocio, diagrama de actividades y modelo de objeto.

Capítulo 3 Requisitos

3.1 Actores del sistema a automatizar

ACTOR	DESCRIPCION
INVITADO	Usuario que solo puede consumir la pantalla para autenticarse y la de registrarse.
INVESTIGADOR	Usuario investigador que pretende valorar un área del conocimiento determinada.
ADMINISTRADOR	Usuario administrador del sistema, desarrolla labores de webmaster.

3.2 Definición de los requisitos funcionales

1. Base para el tratamiento de errores
2. Autenticar
3. Salir
4. Filtro para el control de acceso de usuario
5. Perfil de usuario
6. Asignar a usuario rol de superusuario
7. Asignar a usuario rol de investigador
8. Registrar usuario
9. Activar usuario
10. Desactivar usuario
11. Cambiar contraseña
12. Eliminar usuario
13. Filtro para el control de acceso de referencias
14. Examinar referencia
15. Modificar referencia
16. Importar referencia desde fichero con formato RIS
17. Eliminar referencia
18. Listar referencias
19. Determinar que publicaciones se refieren a una publicación determinada
20. Contar cuantas publicaciones se refieren a una publicación determinada
21. Determinar que autores se refieren a una publicación determinada
22. Generar de forma tabular datos para la red de referencias de una publicación determinada
23. Graficar red de referencias de una publicación determinada
24. Ordenar nodos de la red de referencias de una publicación determinada

25. Filtro para el control de acceso de ejes de investigación
26. Examinar un eje de investigación
27. Crear un eje de investigación
28. Modificar un eje de investigación
29. Eliminar un eje de investigación
30. Listar ejes de investigación
31. Exportar un eje de investigación
32. Crear y descargar un eje de investigación en formato ZIP
33. Listar reportes
34. Determinar segmentos de ambigüedad en un eje de investigación
35. Eliminar *in situ* publicaciones ambiguas
36. Determinar el título de una publicación determinada a partir de un fichero RIS
37. Traducir el tipo de publicación de una publicación determinada de la notación de formato RIS a español
38. Determinar los autores de una publicación determinada a partir de un fichero RIS
39. Determinar el autor principal de una publicación determinada a partir de un fichero RIS
40. Determinar autor(es) secundario(s) de una publicación determinada a partir de un fichero RIS
41. Determinar relación publicación/año de las referencias de un eje de investigación determinado
42. Graficar poligonalmente la relación cantidad de publicaciones / año
43. Determinar que publicaciones citan a un autor determinado
44. Determinar cuántas publicaciones citan a un autor determinado
45. Determinar que autores citan a un autor determinado
46. Determinar la producción de un autor determinado
47. Determinar las colaboraciones de un autor determinado
48. Determinar la producción general de un autor determinado
49. Determinar escalafón de los N autores más productivos
50. Determinar escalafón de los N autores más citados
51. Generar de forma tabular la red de colaboración entre autores
52. Graficar la red de colaboración entre autores
53. Ordenar la red de colaboración entre autores

