UNIVERSIDAD CATÓLICA ANDRÉS BELLO

FACULTAD DE INGENIERÍA

ESCUELA DE INGENIERÍA EN INFORMÁTICA

SISTEMA DE APOYO A LA TOMA DE DECISIONES PARA LA ESCUELA DE INGENIERÍA INFORMÁTICA DE LA UNIVERSIDAD CATÓLICA ANDRÉS BELLO

TRABAJO ESPECIAL DE GRADO

Presentado ante la

UNIVERSIDAD CATÓLICA ANDRÉS BELLO

Como parte de los requisitos para optar al título de

INGENIERO EN INFORMÁTICA

REALIZADO POR **Martinez Farías, Carlos E.**

Moreno Fernández, Francisco J.

PROFESOR GUIA **Susana García**

FECHA **Septiembre 2009**

UNIVERSIDAD CATÓLICA ANDRÉS BELLO

FACULTAD DE INGENIERÍA

ESCUELA DE INGENIERÍA EN INFORMÁTICA

SISTEMA DE APOYO A LA TOMA DE DECISIONES PARA LA ESCUELA DE INGENIERÍA INFORMÁTICA DE LA UNIVERSIDAD CATÓLICA ANDRÉS BELLO

Este Jurado; una vez realizado el examen del presente trabajo ha evaluado su contenido con el resultado: 19 puntos con Mención Honorífica.

J U R A D O E X A M I N A D O R

Firma: Andreina Castello Firma: Firma: Gloria Farrio V.

Nombre: Andreina Castello Nombre: Susana García M Nombre: Gloria Farrio V.

REALIZADO POR **Martínez Farías, Carlos E.**

Moreno Fernández, Francisco J.

PROFESOR GUIA **Susana García**

FECHA **Septiembre 2009**

AGRADECIMIENTOS

En primer lugar a mi tutora Susana García por toda su dedicación y apoyo durante todo el desarrollo de este trabajo, por toda su colaboración y consejos a lo largo de los años que tenemos conociéndonos.

Un especial agradecimiento a la profesora Lúcia Cardoso por todo su apoyo durante el desarrollo de este trabajo, por todos sus consejos, por su disposición a enseñarnos a desarrollar un trabajo verdaderamente profesional y de calidad, sinceramente muchísimas gracias.

A todos los coordinadores por su colaboración y disposición de atendernos y dar lo mejor de sí para que este trabajo logre ayudarles en sus labores diarias.

A todas aquellas personas que me han apoyado y me han dado ánimos para continuar con este trabajo cuando realmente se veía difícil la culminación del mismo.

Francisco J. Moreno F.

Me gustaría agradecer a nuestra tutora Susana García, por toda la ayuda y dedicación durante la realización de este trabajo de grado, gracias por su colaboración y paciencia.

A todos los coordinadores en especial a la profesora Lúcia Cardoso, por siempre estar dispuesta a atendernos con mucha paciencia y por aconsejarnos en todo momento, muchas gracias profesora.

A Francisco y a toda su familia, por hacerme sentir en mi casa y por tratarme como un miembro más de su familia, estoy muy agradecido con ustedes.

A todas las demás personas que siempre me dieron ánimos cuando las cosas se complicaban, muchas gracias a todos.

Carlos E. Martínez F.

DEDICATORIA

En primer lugar este trabajo está dedicado a Dios, por la oportunidad que me ha dado de poder culminar todos los estudios a pesar de los tantos obstáculos que aparecieron en mi vida y de poder lograr ser un profesional.

A mis padres por todo su apoyo y disposición de ayudarme en lo que fuese que necesitara para poder terminar mis estudios y convertirme en profesional.

A todos mis familiares que siempre estuvieron dando ánimos y fuerza para continuar por este camino lleno de obstáculos para poder lograr culminar mis estudios. En especial a mi abuelo que siempre estuvo apoyándome para seguir adelante con mis estudios.

Francisco J. Moreno F.

Primero que nada dedicarle y agradecer a dios por toda la fuerza que me dio para seguir siempre adelante y nunca mirar atrás.

A mi mama y mi papa, sin su apoyo incondicional no lo hubiese podido lograr, este logro es de ustedes, los quiero mucho viejos.

A mi hermana Coralia, Cora gracias por toda la ayuda y apoyo que me brindaste durante toda la carrera, este logro es tuyo. Te quiero muchísimo.

A mi hermana Carlina, Miñi gracias por tu apoyo y consejos claves que siempre me ayudaron y me seguirán ayudando siempre, te dedico este logro. Te quiero mucho hermana.

A mis sobrinas Camila y Carlota, gracias por darle alegría a mi vida, espero que este logro les sirva de ejemplo para su futuro. Las quiero muchísimo.

Carlos E. Martínez F.

ÍNDICE

AGRADECIMIENTOS I
DEDICATORIA II
ÍNDICE III
ÍNDICE DE FIGURAS VI
ÍNDICE DE TABLAS VII
SINOPSIS VIII
INTRODUCCIÓN 1
CAPÍTULO I. PROBLEMA Y OBJETIVO 3
I.1 PLANTEAMIENTO DEL PROBLEMA 3
I.2 OBJETIVO GENERAL 4
I.3 OBJETIVOS ESPECÍFICOS 4
I.4 ALCANCE 5
I.5 JUSTIFICACIÓN 6
CAPÍTULO II. MARCO TEÓRICO 8
II.1 SISTEMAS DE APOYO A LA TOMA DE DECISIONES 8
II.1.1 Sistemas de Apoyo a la Toma de Decisiones 8
II.1.1.1 Componentes de los Sistema de Apoyo a la Toma de Decisiones 8
II.1.2 Tipos de Decisiones 9
II.1.3 Sistemas de soporte a la decisión en la educación, EDSS (Educational Decision Support System) 9
II.1.4 Sistemas de información en la educación 11
II.2 INDICADORES 13
II.2.1 Definición 13
II.2.2 Requisitos de un buen indicador 13
II.2.3 Los indicadores educativos 14
II.3 ALMACÉN DE DATOS (DATA WAREHOUSE) 15
II.3.1 Definiciones 15
II.3.2 Componentes de los almacenes de datos 17
II.3.3 Beneficios de un datawarehouse 19
II.4 DATAMART 21
II.4.1 Definiciones 21
II.4.2 Diferencias entre un DataWarehouse y un DataMart 22
CAPÍTULO III. MARCO METODOLÓGICO 23
III.1 FASES DEL MODELO 25
III.1.1 Concepto de Software 25
III.1.2 Análisis de Requerimientos 25
III.1.2.1 Identificación de indicadores 26
III.1.2.2 Evaluación y selección de la herramienta de desarrollo 26

III.1.3 Diseño Global 26
III.1.4 Diseño e Implementación del almacén de datos 26
III.1.4.1 Análisis del origen y destino de datos. 26
III.1.4.2 Diseño del almacén de datos 27
III.1.4.3 Implementación del almacén de datos 27
III.1.4.4 Definición de los procesos de Extraccion, Transformacion y Transporte (ETT). 27
III.1.4.5 Implementación de los procesos de Extraccion, Tranformacion y Transporte (ETT). 27
III.1.4.6 Elaboración de las pruebas (ETT) y Carga del almacén de datos .. 27
III.1.5 Diseño de la aplicación de reportes. 27
III.1.6 Implementación de la aplicación de reportes 28
III.1.7 Elaboración de las pruebas de la aplicación de reportes. 28
III.2 FASES DEL MODELO 29
III.2.1 Análisis de Requerimientos 29
III.2.2 Diseño Global 29
III.2.3 Diseño Detallado 29
CAPÍTULO IV. DESARROLLO 30
IV.1 INTEGRACIÓN DE LOS SISTEMAS TRANSACCIONALES DE LA ESCUELA DE INGENIERÍA INFORMÁTICA. 30
IV.1.1 Análisis de Requerimientos. 30
IV.1.2 Diseño Global. 31
IV.1.3 Diseño Detallado 31
IV.1.3.1 Documentación de Procesos. 32
IV.1.3.2 Definición de Estándar de Nomenclatura de Base de Datos 33
IV.1.3.3 Diseño de la Base de Datos Integrada. 33
IV.1.3.4 Documentación de la Base de Datos Integrada. 35
IV.1.3.5 Documentación de Funcionalidad. 35
IV.2 DESARROLLO DEL SISTEMA DE APOYO A LA TOMA DE DECISIONES 36
IV.2.1 Concepto de Software 36
IV.2.2 Análisis de Requerimientos. 37
IV.2.2.1 Identificación de Indicadores. 38
IV.2.2.2 Evaluación y selección de la herramienta de desarrollo. 42
IV.2.3 Diseño Global. 46
IV.2.4 Diseño e Implementación del Almacén de Datos. 48
IV.2.4.1 Análisis del origen y destino de datos. 50
IV.2.4.2 Diseño del Almacén de Datos 50
IV.2.4.3 Implementación del Almacén de Datos. 53
IV.2.4.4 Definición de los procesos ETT. 53
IV.2.4.5 Implementación de los procesos ETT. 54
IV.2.4.6 Elaboración de las pruebas ETT y carga del almacén de datos 56

IV.2.5 Diseño de la aplicación de reportes. 57
IV.2.6 Implementación de la aplicación de reportes 57
IV.2.6.1 Diseño e Implementación de los cubos OLAP. 57
IV.2.6.2 Reportes. 60
IV.2.7 Elaboración de las pruebas de la aplicación de reportes. 62
CAPÍTULO V. RESULTADOS 63
CONCLUSIONES 66
RECOMENDACIONES 68
REFERENCIAS BIBLIOGRÁFICAS 69
GLOSARIO 73