3.3 Diagrama de casos de uso del sistema a automatizar

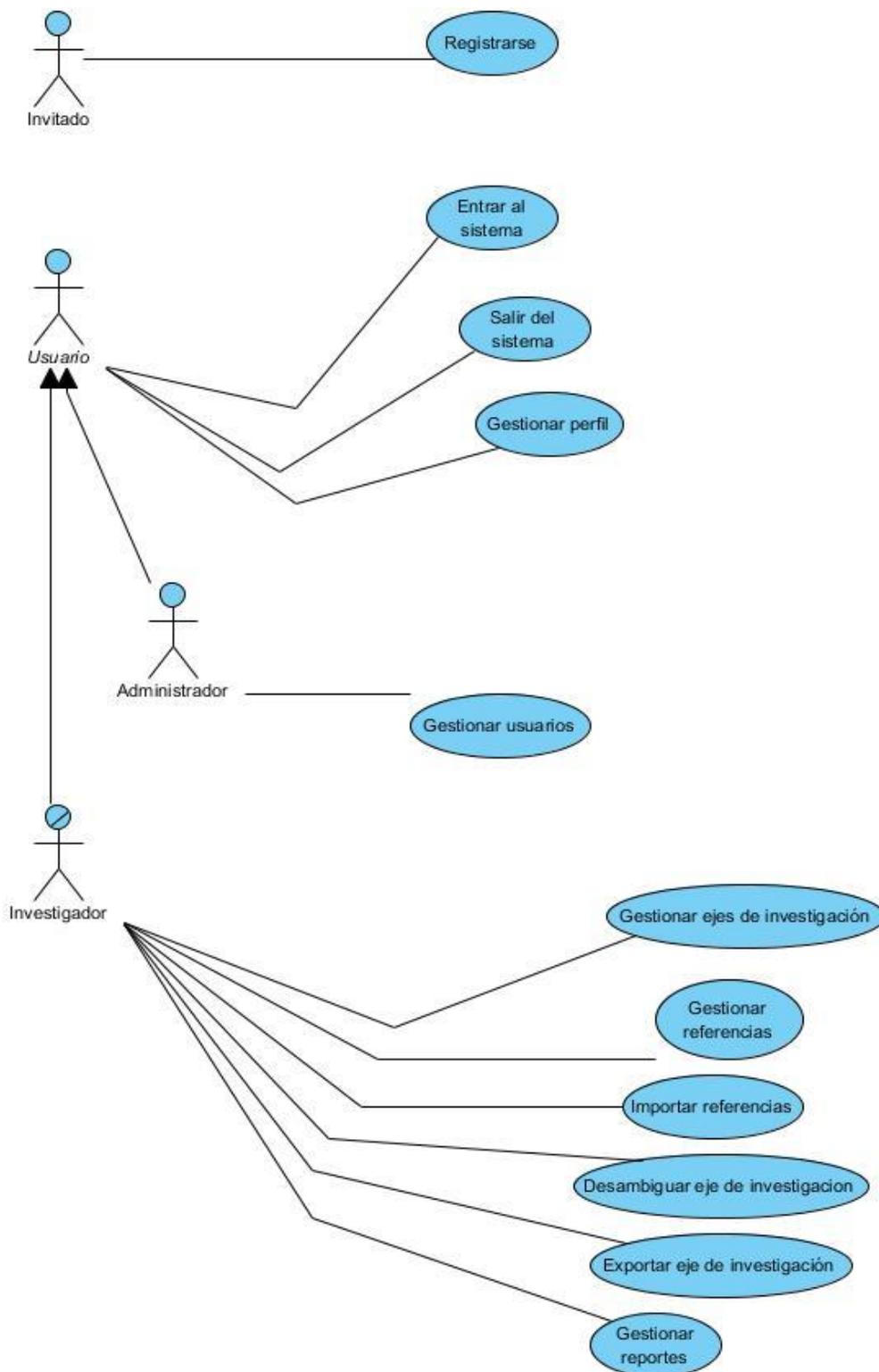


Ilustración 9. Diagrama de casos de uso del sistema

3.4 Definición de los requisitos no funcionales

- **Apariencia o interfaz externa:** El sistema posee una apariencia agradable e intuitiva, usa colores en tonalidades azules y grises en concordancia con los de la Universidad de Cienfuegos.
- **Usabilidad:** El software está dirigido a usuarios cuasi expertos y expertos en informática, capaces de manejar un navegador y navegar en una interfaz gráfica de usuario.
- **Rendimiento:** Es sistema debe realizar las tareas lo más rápido posible, no deberá exceder los 900 segundos de procesamiento.
- **Seguridad:** Es sistema cuenta con protección basada en contraseña alfanumérica, encriptada en la base de datos con un código de 32 bytes derivado de otro de 64 bytes, está protegido contra ataques XSS (*Cross Site Scripting*), CSRF (*Cross Site Request Forgery*) y CI (*Code Injection*). Además cuenta con el aparato de seguridad proporcionado por el framework *yii framework* evitando que se muestren errores técnicos de frente al usuario.
- **Fiabilidad:** El sistema deberá estar dispuesto 24 horas al día, 7 días a la semana.
- **Software:** El sistema está preparado para funcionar en navegadores modernos. Se asegura funcionamiento en Mozilla Firefox v25+ y Google Chrome v30+. Del lado del servidor se requiere Apache2, PHP 5.0+ y MySQL 5+.
- **Hardware:** Es sistema es un gran consumidor de recursos por lo que se recomienda un servidor dedicado con mínimo 10GB de disco duro, 2GB de memoria RAM, Intel Pentium Dual Core 2.6GHz y tarjeta de red con capacidad de transferencia de 100Mbs.

3.5 Descripción de los casos de uso

NOMBRE	ENTRAR AL SISTEMA
ACTORES	Usuario
RESUMEN	Al usuario se le presenta la pantalla de entrada, mostrándose el formulario para insertar nombre de usuario y contraseña, seguido de un botón para continuar y otro para registrarse. El usuario al insertar el usuario y la contraseña el sistema

	valida dichos datos y devuelve un mensaje de error en caso de que alguno de los datos fuesen incorrectos o continua a la pantalla de perfil de usuario. En caso de que el usuario se dirija al enlace para registrarse, se presentara la pantalla para registrarse.
PRECONDICIONES	Sistema en marcha y usuario registrado y activado en el sistema.
POSCONDICIONES	Error de autenticación o autenticación satisfactoria.
RQUISITOS ESPECIALES	Ninguno.

NOMBRE	REGISTRARSE
ACTORES	Invitado
RESUMEN	Al usuario se le presenta la pantalla de registro, mostrándose un mensaje de que un correo electrónico será enviado a la dirección que el usuario provea para activar la cuenta y un formulario para insertar sus datos, algunos de estos datos son de carácter mandatorio y derivaran en error en caso de que el usuario no los introduzca. Finalmente se encuentra un botón de enviar donde el usuario envía su solicitud de registro. Si el usuario inserto incorrectamente algún campo del formulario o dejó algún campo mandatorio vacío, el sistema mostrara un error con instrucciones para corregirlos. En caso de que todo esté en orden, el sistema enviará la solicitud de registro junto con un correo electrónico y esperara a que el usuario active su cuenta.
PRECONDICIONES	Sistema en marcha y usuario no registrado.
POSCONDICIONES	Error en caso de que se introduzca algún dato incorrectamente o éxito en caso de que todo esté en orden.
RQUISITOS ESPECIALES	Ninguno.

NOMBRE	SALIR DEL SISTEMA
ACTORES	Usuario
RESUMEN	Al usuario se le presenta casi en todo momento en el panel de la izquierda un botón con una rueda dentada, metáfora de configuración o de opciones. Al accionar sobre este botón se despliega una lista de opciones donde Salir se encuentra en primer lugar. Al accionar sobre salir

	el usuario dejará el sistema y se le presentara la pantalla para entrar.
PRECONDICIONES	Sistema en marcha y usuario registrado y dentro del sistema.
POSCONDICIONES	Usuario fuera del sistema
RQUISITOS ESPECIALES	Ninguno.

NOMBRE	GESTIONAR PERFIL
ACTORES	Usuario
RESUMEN	Al usuario se le presenta casi en todo momento en el panel de la izquierda un botón con una rueda dentada, metáfora de configuración o de opciones. Al accionar sobre este botón se despliega una lista de opciones donde Cambiar contraseña se encuentra en segundo lugar, al accionar sobre este, se le presentará una pantalla con un formulario para cambiar su contraseña. El usuario debe llenar los campos de los formularios correctamente o derivará en un error que no le permitirá realizar ningún cambio.
PRECONDICIONES	Usuario dentro del sistema.
POSCONDICIONES	Error al insertar los datos con la apropiada descripción e como corregir la entrada o mensaje de éxito al cambiar la contraseña.
RQUISITOS ESPECIALES	El usuario debe recordar la contraseña anterior.

NOMBRE	GESTIONAR EJES DE INVESTIGACION
ACTORES	Investigador
RESUMEN	En casi todo momento en el panel de la izquierda al investigador se le muestra un botón para crear nuevos ejes de investigación y una lista de los ejes de investigación disponibles. Al accionar sobre el botón Nuevo eje de investigación el sistema presentara un formulario para que inserte los datos correspondientes, derivara en un error en caso de que falte algún dato mandatorio o fuese insertado incorrectamente, en caso de éxito de esta operación mostrará una notificación de éxito y añadirá el nuevo eje a la lista de ejes de investigación. Al seleccionar un eje de investigación de la lista el sistema muestra una lista con las publicaciones que pertenecen a dicho eje de investigación y una variedad de opciones como: <ul style="list-style-type: none"> • Importar referencia: es necesario un

	<p>fichero RIS válido, de lo contrario el sistema mostrará una notificación de error.</p> <ul style="list-style-type: none"> • Exportar referencias: El sistema creará y dispondrá un fichero en formato ZIP donde se encontrarán los datos referentes al eje de investigación. • Modificar referencias: El sistema presentará una pantalla con un formulario con los datos referentes al eje de investigación similar al anterior de crearlo. • Eliminar eje de investigación: el sistema muestra un mensaje de confirmación que de resultar positivo es eliminado el eje de investigación y se muestra una notificación que lo confirma. • Buscar ejes de investigación: en el panel de la izquierda al usuario se le muestra un formulario con la entrada para una búsqueda, donde al accionar sobre el alguna palabra se filtrará la lista de ejes de investigación.
PRECONDICIONES	Investigador dentro del sistema.
POSCONDICIONES	Ninguna.
RQUISITOS ESPECIALES	Ninguno.