ÍNDICE DE FIGURAS

FIGURA 1. ARQUITECTURA DE LOS ALMACENES DE DATOS. 17
FIGURA 2. MODELO DEL CICLO DE DESARROLLO ELABORADO PARA LA REALIZACIÓN DEL SISTEMA DE APOYO PARA LA TOMA DE DECISIONES... 24
FIGURA 3 .MODELO DEL CICLO DE DESARROLLO ELABORADO PARA LA INTEGRACIÓN DE LAS APLICACIONES..... 29
FIGURA 4. DIAGRAMA DE FLUJO DE PROCESOS DE LAS MATERIAS ELECTIVAS. 32
FIGURA 5. MODELO ENTIDAD-RELACIÓN DE LA INTRANET INTEGRADA. 34
FIGURA 6. ARQUITECTURA DEL SISTEMA DE APOYO A LA TOMA DE DECISIONES. 47
FIGURA 7. DISEÑO DE ALMACÉN DE DATOS. 52
FIGURA 8. DIAGRAMA ETT PARA LOS INDICADORES RELACIONADOS A LOS PROFESORES. 55
FIGURA 9. DISEÑO DEL CUBO PARA LA TABLA DE FACTDIRECCIONEII. 59
FIGURA 10. DATASOURCE CON LOS REPORTES CREADOS. 60
FIGURA 11. PANTALLA DE INICIO DE LA APLICACIÓN DE REPORTES. 61
FIGURA 12. REPORTE DEL TOTAL DE ESTUDIANTES POR PERÍODO ACADÉMICO. 61
FIGURA 13. PANTALLA CON REPORTE GRÁFICO DEL TOTAL DE ESTUDIANTES..... 62

ÍNDICE DE TABLAS

TABLA 1. DIFERENCIAS ENTRE UN DATAWAREHOUSE Y UN DATAMART. .. 22
TABLA 2. LISTA DE PROCESOS. 31
TABLA 3. DESCRIPCIÓN DE LA TABLA TB_ORGANIZACION. 35
TABLA 4 .LISTA DE CASOS DE USO DEL SISTEMA..... 37
TABLA 5. DESCRIPCIÓN DE INDICADORES..... 41
TABLA 6. COMPARACIÓN DE REQUERIMIENTOS DE LOS MANEJADORES.. 43
TABLA 7. COMPARACIÓN DE CARACTERÍSTICAS DE LOS MANEJADORES. 43
TABLA 8. RESULTADO DE LA EJECUCIÓN DE CARGA DE LOS PERÍODOS.. 56

SINOPSIS

En este trabajo especial de grado se desarrolló un Sistema de Apoyo a la Toma de Decisiones para la Escuela de Ingeniería Informática de la Universidad Católica Andrés Bello.

Se diseñó e implementó un datamart cuyas fuentes de datos fueron archivos proporcionados por las distintas coordinaciones de la escuela y otros reportes del Sistema de Reportes Banner de la UCAB. Se determinaron cuales eran las principales necesidades de información por parte de las coordinaciones y por parte de la dirección de escuela al momento de realizar la toma de decisiones.

Como parte del sistema de apoyo a la toma de decisiones se evaluaron las herramientas de reportes y quedó seleccionada una herramienta web (desarrollada en JSP) para el manejo dinámico de los distintos reportes que debe ofrecer la aplicación.

Es importante destacar que se usó como referencia la metodología de otro Trabajo Especial de Grado (TEG), la misma permite a través de la combinación de modelos metodológicos obtener una serie de etapas que deben ser llevadas a cabo para lograr el desarrollo e implantación exitosa de un sistema de apoyo a la toma de decisiones.

Para una mejor comprensión de las necesidades de información se realizaron las documentaciones respectivas de los procesos que son llevados a cabo en la escuela y a partir de la misma documentación se realizó una integración de las aplicaciones transaccionales de forma tal que pueda servir de fuente al sistema de apoyo a la toma de decisiones.

INTRODUCCIÓN

Los Sistemas de Apoyo a las Decisiones (DSS) se pueden definir como un conjunto de programas y herramientas que permiten obtener oportunamente la información requerida durante el proceso de la toma de decisiones, en un ambiente de incertidumbre.

En general, podemos decir que un DSS es un sistema informático utilizado para servir de apoyo, más que automatizar, el proceso de toma de decisiones. La decisión es una elección entre alternativas basadas en estimaciones de los valores de esas alternativas. El apoyo a una decisión significa ayudar a las personas que trabajan solas o en grupo a reunir inteligencia, generar alternativas y tomar decisiones. Apoyar el proceso de toma de decisión implica el apoyo a la estimación, la evaluación y/o la comparación de alternativas.

Las necesidades de información se agrupan según las unidades organizativas de la institución y las aplicaciones que cada una de ellas lleve a cabo. Resulta importante la necesidad de información sobre el entorno, implicando un mecanismo de observación que provea constantemente información relativa a los principales factores estratégicos: competencia, tecnología y política, entre otros.

Resumiendo, tenemos que el propósito de este trabajo especial de grado es desarrollar un Sistema de Apoyo a la Toma de Decisiones para la Escuela de Ingeniería Informática de la Universidad Católica Andrés Bello, a través del cual los directivos y coordinadores puedan obtener reportes que los ayuden a conocer la información crítica en el momento en que surja la necesidad de tomar una decisión dentro de su respectiva unidad académica, a través una interfaz de usuario amigable

Para la realización de este trabajo especial de grado se va a utilizar como referencia una metodología realizada por otro TEG en el cual también se elaboró un sistema de apoyo a la toma de decisiones, la metodología resultó ser muy exitosa y es por ello que se ha decidido aplicarla. Esta metodología contempla el conjunto de etapas del modelo metodológico de cascada pero en cierto punto de la misma se realiza un modelo iterativo en el cual se va a ir construyendo el almacén de datos final.

En el capítulo I se muestra el problema original que dió origen al presente trabajo, también se podrá observar el objetivo general, los objetivos específicos, la justificación y el alcance.

En el capítulo II se muestra el marco teórico, en el cual se presentan los fundamentos teóricos investigados y que servirán de apoyo al desarrollo.

En el capítulo III se muestra la metodología, explicada detalladamente para poder servir de guía durante la realización del trabajo.

En el capítulo IV se muestra el desarrollo a través del seguimiento de la metodología utilizada.

En el capitulo V se muestran los resultados obtenidos del desarrollo, así como las conclusiones y recomendaciones realizadas para que el mismo sea exitoso. Por último se presentan los apéndices que muestran todos los documentos obtenidos como producto de la realización de este trabajo.

CAPÍTULO I. PROBLEMA Y OBJETIVO

I.1 PLANTEAMIENTO DEL PROBLEMA

Por definición, la Ingeniería Informática es la rama de la ingeniería que consiste en la aplicación de la tecnología a la construcción de equipos capaces de procesar y almacenar la información de manera automática. La ingeniería informática es una carrera que trata sobre el manejo automatizado mediante computadores para la toma de decisiones de las organizaciones, gracias a esto es posible que la información llegue a dichas organizaciones en forma oportuna, actualizada y segura.

La Escuela de Ingeniería Informática de la Universidad Católica Andrés Bello, es la unidad administrativa encargada de la formación del estudiante en los fundamentos de la informática, la tecnología del software, la tecnología de las computadoras, la ingeniería del software, así como la tecnología de las redes de computadoras. Su principal objetivo es formar profesionales con sólidos conocimientos en las áreas de Ciencias Básicas, Ciencias de la Computación, Ingeniería del Software y Telemática; con un alto sentido crítico y una base gerencial que les permita desenvolverse como ingenieros en un mundo de cambios constantes.

La estructura organizacional está conformada por la dirección de escuela, quién a su vez gestiona un conjunto de coordinaciones, cada una de las cuales es responsable de un conjunto de áreas relacionadas con la carrera. Actualmente la Escuela de Ingeniería Informática cuenta con una serie de sistemas, algunos desarrollados internamente y otro propios de la universidad que son utilizados por las distintas facultades para la administración académica, sin embargo a nivel de las coordinaciones y de la dirección, no se cuenta con un sistema que apoye las decisiones que son tomadas en cada uno de los períodos academicos. En la actualidad la

mayoría de las decisiones son tomadas empíricamente, es decir, no es utilizada la información histórica que se tiene, además que dicha información no ha sido validada y no garantiza confianza a la hora de ser utilizada.

Es necesario para las unidades administrativas, (coordinaciones y dirección general) disponer de una documentación de los procesos de gestión que son llevados a cabo por las mismas; también se necesita definir un proceso que permita la integración de los sistemas internos y que evite las discrepancias entre éstos, garantizando mayor confianza en la información que allí se encuentra. Por último, es necesario un sistema que apoye a la toma de decisiones en el área administrativa; dicho sistema se basará en una serie de indicadores que se definirán mediante el proceso de levantamiento de información, la información que se utilizará garantizará confianza, y será obtenida de los sistemas internos y externos con los que cuenta la Escuela de Ingeniería Informática.

I.2 OBJETIVO GENERAL

Desarrollar un sistema de apoyo a la toma de decisiones para la Escuela de Ingeniería Informática de la Universidad Católica Andrés Bello.

I.3 OBJETIVOS ESPECÍFICOS

I.3.1 Integrar las aplicaciones transaccionales existentes en la Escuela de Ingeniería Informática de la Universidad Católica Andrés Bello.

I.3.2 Definir y documentar los procesos administrativos de la Escuela de Ingeniería Informática de la Universidad Católica Andrés Bello.

I.3.3 Definir y determinar a través del proceso de levantamiento de información los indicadores necesarios para apoyar el proceso de toma de decisiones.

I.3.4 Definir e Implementar los procesos de extracción, transformación y transporte de datos de las fuentes de datos a utilizar (Procesos ETT).

I.3.5 Diseñar e implementar el Almacén de Datos (Datamart) que servirá como fuente de datos para el sistema de apoyo para la toma de decisiones.

I.3.6 Evaluar y seleccionar la plataforma de desarrollo (Manejador de Base de Datos, herramientas para los procesos de Extracción, Transformación y Transporte).

I.3.7 Evaluar y seleccionar la(s) herramienta(s) de usuario final con la que se desarrollarán los reportes finales del sistema para la toma de decisiones.

I.3.8 Desarrollar un sistema de apoyo a la toma de decisiones basado en los indicadores, el datamart, los reportes y los procesos ETT.

I.4 ALCANCE

I.4.1 El sistema de apoyo a la toma de decisiones generará reportes basados en los indicadores que se determinarán en el levantamiento de información y con la información extraída del Datamart.

I.4.2 Desarrollo e Implementación del Datamart que almacenará la información a ser utilizada por el Sistema de apoyo a la toma de decisiones.

I.4.3 Realizar la integración de las aplicaciones transaccionales de la Escuela de Ingeniería Informática.