NOMBRE	GESTIONAR REFERENCIAS
ACTORES	Investigador
RESUMEN	<p>Al usuario se le presenta una lista de referencias que podrá filtrar por título, autor, tipo de publicación y palabras clave. Al accionar sobre el título de una referencia se presentará una pantalla para donde se detalla la referencia y son posibles las siguientes opciones:</p> <ul style="list-style-type: none"> • Modificar referencia: El sistema muestra una pantalla con un formulario con los datos correspondientes a una referencia. • Eliminar referencia: el sistema mostrará un mensaje de confirmación que de resultar

	positivo eliminara la referencia notificando el éxito de la operación.
PRECONDICIONES	Investigador dentro del sistema.
POSCONDICIONES	Ninguna.
RQUISITOS ESPECIALES	Ninguno.

NOMBRE	IMPORTAR REFERENCIAS
ACTORES	Investigador
RESUMEN	Al usuario se le presenta un formulario en el que deberá insertar fichero RIS válido, de no ser válido el fichero, el sistema derivará en un error. De ser correcto el fichero el sistema importara las referencias y las presentará automáticamente en la pantalla.
PRECONDICIONES	Investigador dentro del sistema.
POSCONDICIONES	Ninguna.
RQUISITOS ESPECIALES	El investigador debe poseer un fichero RIS válido.

NOMBRE	DESAMBIGUAR EJE DE INVESTIGACION
ACTORES	Investigador
RESUMEN	Al seleccionar un eje de investigación, una de las opciones que se le presenta al usuario es la de desambiguar eje de investigación. Al accionar sobre esta opción se presenta una pantalla una lista de segmentos de publicaciones donde se encontraron posibles ambigüedades. A la derecha de cada publicación se muestra un botón de eliminar para reducir las duplicidades <i>in situ</i> .
PRECONDICIONES	Investigador dentro del sistema.
POSCONDICIONES	Ninguna.
RQUISITOS ESPECIALES	Ninguno.

NOMBRE	EXPORTAR EJE DE INVESTIGACION
ACTORES	Investigador
RESUMEN	Al seleccionar un eje de investigación, entre las opciones que se muestran esta Exportar eje de investigación. Al accionar sobre este botón al usuario se le presenta un mensaje de descarga propio del navegador en el que se delegan el resto de las tareas.
PRECONDICIONES	Investigador dentro del sistema.
POSCONDICIONES	Ninguna.

RQUISITOS ESPECIALES	Debe de haber espacio libre en el disco duro.
-----------------------------	---

NOMBRE	GESTIONAR REPORTES
ACTORES	Investigador
RESUMEN	Al seleccionar un eje de investigación, entre las opciones que se muestran esta Reportes. Al accionar sobre este botón, se muestra una lista de reportes previamente confeccionados por el sistema por los que el usuario podrá navegar y hacer conclusiones generales.
PRECONDICIONES	Investigador dentro del sistema.
POSCONDICIONES	Ninguna.
RQUISITOS ESPECIALES	Ninguno.

3.6 Conclusiones

Durante el capítulo se definieron los actores del sistema, el diagrama de casos de uso del sistema, los requisitos funcionales y no funcionales del sistema y se describieron en detalle los casos de uso del sistema.

Capítulo 4 Descripción de la solución propuesta

4.1 Introducción

En la metodología utilizada (RUP) se emplean los diagramas de clases de diseño para una descripción general del sistema, un representación de un modelo lógico y otro modelo físico de los datos y un diagrama de despliegue que describe la distribución del sistema.