I.5 JUSTIFICACIÓN

Cualquier empresa que desee lograr un alto nivel de competitividad en el mundo actual debe disponer de un sistema que apoye aquellas decisiones que sean tomadas a nivel estratégico, táctico y operativo; un sistema de esa envergadura ayuda a cualquier empresa a responder de forma rápida y confiable a las condiciones cambiantes del mercado y en particular a las necesidades de sus clientes.

Para que un sistema de apoyo a la toma de decisiones sea exitoso en la empresa, la información que debe manipular debe ser de buena calidad, es decir, la información obsoleta, inexacta o muy difícil de entender no será de utilidad ni tendrá ningún valor para ningún profesional de los negocios.

En su meta de ofrecer al mercado profesionales exitosos y capaces de responder a las necesidades del mercado la escuela de Ingeniería Informática de la Universidad Católica Andrés Bello necesita a nivel interno disponer de herramientas que les permitan realizar una buena gestión de recursos y tener sustento en informaciones históricas para la elección de una buena decisión que les permita cumplir su meta y continuar en su lucha por innovar, optimizar y mejorar la calidad de los conocimientos que son impartidos a los futuros profesionales del área de la informática.

En la escuela la población estudiantil ha aumentado potencialmente en los últimos años (se estima que seguirá creciendo aún más de forma considerable), es por esto que se hace necesario mantener una ventaja competitiva respecto a las demás universidades que ofrecen esta carrera a nivel nacional. Para ello, contar con un sistema de apoyo a la toma de

decisiones generará información necesaria para desarrollar estrategias para mejorar los procesos de gestión, ofrecerá información confiable para dar soporte a las decisiones. Motivados por las razones expuestas y por los beneficios que ofrecen los sistemas de apoyo a la toma de decisiones para cualquier organización es que se ha decidido desarrollar e implementar este proyecto de innovación dentro de esta institución universitaria.

Durante la realización de las etapas del desarrollo del sistema será necesaria la aplicación de todos los conocimientos adquiridos a lo largo de nuestra carrera universitaria y de nuestras experiencias personales, así como también será necesario realizar investigaciones que nos orienten a la utilización de nuevas metodologías y tecnologías utilizadas en el mercado actual; permitiéndonos ampliar nuestros conocimientos en el área de sistemas de información, una de las ramas de la tecnología que está siendo muy utilizada por las empresas a nivel mundial, convirtiéndonos así en profesionales más competitivos.

CAPÍTULO II. MARCO TEÓRICO

II.1 SISTEMAS DE APOYO A LA TOMA DE DECISIONES.

II.1.1 Sistemas de Apoyo a la Toma de Decisiones [8]

Son sistemas de información basados en computadora que proporcionan apoyo de información interactiva a administradores y profesionales durante el proceso de negocio. Son usados por los responsables de las tomas de decisiones a nivel administrativo estratégico, ya que éstos les proporcionan informes, pronósticos resumidos, específicos y no programados e inteligencia extrema para apoyar sus responsabilidades no estructuradas de planeación y diseño de políticas. Los sistemas de apoyo a la toma de decisiones usan:

- Modelos Analíticos.
- Bases de Datos Especializadas.
- Las propias percepciones y juicio de quién toma las decisiones.
- Un proceso de modelación interactivo basado en computadora para apoyar la toma de decisiones semiestructuradas de negocios.

II.1.1.1 *Componentes de los Sistema de Apoyo a la Toma de Decisiones* [1,8]

A diferencia de los sistemas de información administrativa, los sistemas de apoyo a la toma de decisiones dependen de los siguientes componentes:

Base de Modelo

Una base de modelos para un sistema de apoyo a la toma de decisiones es un componente de software que consiste en modelos usados

en rutinas computacionales y analíticas que expresan matemáticamente las relaciones entre variables.

Base de Datos

Una base de datos es un conjunto de elementos relacionados entre sí que reflejan o modelan la información de una organización. Quiere decir que el contenido de cualquier base de datos refleja una realidad significativa para alguien.

II.1.2 Tipos de Decisiones [8]

- Decisiones Estructuradas: son aquellas decisiones que se toman al nivel de la administración operativa. Implican situaciones en las que se pueden especificar los procedimientos a seguir por adelantado cuando se necesita tomar una decisión.

- Decisiones No Estructuradas: son aquellas decisiones que se toman al nivel de la administración estratégica. Conllevan situaciones de decisión en las que no es posible especificar por adelantado la mayoría de los procedimientos de decisión a seguir.

- Decisiones Semiestructuradas: son aquellas decisiones que se toman al nivel de la administración táctica. Se pueden predefinir algunos procedimientos de decisión, pero no lo suficiente como para conducir a una decisión recomendada definitiva.

II.1.3 Sistemas de soporte a la decisión en la educación, EDSS (Educational Decision Support System) [12]

Normalmente cuando hablamos de DSS lo enfocamos a grandes empresas o a industrias que son muy complejas. Sin embargo, la aplicación

de sistemas de soporte ha crecido en gran manera en los últimos años siendo una herramienta aplicable a otros sectores de la sociedad como lo es: La Educación.

La educación es una noble tarea cuya materia prima son seres humanos. En este sentido, la toma de decisiones adquiere una verdadera importancia, pues las decisiones que tomemos ahora, afectarán o beneficiarán a las futuras generaciones. La introducción de sistemas de soporte a la decisión en el sector educativo facilitará la toma de decisiones, potenciando la capacidad de liderazgo de la autoridad educativa sobre el elemento humano y organizacional.

A partir de 1992, con el Acuerdo Nacional para la Modernización de la Educación Básica (ANMEB), se descentralizó el sistema educativo a los gobiernos estatales, estableciendo una plataforma como punto de partida para elevar la calidad de la educación.

Esta plataforma supone que las decisiones estratégicas y operativas del sector educativo deberán apoyarse en un sistema de soporte a la toma de decisión, bautizado en el medio educativo como sistema de información, que debe estar alimentado con información precisa y oportuna del desempeño del sistema educativo y basado en indicadores de equidad, eficiencia y aprovechamiento. En este sentido, las entidades son responsables de diseñar mecanismos que aseguren la generación sistemática de información, relativa al desempeño de sus sistemas educativos con la precisión y oportunidad suficientes para que las autoridades de cualquier ámbito de gobierno respondan adecuadamente para corregir deficiencias y reforzar aciertos.

Considerando el teorema de la administración: "Lo no mensurable, no es administrable", entonces para administrar con calidad un sistema, es indispensable hacerlo sobre la base de información veraz y oportuna acerca de su desempeño.

La información adquiere valor real cuando, en forma de indicadores, es empleada para diagnosticar el desempeño del sistema educativo y, con ello, establecer políticas que impacten la calidad de la educación. Cualquier política educativa que se asuma sin tener información suficiente, será de poca o ninguna utilidad para el sistema.

El valor de la información como elemento de apoyo a la toma de decisiones, crece con su nivel de detalle, porque permite diagnosticar con mayor precisión el comportamiento del sistema educativo. Si la autoridad educativa desea tomar decisiones de calidad, debe preocuparse por conformar bases de datos a nivel individual, es decir, estudiante, profesor o escuela, que le permitan este máximo grado de detalle para el tomador de decisiones.

II.1.4 Sistemas de información en la educación [12]

Un sistema educativo tiene como propósito facilitar el aprendizaje de los miembros de una sociedad. Por tanto en la educación, un sistema de información tiene como propósito servir de mecanismo informativo para que todos los componentes del sistema educativo, tomen la mejor decisión en el momento exacto. Así pues, la intención de un sistema de información en la educación debería de:

- Promover la construcción de sistemas integrados que distribuyan información y conocimiento para mejorar la toma de decisiones.

- Conformar una sola instancia generadora de información, como fuente única que elimine la duplicidad de esfuerzos entre distintas áreas de la organización.
- Incrementar la productividad del personal administrativo y docente.
- Fomentar el trabajo en grupo compartiendo información.
- Presentar oportunamente información relevante para la toma de decisiones a distintos niveles de la autoridad educativa.
- Formalizar el proceso de monitoreo de las variables críticas, que impactan directamente en la calidad de la educación.

El impacto en el sector educativo, de un sistema de información como apoyo a la toma de decisiones, consistiría en primer término, en que los altos ejecutivos y tomadores de decisión tendrían más tiempo para la planeación estratégica y la función docente. Adicionalmente brindaría los siguientes beneficios:

- La automatización de decisiones rutinarias o fases del proceso de toma de decisiones.
- Menor experiencia requerida para tomar decisiones.
- Toma de decisiones más rápida.
- Menos apoyo en expertos para proveer soporte a las altas autoridades.
- Redistribución del poder entre los directivos y tomadores de decisiones.
- Soporte a decisiones complejas, tomándolas más rápido y de mejor calidad.
- Contar con información pertinente para decisión de alto nivel.

II.2 INDICADORES

II.2.1 Definición [13]

Es un dato que pretende reflejar el estado de una situación, o de algún aspecto particular, en un momento y un espacio determinados. Habitualmente se trata de un dato estadístico (porcentajes, tasas, razones...) que pretende sintetizar la información que proporcionan los diversos parámetros o variables que afectan a la situación que se quiere analizar.

II.2.2 Requisitos de un buen indicador [14]

Un buen indicador debe abarcar, como sea posible, el mayor número de las siguientes características:

- Disponibilidad: los datos básicos para la construcción del indicador deben ser de fácil obtención sin restricciones de ningún tipo.
- Simplicidad: el indicador debe ser de fácil elaboración.
- Validez: la validez de los indicadores significa que éstos deben tener la capacidad de medir realmente el fenómeno que se quiere medir y no otros.
- Especificidad: si un indicador no mide realmente lo que se desea medir, su valor es limitado, pues no permite la verdadera evaluación de la situación al reflejar características que pertenecen a otro fenómeno paralelo.
- Confiabilidad: los datos utilizados para la construcción del indicador deben ser fidedignos (fuentes de información satisfactorias).
- Sensibilidad: el indicador debe ser capaz de poder identificar las distintas situaciones aún en áreas con distintas particularidades, independientemente de la magnitud que ellas tengan en la comunidad.

- Alcance: el indicador debe sintetizar el mayor número posible de condiciones o de distintos factores que afectan la situación descrita por dicho indicador. En lo posible el indicador debe ser globalizador.

II.2.3 Los indicadores educativos. [21]

La necesidad de definir y medir las cuestiones relacionadas con el desarrollo, los niveles de vida y las condiciones sociales y económicas motivó la aparición, desde los años setenta, de diversos programas de indicadores sociales elaborados por organismos internacionales y países. En concreto los educativos, destinados a comparar y juzgar el contexto y el funcionamiento de la enseñanza y sus resultados, han experimentado un gran desarrollo.

En 1987 el Centro para la Investigación e Innovación Educativa (CERI) en cooperación con la Unidad de Indicadores y Estadísticas de la Educación de la OCDE (Organización para la Cooperación y el Desarrollo Económicos) elaboró un sistema internacional de indicadores de la situación de la educación, la estructura general se mantiene en torno a cuatro grupos: contexto, recursos, proceso y resultados.

Los indicadores de contexto sirven para caracterizar las condiciones demográficas, socioeconómicas y culturales en las que se desarrolla la educación. El primero es el nivel general de formación de la población que está relacionado con las capacidades y competencia de la mano de obra, con la creación de puestos de trabajo y con el desarrollo económico y social de los países.

Los indicadores de recursos evalúan los recursos materiales y los humanos. Por un lado, las inversiones económicas y los gastos de

infraestructura y, por otro, el número de alumnos por niveles y la formación de profesores, remuneración, características, etc.

Los indicadores de procesos son los más recientes e informan sobre la organización y funcionamiento de los centros, práctica educativa y clima escolar. Están más relacionados con aspectos cualitativos que cuantitativos, de ahí que sean los más difíciles de utilizar.

Los indicadores de resultados manifiestan los logros del sistema educativo a través del éxito escolar de los alumnos mediante las pruebas de evaluación, certificados y títulos que consiguen y su inserción en el mercado del trabajo.

II.3 ALMACÉN DE DATOS (DATA WAREHOUSE)

II.3.1 Definiciones [7,15,16,18]

- Es un conjunto de datos orientado a temas, integrado, no volátil, variante en el tiempo, como soporte a la toma de decisiones de dirección.

- Es la integración de datos consolidados, almacenados en un dispositivo de memoria no volátil, proveniente de múltiples y posiblemente diferentes fuentes de datos. Con el propósito del análisis y a partir de este tomar decisiones en función de mejorar la gestión del negocio. Contiene un conjunto de cubos de datos que permiten a través de técnicas de OLAP consolidar, ver y resumir los datos acorde a diferentes dimensiones de estos. **[20]**

Un almacén de datos guarda datos que se han extraído desde diversas bases de datos operativas, externas y de otro tipo de una

organización. Es una fuente central de datos que han sido limpiados, transformados y catalogados, de tal manera que los administradores y otros profesionales de negocios puedan utilizarlos para la minería de datos, el procesamiento analítico en línea y otras formas de análisis de negocio, investigación de mercados y apoyo a la toma de decisiones.

Según se definió en **[4]**, un data warehouse se caracteriza por ser:

- Integrado: los datos almacenados en el data warehouse deben integrarse en una estructura consistente, por lo que las inconsistencias existentes entre los diversos sistemas operacionales deben ser eliminadas. La información suele estructurarse también en distintos niveles de detalle para adecuarse a las distintas necesidades de los usuarios.

- Temático: sólo los datos necesarios para el proceso de generación del conocimiento del negocio se integran desde el entorno operacional. Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del data warehouse. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar.

- Histórico: el tiempo es parte implícita de la información contenida en un data warehouse. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. Por el contrario, la información almacenada en el data warehouse sirve, entre otras cosas, para realizar análisis de tendencias. Por lo

tanto, el data warehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.

- No volátil: el almacén de información de un data warehouse existe para ser leído, pero no modificado. La información es por tanto permanente, significando la actualización del data warehouse la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

II.3.2 Componentes de los almacenes de datos [9]

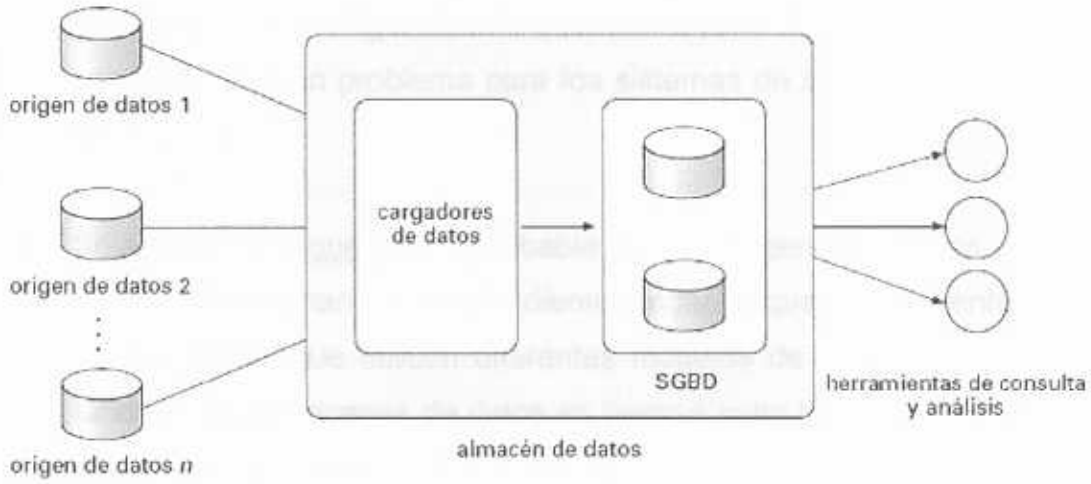


Figura 1. Arquitectura de los almacenes de datos.
Fuente: [9]

La Figura 1 muestra la arquitectura de un almacén de datos típico e ilustra la recogida de los datos, su almacenamiento y el soporte de las consultas y del análisis de datos. Entre los problemas que hay que resolver al crear un almacén de datos están los siguientes:

- Momento y modo de la recogida de datos. En una arquitectura dirigida por los orígenes para la recogida de los datos, los orígenes de los

datos transmiten la información nueva, bien, de manera continua (a medida que se produce el procesamiento de las transacciones) o de manera periódica (de noche, por ejemplo). En una arquitectura dirigida por el destino, el almacén de datos envía de manera periódica solicitudes de datos nuevos a los orígenes de datos.

A menos que las actualizaciones de los orígenes de datos se repliquen en el almacén de datos mediante un compromiso de dos fases, el almacén de datos nunca estará actualizado respecto a los orígenes de datos. El compromiso de dos fases suele resultar demasiado costoso para ser una opción aceptable, por lo que los almacenes de datos suelen tener datos ligeramente desactualizados. Eso, no obstante, no suele suponer un problema para los sistemas de ayuda a la toma de decisiones.

- Selección del esquema. Es probable que los orígenes de datos que se han creado de manera independiente tengan esquemas diferentes. De hecho, puede que utilicen diferentes modelos de datos. Parte de la labor de los almacenes de datos es llevar a cabo la integración de los esquemas y convertir los datos al esquema integrado antes de almacenarlos. En consecuencia, los datos almacenados en el almacén de datos no son una mera copia de los datos de los orígenes de datos. Por el contrario, se pueden considerar como una vista materializada de los datos de los orígenes de datos.

- Limpieza de los datos. La labor de corregir y realizar un procesamiento previo de los datos se denomina limpieza de los datos. Los orígenes de datos suelen entregar datos con numerosas inconsistencias de carácter menor, que pueden corregirse. Por ejemplo, los nombres suelen estar mal escritos y puede que las direcciones tengan mal

escritos los nombres de la calle, del distrito o de la ciudad, o puede que los códigos postales se hayan introducido de manera incorrecta. Esto puede corregirse en un grado razonable consultando una base de datos de los nombres de las calles y de los códigos postales de cada ciudad. Las listas de direcciones recogidas de varios orígenes pueden tener duplicados que haya que eliminar en una operación de mezcla-purga.

- Propagación de las actualizaciones. Las actualizaciones de las relaciones en los orígenes de datos deben propagarse a los almacenes de datos. Si las relaciones en los almacenes de datos son exactamente las mismas que en los orígenes de datos, la propagación es directa.

- Resúmenes de los datos. Los datos brutos generados por un sistema de procesamiento de transacciones pueden ser demasiado grandes para almacenarlos en línea. No obstante, se pueden responder muchas consultas manteniendo únicamente datos resumen obtenidos por agregación de las relaciones, en lugar de mantener las relaciones enteras. Por ejemplo, en lugar de almacenar los datos de cada venta de ropa, se pueden almacenar las ventas totales de ropa por nombre de artículo y por categoría.

II.3.3 Beneficios de un datawarehouse [9,17]

- Proporciona una herramienta para la toma de decisiones en cualquier área funcional, basándose en información integrada y global del negocio.

- Facilita la aplicación de técnicas estadísticas de análisis y modelización para encontrar relaciones ocultas entre los datos del almacén; obteniendo un valor añadido para el negocio de dicha información.
- Proporciona la capacidad de aprender de los datos del pasado y de predecir situaciones futuras en diversos escenarios.
- Simplifica dentro de la empresa la implantación de sistemas de gestión integral de la relación con el cliente.
- Supone una optimización tecnológica y económica en entornos de Centro de Información, estadística o de generación de informes con retornos de la inversión espectaculares.

El almacenamiento y recuperación de los datos para la ayuda a la toma de decisiones plantea varios problemas:

- Aunque muchas consultas para ayuda a la toma de decisiones pueden escribirse en SQL, otras no pueden expresarse en SQL o no pueden hacerlo con facilidad. En consecuencia, se han propuesto varias extensiones de SQL para facilitar el análisis de los datos. El área de procesamiento analítico en línea (Online Analytical Processing, OLAP) trata de las herramientas y de las técnicas para el análisis de los datos que pueden dar respuestas casi instantáneas a las consultas que soliciten datos resumidos, aunque la base de datos sea extremadamente grande.

- Los lenguajes de consulta de bases de datos no resultan adecuados para el rendimiento de los análisis estadísticos detallado de los datos. Hay varios paquetes, como SAS y S++, que ayudan en el análisis estadístico. A estos paquetes se les han añadido interfaces con las bases de datos para permitir que se almacenen en la base de datos grandes volúmenes de datos y se recuperen de manera eficiente para su análisis.

- Las técnicas de búsqueda de información intentan descubrir de manera automática las reglas y las pautas estadísticas de los datos. El campo de la recopilación de datos combina las técnicas de búsqueda de información creadas por los investigadores en inteligencia artificial y los expertos en análisis estadístico con las técnicas de implementación eficiente que permiten utilizarlas en bases de datos extremadamente grandes.