4.2 Diagrama de clases del diseño

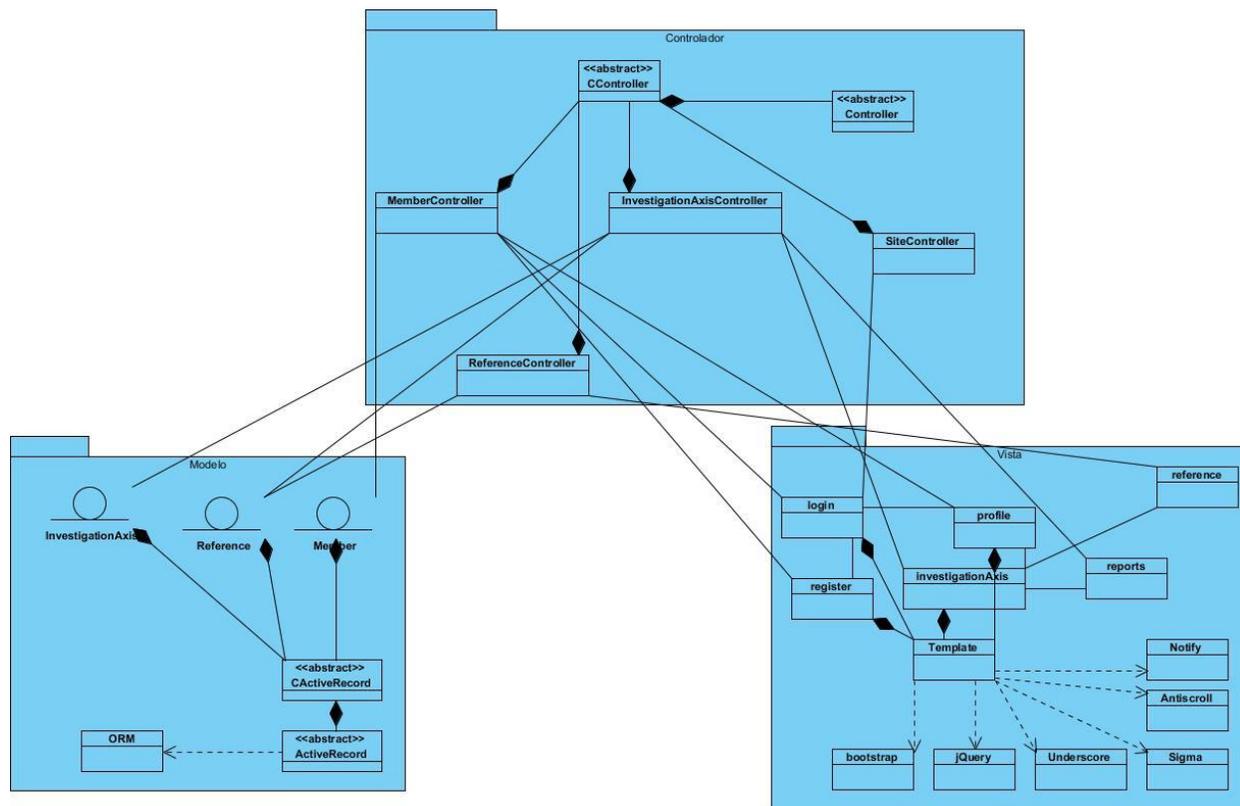


Ilustración 10. Diagrama de clases del diseño

4.3 Principios de diseño

4.3.1 Patrones de diseño

- Experto: es el principio básico de asignación de responsabilidades. A una clase le corresponde la responsabilidad de lo referente a un objeto.
- Creador: asigna responsabilidades de creación de un objeto a la clase que contiene la información necesaria para crearlo.

- Alta cohesión: expresa que la información que almacena una clase debe ser coherente y está en la mayor medida de lo posible relacionada con la clase.
- Bajo acoplamiento: define que las clases deben depender lo menos posible una de otra, para que las modificaciones resulten simples y limpias.

4.3.2 Interfaz de usuario

- Vista principal: es una pantalla dividida en 2 a la izquierda un menú con opciones generales del sistema y a la derecha lo que se está examinando en el momento, las opciones disponibles en el momento se muestran en la barra superior de la pantalla derecha.
- Vista auxiliar: diseñada para ganar más espacio, es utilizada en pantallas donde no es necesaria mucha navegación. Tiene una barra superior visible solo al principio de la página para volver a la vista principal.

Se utilizan colores azules y grises, fuentes blancas y negras para apoyar contraste y texturas para suavizar el contenido y aparentar profundidad.