- Las grandes empresas tienen varios orígenes de datos que necesitan utilizar para adoptar decisiones empresariales. Los orígenes pueden almacenar los datos según diferentes esquemas. Por motivos de rendimiento (así como por motivos de control de la organización) los orígenes de datos no suelen permitir que otras partes de la empresa recuperen datos a petición.

- Para ejecutar de manera eficiente las consultas sobre datos tan diferentes las empresas han creado almacenes de datos. Los almacenes de datos reúnen los datos de varios orígenes bajo un esquema unificado en un solo sitio. Por tanto, ofrecen al usuario una sola interfaz uniforme para los datos.

II.4 DATAMART

II.4.1 Definiciones [17,18]

- Según define **[2]**, "un datamart es un pequeño datawarehouse diseñado para una unidad estratégica del negocio (SBU), o un nivel departamental". Los datamarts, tienen las mismas características de integración, no volatilidad, orientación temática y no volatilidad que el datawarehouse. Representan una estrategia de "divide y vencerás" para ámbitos muy genéricos de un datawarehouse.

- Es un subconjunto del DataWarehouse, usado normalmente para el análisis parcial de los datos. El objetivo de subdividir está dado por la complejidad computacional del análisis global de todas las dimensiones del DataWarehouse y por la necesidad de rapidez. **[6]**

La función principal de un datamart es apoyar a otros sistemas para la toma de decisiones, además de disponer de la información solicitada en el momento preciso **[3]** , permitiendo un mejor control de aquella que se abarca. Esta estrategia es particularmente apropiada cuando el datawarehouse central crece muy rápidamente y los distintos departamentos requieren sólo una pequeña porción de los datos contenidos en él.

II.4.2 Diferencias entre un DataWarehouse y un DataMart

Un Datamart es exactamente lo mismo que un DataWarehouse, siendo las diferencias más significativas las relacionadas con el alcance y el contenido de información. En la Tabla 1 se pueden apreciar dichas diferencias.

Características	DataMart	DataWarehouse
Alcance	Departamento	Toda la organización
Temas	Único	Muchos
Fuentes de Datos	Pocas	Muchas
Tamaño Promedio	Menos de 100 GB	Más de 100 GB
Tiempo de Implementación	Meses	Meses a Años

Tabla 1. Diferencias entre un Datawarehouse y un datamart.

Fuente: Adaptación de [1].

CAPÍTULO III. MARCO METODOLÓGICO

Para el desarrollo de cualquier software, es necesario usar una metodología ya que ésta permite mostrar un seguimiento lógico y entendible de lo que se desea desarrollar. Realizar el desarrollo de cualquier software usando una metodología, nos permite mantener un orden y estructura sobre cómo hacer el proceso, además permite concentrarnos en la lógica y el giro del negocio al que nos abarcamos y nos permitirá tener una guía para saber cuándo y de qué manera obtener los requerimientos de los clientes.

Para la realización del Sistema de Apoyo a la Toma de Decisiones se adapto una metodología que fue utilizada en otro Trabajo Especial de Grado (Berroterán, D. y Gamboa, E. (2007). *Sistema de Información de Apoyo para la toma de decisiones relacionadas con la facturación, pedidos y necesidades de recursos humanos de la empresa Flamberge*. Trabajo de grado, Ingeniería Informática, Universidad Católica Andrés Bello, Caracas) y resultó exitoso su uso, aplicaremos algunas modificaciones para adaptarla a las necesidades del proyecto de forma tal que cumpla con los aspectos necesarios para la realización del proyecto.

La metodología está basada en la combinación de dos modelos metodológicos muy reconocidos, uno de ellos es cascada del cual tomamos las etapas básicas de forma tal que proporcionen una línea base y permitan establecer cada una de las fases que deberán ser realizadas en el proyecto, el otro modelo utilizado fue entrega evolutiva del cual tomamos las fases iterativas para la creación del almacén de datos, dichas fases fueron incorporadas en conjunto con el modelo de cascada para formar la metodología que será utilizada en la realización del proyecto. El detalle de la metodología, fue tomado de un TEG el cual a su vez basó cada una de las

etapas en el ejemplo del libro de McConnell Steve (1996), titulado Desarrollo y Gestión de Proyectos Informáticos.

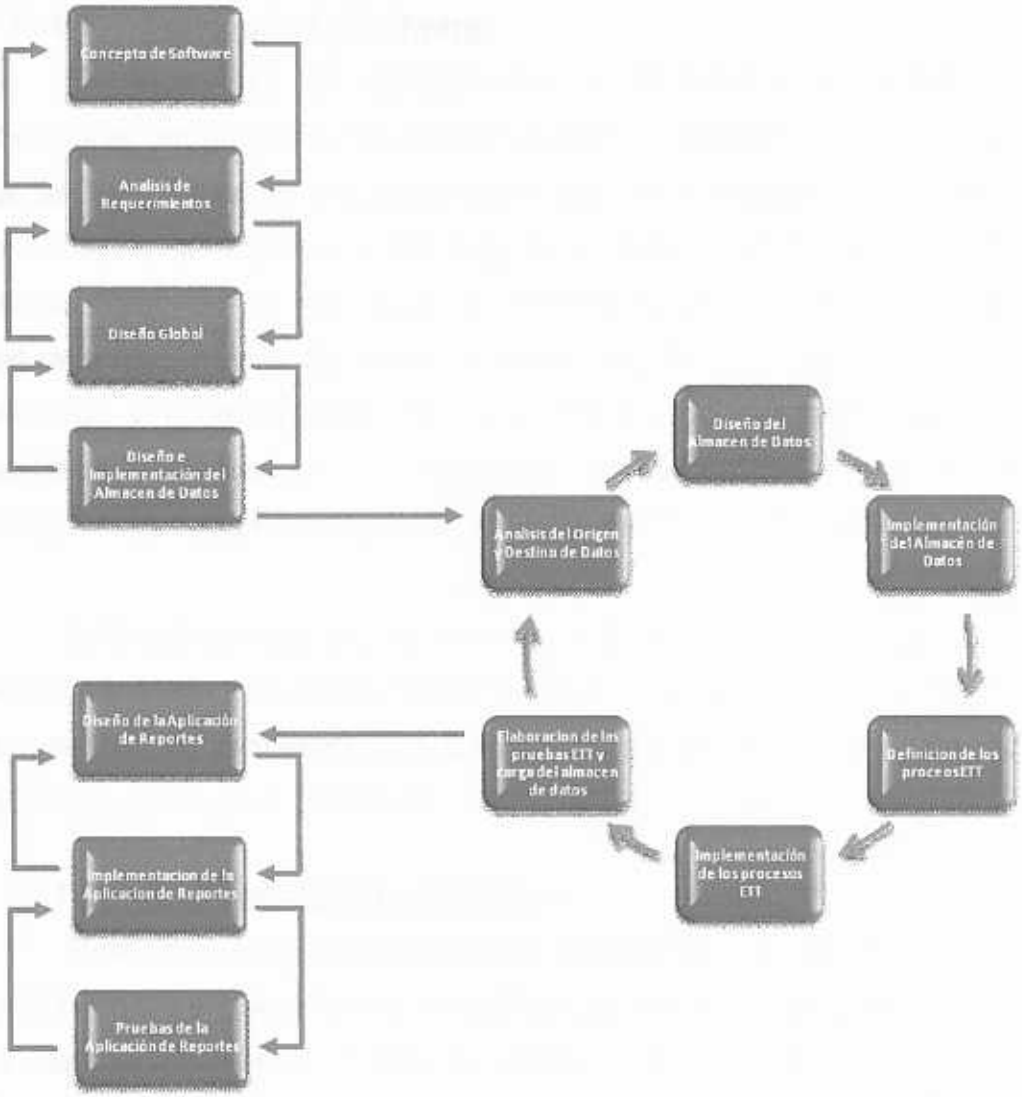


Figura 2. Modelo del ciclo de desarrollo elaborado para la realización del sistema de apoyo para la toma de decisiones.

Fuente: Adaptación de [19].

III.1 FASES DEL MODELO [19]

III.1.1 Concepto de Software

En esta fase se recogen los requerimientos funcionales y no funcionales del proyecto. La idea de esta fase es tener en claro el proyecto que se va a construir. Se debe tener una visión general del sistema y, a través del levantamiento de información se deben establecer los elementos que conformarán el mismo. El levantamiento de información puede ser realizado por medio de entrevistas con el usuario, entrevistas con los directivos y coordinadores de la institución, documentación que permita establecer criterios en la realización del proyecto y la creación de cuestionarios que permitan planificar el desarrollo de las entrevistas.

Esta fase se realiza para entender el problema y buscar soluciones, en ella se hace una lista de los requerimientos, la lista de los casos de uso que representan los requerimientos funcionales del proyecto y de esta manera saber lo que se quiere construir.

III.1.2 Análisis de Requerimientos

Esta fase tiene como propósito determinar qué se va a construir y cuáles son los elementos que intervienen en el sistema, así como los detalles de sus funcionalidades, y de esta manera poder obtener una descripción clara de qué sistema se va a desarrollar. Ésta se basa en la lista de requerimientos y lista de casos de uso de la fase anterior, para realizar los distintos documentos que servirán para realizar el diseño del sistema. Adicionalmente se pueden realizar las especificaciones de casos de uso, los modelos de casos de uso, es decir, la documentación que sea necesaria para lograr una mejor comprensión sobre lo que desea el usuario y la manera de cómo lograrlo. En la fase de análisis del sistema de soporte para la toma de decisiones se incorporaron adicionalmente otras actividades que son

necesarias realizar para cubrir las necesidades del proyecto. Estas actividades son las siguientes:

III.1.2.1 *Identificación de indicadores*

En esta actividad se seleccionarán los indicadores del sistema según los requerimientos obtenidos en la fase de concepto de software.

III.1.2.2 *Evaluación y selección de la herramienta de desarrollo*

Se analizarán las distintas herramientas disponibles en el mercado que permitan la realización del proyecto, estas herramientas serán comparadas y evaluadas para seleccionar una o un conjunto de ellas que soporten la estructura del proyecto y que mejor se ajusten a las características de hardware con las que cuenta la institución para la cual se desarrollará el sistema.