4.3.3 Formato de salida de los reportes

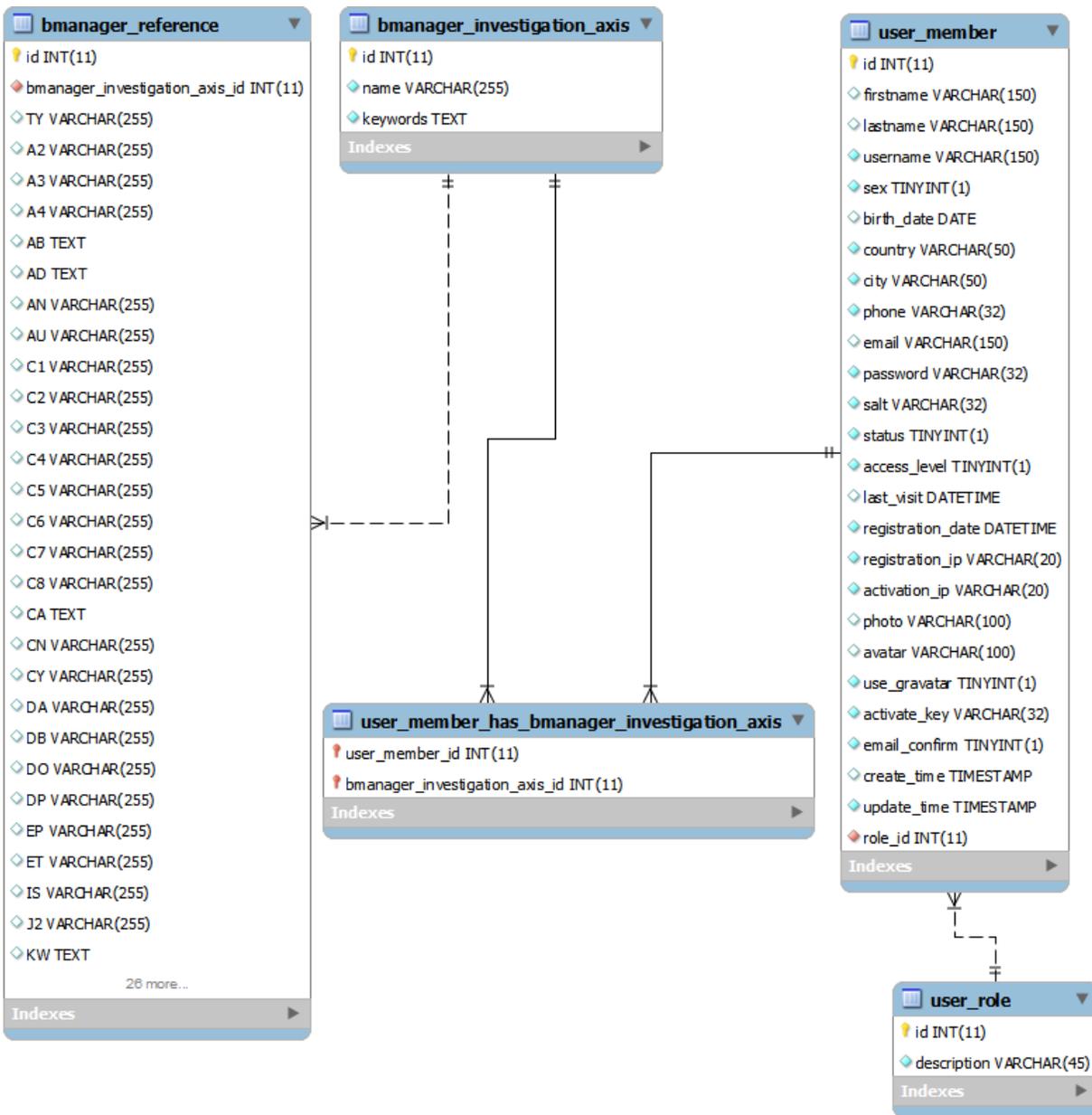
Los reportes tabulares se representan con hasta 20 filas en tablas interlineadas de blanco y gris. Los reportes gráficos poligonales se representan en azul con una leyenda a la derecha. Los reportes de redes se representan con nodos principales de color verde y nodos secundarios de color marrón, con posibilidad de zoom y etiqueta y un filtro que vuelve gris los nodos no conectados al seleccionado.

4.4 *Tratamiento de errores*

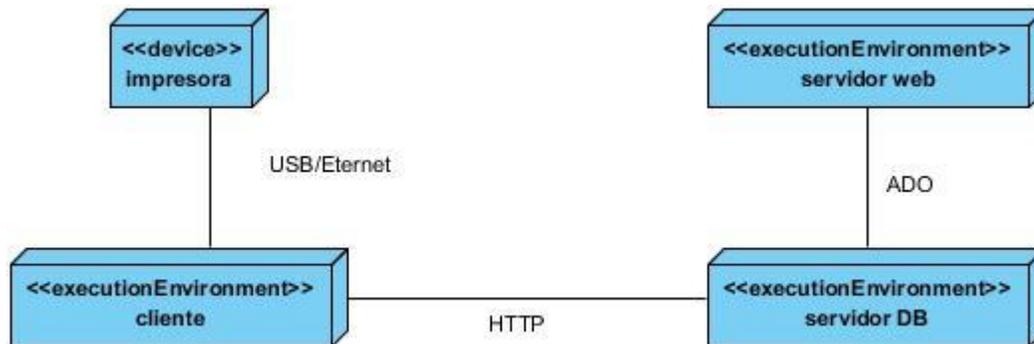
Los errores del sistema son presentados al usuario con una página de error advirtiéndole que el sistema es sensible o mediante una notificación roja en la esquina superior derecha.

4.5 Diseño de la base de datos

4.5.1 Modelo físico de datos



4.6 Diagrama de despliegue



4.7 Conclusiones

Durante el capítulo se detallan el diseño de clases y los patrones de diseño aplicados, se describe el tratamiento de errores adoptado por el sistema así como el diseño de la base de datos, para finalmente diagramar el sistema desplegado

Capítulo 5 Estudio de factibilidad

5.1 Introducción

En este capítulo se mediante un estimado de recursos humanos y tiempo un estudio de factibilidad al software con el objetivo de concluir si su desarrollo es factible o no.