III.1.3 Diseño Global

El propósito de esta fase es elaborar el diseño global de la arquitectura del sistema completo que se va a desarrollar.

III.1.4 Diseño e Implementación del almacén de datos

Esta fase se dividió a su vez en varias fases que pueden realizarse iterativamente, hasta obtener el diseño e implementación del almacén de datos que mejor se ajuste a los requerimientos del proyecto.

III.1.4.1 *Análisis del origen y destino de datos.*

El propósito de esta fase es determinar cuáles son las posibles fuentes de datos, así como el análisis del destino de los datos que se utilizarán en el sistema de reportes.

III.1.4.2 *Diseño del almacén de datos*

En esta fase se identificarán las tablas de bases de datos de las distintas aplicaciones y los archivos de las diversas fuentes de datos que se utilizarán para extraer los datos que serán cargados en el almacén, además, se diseñará el modelo de datos necesario para almacenar la información sumarizada que, servirá como repositorio de datos para la generación de reportes de acuerdo a los requerimientos del sistema.

III.1.4.3 *Implementación del almacén de datos*

Se desarrollará el almacén de datos. Se procederá a implementar el almacén que se definió en la etapa de diseño.

III.1.4.4 *Definición de los procesos de Extraccion, Transformacion y Transporte (ETT).*

El propósito de esta fase es determinar los procesos que serán necesarios para extraer, transformar y cargar los datos en el almacén.

III.1.4.5 *Implementación de los procesos de Extraccion, Tranformacion y Transporte (ETT).*

Se codifican y programan los procesos de Extracción, Transformación y Transporte del sistema.

III.1.4.6 *Elaboración de las pruebas (ETT) y Carga del almacén de datos*

Esta fase surge por la necesidad de confirmar que los procesos de extracción implementados en la fase anterior son correctos y así, estar seguros de que los datos que se cargarán en el almacén son consistentes.

III.1.5 Diseño de la aplicación de reportes.

En esta fase se realizarán los diagramas que sean necesarios para comprender la aplicación de reportes a desarrollar y que sirvan de guía en el

proceso de implementación del proyecto. Además se diseñara la estructura de los reportes para la presentación de la información, así como las interfaces graficas de acceso a los reportes.

III.1.6 Implementación de la aplicación de reportes

El propósito de esta fase es desarrollar el código que permita crear la estructura adecuada de la aplicación junto con los reportes y las interfaces diseñadas.

III.1.7 Elaboración de las pruebas de la aplicación de reportes.

Esta fase consiste en asegurar el buen funcionamiento y calidad del producto a través de la aplicación de varios casos de pruebas al producto final, de esta manera se asegura que el sistema integrado por las interfaces y los reportes funcione correctamente. Además se realizan las pruebas con los usuarios del sistema.

Para uno de los productos del presente TEG que es la Integración de las Aplicaciones transaccionales existentes, se adaptó el modelo metodológico de cascada con retroalimentación. Algunas fases no fueron desarrolladas ya que esta metodología se utilizó solo para desarrollar el objetivo "*Integrar las aplicaciones transaccionales existentes en la Escuela de Ingeniería Informática de la Universidad Católica Andrés Bello*" por lo tanto no conlleva a desarrollar las etapas de Codificación y Depuración, Pruebas del Sistema e Implantación, sino al análisis de las aplicaciones existentes para rediseñar la nueva funcionalidad y base de datos de forma tal que se adapte a los procesos levantados y documentados, a continuación se detallarán cada una de sus etapas:

Figura 3 .Modelo del ciclo de desarrollo elaborado para la integración de las aplicaciones. Fuente: Adaptación de [4].

III.2 FASES DEL MODELO

III.2.1 Análisis de Requerimientos

En esta fase del modelo se analizan las necesidades de los usuarios finales de la aplicación con la finalidad de establecer que objetivos debe cubrir. En esta etapa se determinan todas las características que debe tener la aplicación final (resultado de la integración).

III.2.2 Diseño Global

El propósito de esta fase es elaborar el diseño global de la arquitectura del sistema completo resultado de la integración de las aplicaciones transaccionales.

III.2.3 Diseño Detallado

En esta fase se realizarán los diagramas que sean necesarios para comprender la aplicación y que sirvan de guía en el proceso de implementación del proyecto.

CAPÍTULO IV. DESARROLLO

En este capítulo se describe la elaboración de cada uno de los productos del trabajo especial de grado según las metodologías mencionadas en el capítulo III.

IV.1 INTEGRACIÓN DE LOS SISTEMAS TRANSACCIONALES DE LA ESCUELA DE INGENIERÍA INFORMÁTICA.

IV.1.1 Análisis de Requerimientos.

La fase de análisis de requerimientos representa la primera fase para realizar el proceso de integración que se desea realizar entre los distintos sistemas transaccionales de la Escuela de Ingeniería Informática. En esta etapa se realizó una fase de entrevistas en la cual básicamente los coordinadores especificaron como son las actividades que llevaban a cabo en su coordinación, cuáles eran los requisitos de cada actividad, las acciones que implicaban, etc. En dichas entrevistas también se consultó sobre la funcionalidad actual que conocían de los sistemas transaccionales y que funcionalidad nueva querían agregar a los sistemas de forma tal que les facilitará la realización de sus actividades. Como producto de esta etapa se obtuvieron los requerimientos de funcionalidad que deberían cubrir los sistemas transaccionales actuales, y una especificación de cada proceso realizado por cada coordinación. Luego de diseñado los procesos y la documentación, se realizaron entrevistas para confirmar que la información contenida en los documentos reflejaba los procesos y los datos necesarios para los mismos.

PROCESOS	
Servicio Comunitario.	Realización de Cursos.
Electivas.	Registro de actividades de los Beca-Trabajo.
Compra de Material	Selección de Preparadores/Ayudantes de Cátedra.
Préstamo de Material de Laboratorio	Aprobación de la propuesta de TEG.
Verificación de Inventario de Laboratorios.	Culminación de TEG
Pasantías	

Tabla 2. Lista de Procesos.

Fuente: Elaboración Propia.

IV.1.2 Diseño Global.

En esta etapa se determinaron los elementos que deberán ser tomados en cuenta a la hora de realizar la integración de los sistemas.

- Procesos de cada coordinación: determinar el funcionamiento actual de cada proceso que realiza cada coordinación dentro de la institución.

- Base de datos de los sistemas transaccionales: verificar la estructura de datos actual utilizada en los sistemas en producción.

- Sistemas transaccionales en producción: verificar la funcionalidad actual de los sistemas en producción en relación a los procesos levantados.

IV.1.3 Diseño Detallado.

En esta etapa se elaboraron los diagramas correspondientes a cada proceso de cada coordinación basándose en la información suministrada por los coordinadores de la escuela de ingeniería informática y por los reglamentos internos existentes. Se diseño el modelo de datos integrado que almacenará toda la información que es manipulada en los distintos sistemas transaccionales de la institución. Además, como parte de esta etapa se

generaron documentos de nomenclatura de bases de datos y un documento de funcionalidad de los sistemas transaccionales en producción.

IV.1.3.1 *Documentación de Procesos.*

En esta actividad se realizó la documentación detallada de cada proceso levantando en el diseño global, esta documentación permite obtener información de quién(es) es(son) el(los) responsable(s) de cada actividad a desarrollar en los procesos así como las fuentes que establecen que acciones debes realizarse en cada actividad. En el Apéndice C se muestra la documentación de los procesos.

En la Figura 4 se muestra el diagrama de flujo de procesos para la coordinación de materias electivas, en el mismo se observa el conjunto de pasos que se deben realizar durante el proceso de postulación y votación para determinar las materias electivas que serán dictadas en el período académico.

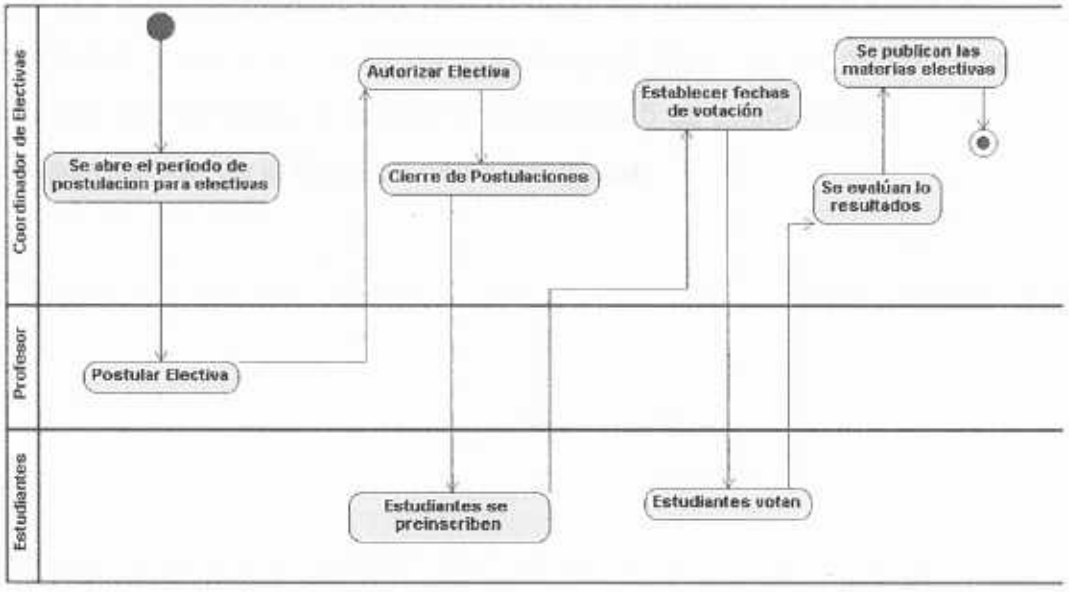


Figura 4. Diagrama de Flujo de Procesos de las Materias Electivas.

Fuente: Elaboración Propia.

IV.1.3.2 *Definición de Estándar de Nomenclatura de Base de Datos.*

Uno de los problemas observados en las etapas anteriores fue que las bases de datos utilizaban distintas nomenclaturas a pesar de pertenecer a la misma institución, esto genera inconsistencias y duplicación de tablas que son comunes. Por las razones antes mencionadas se decidió la realización de un estándar de nomenclatura que deben utilizar todas las bases de datos que se creen dentro de la Escuela de Ingeniería Informática, dicho estándar es mostrado en el Apéndice F.

IV.1.3.3 *Diseño de la Base de Datos Integrada.*

En esta etapa se procedió a la elaboración de los diagramas entidad - relación de cada sistema transaccional, dichos diagramas pueden ser vistos en el Apéndice B, estos diagramas contienen los datos que actualmente se manejan en los sistemas transaccionales. Se diseño un modelo entidad - relación con todos los datos necesarios para cubrir la integración de los sistemas transaccionales. En ese modelo se tomaron en cuenta los datos analizados y los nuevos datos obtenidos a partir de la definición de los procesos de las coordinaciones. En la Figura 5 se puede observar el modelo entidad-relación de la base de datos integrada.

IV.1.3.4 *Documentación de la Base de Datos Integrada.*

A partir de la información de las etapas anteriores y del modelo entidad-relación resultado de la integración de los sistemas transaccionales se realizó un diccionario de datos de las tablas del modelo, con la finalidad de que futuros desarrolladores pueden conocer el significado de las tablas y la funcionalidad que cumple dentro del nuevo sistema que será el resultado de la integración. En el Apéndice D se muestra la documentación de la base de datos integrada.

En la Tabla 3 se muestra un ejemplo de la documentación realizada a las tablas de la base de datos integrada.

NOMBRE	PRIMARIA	TIPO	CLAVE FORÁNEA	LONGITUD	NULL	COMENTARIOS
organizacionId	SI	Int		11	NO	Identificador de la organización que está realizando ofertas a los estudiantes.
nombre		Varchar		255	NO	Nombre de la organización.
direccion		Varchar		255	SI	Dirección fisica en la que se encuentra dicha empresa.
nombreContacto		Varchar		255	SI	Nombre de la persona de contacto dentro de la organización.
email		Varchar		255	SI	Correo electrónico de la persona de contacto.

Tabla 3. Descripción de la tabla tb_organizacion.

Fuente: Elaboración Propia.

IV.1.3.5 *Documentación de Funcionalidad.*

Con todo lo estudiado y obtenido en las fases anteriores se genera el documento de funcionalidad, que explica los cambios y/o mejoras que deben realizarse sobre los sistemas actuales de forma tal que logren quedar unificados, cumplan con las actividades de los procesos a los que pertenecen y permitan obtener información en estado consistente para los futuros

análisis. En el Apéndice E se muestra el documento de funcionalidad resultado de la integración de los sistemas.

IV.2 DESARROLLO DEL SISTEMA DE APOYO A LA TOMA DE DECISIONES

IV.2.1 Concepto de Software

En esta primera etapa del modelo para el desarrollo del sistema de apoyo a la toma de decisiones se realizaron diversas reuniones con las coordinaciones y con la dirección de la escuela, utilizando parte de la información obtenida para el análisis de los requerimientos de la integración de los sistemas transacciones. Se pudo determinar en líneas generales cuáles eran las necesidades de información para el momento de la toma de decisión en cada una de las coordinaciones y en la dirección de escuela.

Una vez realizada las entrevistas con las coordinaciones y con la dirección de la institución se pudo conocer que la principal necesidad de información hace referencia al rendimiento estudiantil en cada una de las áreas que mantienen un alto nivel de importancia para que el estudiante pueda culminar exitosamente su carrera. Estas áreas son principalmente: los trabajos especiales de grado, las pasantías y la realización del servicio social comunitario; aunado a esto son de interés para la dirección de la institución la información respecto a la población estudiantil, al conjunto de profesores, los materiales de laboratorios y el tiempo de trabajo por parte de los beca-trabajo que laboran en la escuela.

Como parte de las entrevistas y con la ayuda del trabajo [22] realizado por una coordinadora se fueron determinando los posibles indicadores que serían de interés y que deberían de ser tomados en cuenta durante la elaboración del sistema de apoyo a la toma de decisiones. Como característica implícita de la mayoría de los sistemas se desea que el mismo

sea fácil de utilizar para los usuarios de forma tal que les permita de manera sencilla obtener la información que desean en el menor tiempo posible. Otra de las características que debe tener el sistema es su posibilidad de ser enlazado con los sistemas transaccionales, es decir, poder tener un enlace en el menú de los sistemas transaccionales (Intranet) que les permita a los usuarios autorizados visualizar el sistema de reportes (DSS). Para cumplir con estos requerimientos se procede a la búsqueda de herramientas (en su mayoría open source) que permitan manipular la información que será mostrada por el sistema de apoyo a la toma de decisiones. Los principales casos de uso que deberían ser realizados y que posiblemente debería tener la herramienta a utilizar son los siguientes:

NOMBRE DEL CASO DE USO	DESCRIPCIÓN
Iniciar Sesión	Permite al Director(a) de la escuela y a los coordinadores ingresar al sistema de reportes.
Consultar Reporte	Permite al Director(a) de la escuela y a los coordinadores ver el conjunto de reportes predefinidos en el sistema.
Exportar Reporte	Permite al Director(a) de la escuela y a los coordinadores exportar los reportes predefinidos en formato Excel y PDF.
Modificar Reporte	Permite al Director(a) de la escuela y a los coordinadores modificar los reportes predefinidos en el sistema.
Cerrar Sesión	Permite al Director(a) de la escuela y a los coordinadores desconectarse del sistema.
Consultar Informe ETT	Permite al Director(a) de la escuela y a los coordinadores consultar las fechas de las distintas ejecuciones de las cargas de ETT.
Parametrizar Reporte	Permite al Director(a) de la escuela y a los coordinadores seleccionar nuevos parámetros o medidas para los reportes predefinidos.

Tabla 4 .Lista de Casos de Uso del Sistema.

Fuente: Elaboración Propia.

IV.2.2 Análisis de Requerimientos.

Una vez determinado los requerimientos que debe tener el sistema a desarrollar se procede a la fase de análisis de dicha información para esto se desarrollaron los modelos de caso de uso y la descripción de los casos de uso, los mismo pueden ser consultados en el Apéndice G.3 y G.4 respectivamente. En la descripción de los casos de uso se puede observar

cómo será la interacción de los usuarios finales con el sistema de apoyo a la toma de decisiones.

IV.2.2.1 *Identificación de Indicadores.*

Partiendo de la información obtenida en la fase de concepto de software se planteó a las coordinaciones algunos indicadores que pudieran ser de su interés, se realizó una reunión con cada coordinación para determinar cuáles de los indicadores planteados iban a ser tomados en cuenta y que indicarán algún otro posible indicador que era de su interés. Una vez determinados los posibles indicadores se realizó una reunión con la dirección de la escuela la cuál suministro una lista de los indicadores que le interesaba fueran tomados en cuenta en el sistema de apoyo a la toma de decisiones, así mismo se le entregó la lista de indicadores planteados por cada coordinación para finalmente determinar la lista definitiva de indicadores que deben ser utilizados en el sistema.

La Tabla 5 contiene los indicadores definitivos que serán utilizados en el sistema de apoyo a la toma de decisiones a implementar:

CLASIFICACIÓN	INDICADOR	DESCRIPCIÓN
FINANCIEROS	Porcentaje de Ítems Consumidos.	Contiene el porcentaje de los materiales de laboratorio que han sido consumidos (gastados) basándose en la cantidad inicial de materiales que había.
	Porcentaje de Pérdida de Material.	Contiene el porcentaje de los materiales de laboratorio que se pierden por periodo académico basándose en la cantidad inicial de materiales que había.
	Porcentaje de Cumplimiento de Horas.	Contiene el porcentaje de Cumplimiento de Horas por parte de los beca-trabajo de la escuela. Basándose en la cantidad de horas por la que fue contratado y la cantidad de horas que cumplió.
	Porcentaje de Tiempo por Actividad.	Contiene el porcentaje de dedicación de tiempo a cada actividad en la escuela por parte de los beca-trabajo, esto respecto a la planificación realizada inicialmente por el director de escuela.
MATERIAS ELECTIVAS	Cantidad de materias Electivas Nuevas.	Contiene el total y el porcentaje de materias electivas nuevas que se abren en un periodo

		académico. Se diferencia entre electivas internas (ofertadas por la escuela de ingeniería informática) y externas (ofertadas por otras escuelas o facultades).
	Cantidad de Materias Electivas Repetidas.	Contiene el total y el porcentaje de materias electivas que se repiten en un período académico (es decir, materias electivas que ya hayan sido dictadas anteriormente). Se diferencia entre electivas internas (ofertadas por la escuela de ingeniería informática) y externas (ofertadas por otras escuelas o facultades).
	Variación en la Oferta de Electivas.	Contiene la variación (tanto en cantidad como en porcentaje) en la oferta de las materias electivas entre dos periodos académicos.
TRABAJO ESPECIAL DE GRADO (TEG)	Porcentaje de TEG con Menciones.	Contiene el porcentaje de los TEG que obtienen alguna mención (honorífica y/o publicación) en un período académico.
	Nota promedio de los TEG.	Contiene la nota promedio de todos los TEG defendidas en un período académico.
	Porcentaje de Estudiantes sin entregar propuesta.	Contiene el porcentaje de estudiantes que tiene los requisitos necesarios para presentar una propuesta de trabajo especial de grado pero que no la han entregado. El requisito para poder entregar una propuesta de TEG es haber aprobado Seminario de TEG.
	Porcentaje de estudiantes con pérdida de tema de TEG.	Contiene el porcentaje de estudiantes que abandona un trabajo especial de grado, es decir, que pasan los 2 años desde la aprobación y no entregan el trabajo.
	Porcentaje de estudiantes con TEG Reprobados.	Contiene el porcentaje de estudiantes con trabajo de grado reprobado (nota menor a 10 puntos) en un período académico.
	Porcentaje de estudiantes con TEG Aprobados.	Contiene el porcentaje de estudiantes con trabajo de grado aprobado (nota mayor o igual a 10 puntos) en un período académico.
	Porcentaje de estudiantes con TEG Cancelados.	Contiene el porcentaje de estudiantes con trabajo de grado que son cancelados por los estudiantes, la empresa, o la escuela.
	Cantidad de propuestas entregadas por estudiante (antes de tener un tema válido).	Contiene la cantidad de propuestas que entrega un estudiante antes que tenga un tema válido, es decir, que sea aprobada la propuesta de TEG.
	Porcentaje de Propuestas Aprobadas.	Contiene el porcentaje de propuestas de TEG que son aprobadas en un periodo académico.
	Porcentaje de Propuestas Rechazadas.	Contiene el porcentaje de propuestas de TEG que son rechazadas en un período académico.
	Porcentaje de Propuestas Aceptadas (con seminario de TEG).	Contiene el porcentaje de propuestas de TEG que son aprobadas y provienen de estudiantes que han aprobado el Seminario de TEG.
PREPARADURÍAS	Número de veces dictada una preparaduría/asistencia de cátedra (por preparador,	Contiene la cantidad de veces que un preparador o asistente dicto una materia. También se puede consultar la cantidad de