5.2 Planificación basada en casos de uso

$$UUCP = UAW + UUCW$$

Conociendo que:

UUCP: Puntos de casos de uso sin ajustar

UAW: Factor de peso de los actores sin ajustar

UUCW: Factor de peso de los casos de uso sin ajustar

CASO DE USO DEL CLASIFICACION SISTEMA

REGISTRAR	Medio
ENTRAR	Simple
SALIR	Simple
GESTIONAR PERFIL	Simple
GESTIONAR USUARIOS	Simple
GESTIONAR EJES DE INVESTIGACION	Simple
GESTIONAR REFERENCIAS	Medio
IMPORTAR REFERENCIAS	Complejo
DESAMBIGUAR EJE DE INVESTIGACION	Complejo
EXPORTAR EJE DE INVESTIGACION	Medio
GESTIONAR REPORTES	Complejo

Tabla 6. Clasificación de los casos de uso

CLASIFICACION	CANTIDAD
SIMPLE	5
MEDIO	3
COMPLEJO	3
TOTAL	11

Tabla 7. Conteo de las clasificaciones de los casos de uso.

CLASIFICACION	FACTOR DE PESO
SIMPLE	5

MEDIO	10
COMPLEJO	20

Tabla 8. Relación de clasificación y peso

$$UUCW = 5 \times 5 + 3 \times 10 + 3 \times 20$$

$$UUCW = 25 + 30 + 60$$

$$UUCW = 115$$

CANTIDAD DE ACTORES DE SISTEMA FACTOR DE PESO

4	5
----------	----------

$$UAW = 4 \times 5$$

$$UAW = 20$$

Por tanto:

$$UUCP = 20 + 115$$

$$UUCP = 135$$

Ajustando:

$$UCP = UCCP \times TCF \times EF$$

Conociendo que:

UCP: Puntos de casos de uso ajustados.

UUCP: Puntos de casos de uso sin ajustar.

TCF: Factor de complejidad técnica.

EF: Factor de ambiente.

FACTOR	DESCRIPCION	PESO	VALOR	PESO X VALOR
T1	Sistema distribuido.	3	2	6
T2	Objetivos de performance o tiempo de respuesta.	3	3	9
T3	Eficiencia del usuario final.	1	2	2
T4	Procesamiento interno complejo.	2	3	6
T5	El código debe ser reutilizable.	1	3	3
T6	Facilidad de instalación.	.5	2	1
T7	Facilidad de uso.	.5	2	1
T8	Portabilidad.	2	4	8
T9	Facilidad de cambio.	1	3	3
T10	Concurrencia.	1	3	3
T11	Incluye objetivos especiales de seguridad.	1	3	3
T12	Provee acceso directo a terceros	1	3	3

T13	partes. Se requieren facilidades especiales c 1 entrenamiento a usuarios.	3	3
------------	---	---	---

Tabla 9. Relación de complejidad técnica.

$$TFC = 0.6 + 0.01 \times \sum P_i \times V_i$$

$$TFC = 0.6 + 0.01 \times (6 + 9 + 2 + 6 + 3 + 1 + 1 + 8 + 3 + 3 + 3 + 3 + 3)$$

$$TFC = 1.11$$

FACTOR	DESCRIPCION	PESO	VALOR	PESO X VALOR
E1	Familiaridad con el modelo de proyecto utilizado.	2	4	8
E2	Experiencia en la aplicación.	1	4	4
E3	Experiencia en orientación a objetos.	1	4	4
E4	Capacidad del analista líder.	1	4	4
E5	Motivación.	1	5	5
E6	Estabilidad de los requerimientos.	1	5	5
E7	Personal part-time.	0	3	0
E8	Dificultad del lenguaje de programación.	0	3	0

Tabla 10. Relación de factor de ambiente.

$$FE = 1.4 - 0.03 \times \sum P_i \times V_i$$

$$FE = 1.4 - 0.03 \times (8 + 4 + 4 + 4 + 5 + 0 + 0)$$

$$FE = 0.65$$

Resultando en:

$$UCP = 115 \times 1.11 \times 0.65$$

$$UCP = 82.9725$$

Estimando esfuerzo:

$$E = UCP \times CF$$

Conociendo que:

E: esfuerzo (horas/hombre)

CF: factor de conversión: CF = 20 horas / hombre

$$E = 82.9725 \times 20$$

$$E = 1659.45$$

Por tanto:

Trabajando 8 horas diarias como promedio 24 días al mes un grupo de 2 personas el desarrollo de proyecto se prolonga:

Horas: 829.725

Días: 103.715625

Meses: 4.321484375

Tomando como salario mensual un promedio de \$1400

Costo: \$12100.15625

5.3 Beneficios tangibles e intangibles

El sistema informático desarrollado, como resultado del presente trabajo de diploma, no implica costo alguno para la Universidad de Cienfuegos o cualquier entidad donde se pretenda implantar, sin embargo, al desarrollo de todo producto informático va asociado un costo y su justificación económica viene dado por los beneficios tangibles e intangibles que este produce.

El uso de este sistema informático permite la agilización de un proceso de investigación o una revisión bibliográfica en cuanto al tiempo. Además garantiza una calidad en el resultado aprobada por la comunidad científica.

5.4 Conclusiones

Durante el capítulo se realizó un estudio de factibilidad con una planificación basada en casos de uso, determinando que un equipo de 2 personas durante un periodo aproximado de 4 meses finalizan el software con un costo de aproximadamente \$12100. También se describen los beneficios que produce el software de implantarse en alguna entidad académica, concluyendo que es factible su desarrollo.

Conclusiones

A lo largo del análisis de todo el documento podemos concluir que:

- Los estudios bibliométricos tienen particularidades, leyes e índices específicos y una gran utilidad para la ciencia actual. Brinda una solución viable para valorar la creciente información que hoy día nos inunda.
- El proceso actual para realizar estudios bibliométricos es largo, engorroso y puede llevar a errores propios de procesos de esta índole. Está necesitado de una metodología bibliométrica que le guíe a través de todo el proceso para garantizar la calidad del estudio.
- Se analizaron las tecnologías y herramientas para el desarrollo web y se seleccionaron las mismas con el fin de que el sistema fuese escalable, mantenible y durable.
- Se determinó que es factible el desarrollo del sistema en un periodo relativamente corto de tiempo, con un equipo relativamente pequeño con un costo comparativamente bajo.
- Se diseñó e implementó un sistema informático capaz de realizar valoraciones al progreso actual de un campo del saber determinado.

Recomendaciones

1. Optimizar procesos largos realizando pre cálculo para reducir el tiempo de espera del usuario.
2. Implementar un módulo para cálculo de índices, de forma que con el transcurrir del tiempo y al cambiar los índices, estos puedan ser actualizados de manera simple.
3. Implementar un módulo para la generación de reportes para que sean actualizados de manera simple
4. Implantar el sistema en la Universidad de Cienfuegos para que los investigadores puedan consumirlo como servicio.

Referencias bibliográficas

- [1] D. de Solla Price, "A General Theory of Bibliometric and Other Cumulative Advantage Processes," 1976.
- [2] A. Cornella, "Cómo sobrevivir a la infoxicación." [Online]. Available: <http://www.uoc.es/web/esp/articles/cornella/acornella.htm>. [Accessed: 22-May-2015].
- [3] E. Spinak, "Indicadores cientificos," 1998.
- [4] J. Ardanuy, "Breve introducción a la bibliometría," 2012.
- [5] "Science español," 2012. [Online]. Available: <http://ip-science.thomsonreuters.com/es>. [Accessed: 22-May-2015].
- [6] "About Scopus," 2012. [Online]. Available: <http://www.info.sciverse.com/scopus/about>. [Accessed: 22-May-2015].
- [7] E. Cabezas-Clavijo and Igado-López-Cózar, "Scholar Metrics: el impacto de las revistas según Google, ¿un divertimento o un producto científico aceptable?," *EC3noticias*, 2012. .
- [8] DRAE, "Academia." [Online]. Available: <http://lema.rae.es/drae/?val=academia>. [Accessed: 22-May-2015].
- [9] D. de Solla Price, "Little science, big science," 1963.
- [10] Y. Alvarez Betancourt and M. Garcia-Silvente, "An overview of iris recognition," Dec. 2013.
- [11] I. Jacobson, B. Grady, and J. Rumbaugh, *El Proceso Unificado de Desarrollo de Software*. .
- [12] *MySQL 5.1 Reference Manual*. 2007.
- [13] *PHP 5.6.0 released*. 2014.
- [14] "Usage of JavaScript libraries for websites." [Online]. Available: http://w3techs.com/technologies/overview/javascript_library/all.
- [15] I. Hickson and D. Hyatt, *HTML5*. .
- [16] "GitHub: Popular Watched Repositories." [Online]. Available: <https://github.com/popular/watched>. [Accessed: 26-Jul-2012].
- [17] "https://netbeans.org/community/releases/roadmap.html." [Online]. Available: <https://netbeans.org/community/releases/roadmap.html>.

Glosario de términos

Cienciometría: rama del saber que estudia la actividad científica con el fin de medir y analizarla.

Bibliometría: aplica métodos matemáticos y estadísticos a toda la literatura de carácter científico y a los autores que la producen.