	asistente, materia).	preparadores que tiene una materia.
	Evaluación de desempeño (preparadores, asistentes).	Contiene la evaluación obtenida por un preparador o asistente en un período académico según el criterio de los estudiantes.
	Porcentaje de Asistencias.	Contiene el porcentaje de Asistencias de los preparadores en un período académico.
PASANTÍAS	Porcentaje de Estudiantes sin pasantías.	Contiene el porcentaje de estudiantes que posee los requisitos para realizar pasantías pero que no tienen una propuesta aprobada o no ingresaron propuesta. El requisito para poder realizar pasantías es haber aprobado Desarrollo del Software.
	Porcentaje de Propuestas rechazadas.	Contiene el porcentaje de propuestas de pasantías rechazadas por el consejo de escuela.
	Porcentaje de Propuestas aprobadas.	Contiene el porcentaje de propuestas de pasantías aprobadas por el consejo de escuela.
	Empresas con más pasantes.	Contiene las empresas que poseen más pasantes en un período académico.
	Empresas empleadoras que repiten.	Contiene las empresas que repiten la contratación de pasantes en distintos períodos académicos.
	Nota promedio de las Pasantías.	Contiene la nota promedio de las pasantías en un período académico.
	Porcentaje de Pasantías Aprobadas	Contiene la cantidad y porcentaje de estudiantes que aprueban las pasantías.
	Porcentaje de Pasantías Reprobadas	Contiene la cantidad y porcentaje de estudiantes que reprueban las pasantías.
SERVICIO COMUNITARIO	Porcentaje de estudiantes con seminario de Servicio Comunitario Aprobado.	Contiene el porcentaje de estudiantes que tiene el seminario de servicio comunitario aprobado.
	Porcentaje de Estudiantes con Servicio Comunitario culminado.	Contiene el porcentaje de estudiantes que ya culminaron su servicio comunitario.
	Porcentaje de Estudiantes con Servicio Comunitario iniciado pero no culminado.	Contiene el porcentaje de estudiantes que iniciaron su servicio comunitario pero que no lo han concluido.
	Porcentaje de Propuestas Aprobadas.	Contiene el porcentaje de propuestas de servicio comunitario aprobadas.
	Porcentaje de Propuestas Rechazadas.	Contiene el porcentaje de propuestas de servicio comunitario rechazadas.
	Cantidad de estudiantes que pueden realizar el seminario de Servicio Comunitario pero no lo han realizado.	Contiene la cantidad de estudiantes que tienen los requisitos para realizar el seminario de servicio comunitario (estudiantes con más de 100 UC) pero que no lo han realizado.

	Porcentaje de Estudiantes con seminario aprobado pero no han iniciado el servicio comunitario.	Contiene el porcentaje de estudiantes que tienen el seminario de servicio comunitario aprobado pero que no han iniciado el servicio comunitario.
	Porcentaje de Cobertura.	Porcentaje de asistencia al seminario de servicio comunitario en base a la cantidad de estudiantes convocados. *Estudiantes Convocados:* estudiantes que tienen más de 100 UC y que pueden realizar el seminario.
PROFESORES	Porcentaje de profesores ordinarios.	Contiene el porcentaje de profesores que son ordinarios en al menos una materia en la escuela para un período académico determinado.
	Porcentaje de profesores por Escalafón.	Contiene el porcentaje de profesores por escalafón en la escuela para un período académico determinado.
	Porcentaje de Profesores según Nivel de Educación.	Contiene el porcentaje de profesores según el nivel de estudios que tengan para un período académico determinado.
	Numero de Tutorias (TEG, Pasantías).	Número de trabajos en los que el profesor es tutor, tanto de pasantías como de TEG. Este indicador se realiza por profesor y un promedio de todas las tutorías.
ESTUDIANTES	Porcentaje de Crecimiento de Estudiantes.	Contiene el porcentaje de incremento de la población estudiantil entre dos períodos académicos.
	Tiempo promedio para graduarse de Ingeniero en Informática.	Contiene el tiempo promedio (cantidad de años) que tarda un estudiante para graduarse de Ingeniero en Informática.
	Porcentaje de Estudiantes que culminan la carrera.	Contiene el porcentaje de estudiantes que culminan la carrera. Basándose en la cantidad de estudiantes que ingresan propuestas aprobadas de trabajo especial de grado y la cantidad de los estudiantes que aprueban trabajos especiales de grado.
EGRESADOS	Porcentaje de Egresados que son profesores en su escuela.	Contiene el porcentaje de egresados que ha sido profesores en la escuela para un determinado período académico.

Tabla 5. Descripción de Indicadores.

Fuente: Elaboración Propia.

Adicionalmente, se diseñaron los documentos: el documento *Fórmulas de los Indicadores* permite establecer cuál es la fórmula aplicada a los datos para poder obtener el indicador deseado, el mismo puede ser consultado en el Apéndice K y el documento *Procesos asociados a los indicadores* permite establecer en que parte del proceso de toma de decisiones de cada

coordinación influencia cada uno de estos indicadores, dicho documento puede ser consultado en el Apéndice L.

IV.2.2.2 *Evaluación y selección de la herramienta de desarrollo.*

En todo desarrollo de sistema es necesario realizar un estudio de las herramientas a utilizar, esto permite garantizar que las herramientas seleccionadas cumplen con las necesidades mínimas para la implementación. Para este trabajo se analizaron un conjunto de herramientas, tanto de base de datos como software para el diseño de los reportes.

La primera evaluación que se hizo fue para determinar el manejador de base de datos que se utilizará en la implantación del Datamart, el manejador seleccionado determinará el posible conjunto de herramientas que puedan ser utilizadas para el desarrollo del sistema de apoyo a la toma de decisiones, por eso el análisis fue realizado recolectando la mayor cantidad de información que nos permitiera determinar en base a las características del sistema a desarrollar cuál manejador ofrecerá mejor rendimiento.

Como parte de la evaluación del manejador de base de datos se crearon dos tablas comparativas, la primera de ella con las características de hardware de los manejadores de base de datos y la segunda tabla un estudio de factibilidad de cada herramienta basado en ciertos criterios.

Comparación de requerimientos hardware.

Manejador de base de datos	Espacio en disco	Memoria RAM	Procesador
PostgreSQL 8.3	39 MB	8 MB	Pentium III o posteriores.
SQL Server 2008	360 MB (se recomienda 1GB)	512 MB (se recomienda 2 GB)	1 GHz (se recomienda 200 GHz)
Oracle Database 11g	4,92 GB	1 GB	550 MHz

My SQL 5.1.30	Mínimo 200 MB	100 MB	Cualquier procesador actual.

Tabla 6. Comparación de requerimientos de los manejadores.

Fuente: Elaboración propia.

Como se puede observar, en la Tabla 6 podemos observar que los manejadores de bases de datos que representan una buena opción para la implementación del datamart son PostgreSQL y MySQL, ambos necesitan muy poca cantidad de espacio de disco duro, la cantidad de memoria RAM es sumamente baja, y la aceptación de cualquier procesador, a partir de esta tabla se procede a la elaboración de otra en la cual podremos ver otros criterios que nos permitan tomar una decisión que realmente cubra las necesidades de nuestro datamart.

Comparación de características de los manejadores de base de datos.

Manejador de base de datos	Costo de Licencia	Seguridad	Costo de Soporte	Disponibilidad del Soporte	Acceso en la Universidad	Consumo de recursos	Instalación/ Proceso de mantenimiento	Total
PostgreSQL 8.3	1	3	1	3	5	3	3	19
SQL Server 2008	3	4	4	5	1	4	4	25
Oracle Database 11g	5	5	5	5	3	5	5	33
My SQL 5.1.30	1	3	3	3	5	1	1	17

Tabla 7. Comparación de características de los manejadores.

Fuente: Elaboración propia.

Leyenda (Escala del 1 al 5):

1. Muy bajo. 2. Bajo 3. Regular 4. Alto 5. Muy Alto

Tal como se muestra en la Tabla 7, podemos observar que el manejador que satisface las necesidades mínimas para el Datamart es MySQL, por su bajo costo en licencia (es Open Source), ofrece un buen nivel de seguridad a nivel de base de datos, el costo del soporte es un costo moderado, sin embargo se cuenta con suficiente información en la página

web del proveedor de MySql que ayuda a los desarrolladores a una óptima utilización del manejador, la disponibilidad del soporte es moderada pero como se mencionó anteriormente se cuenta con suficiente material en la página del proveedor, el acceso en la institución en la cual se va a implantar el datamart es totalmente factible ya que el mismo manejador es utilizado en los sistemas transaccionales que utiliza la institución lo cual facilitaría una integración y manejo por parte de los desarrolladores que allí laboran. El consumo de recursos es sumamente bajo (dependiendo del tamaño de la base de datos), su proceso de instalación es relativamente sencillo y en la página web del proveedor se ofrecen documentos de los pasos a seguir para la instalación en los distintos sistemas operativos.

Según la totalización de los valores asignados en las distintas características de los manejadores, se debe considerar como recomendable el que posea menor cifra, en este caso como podemos observar el manejador que cumple dicha condición es MySql, por lo tanto se recomienda utilizar dicho manejador.

Por todas las razones antes mencionadas se ha seleccionado MySQL como el manejador de base de datos que se utilizará en la implementación del datamart. Para más detalle de las características de los manejadores de base de datos consultar el Anexo A.

El manejador seleccionado permite establecer el conjunto de herramientas que permitirán conformar el sistema de apoyo a la toma de decisiones. Según la investigación realizada no existen muchas herramientas Open Source para el desarrollo de un datamart en MySql, por lo tanto gracias a la investigación se determinó lo siguiente:

- ✓ Como MySql ofrece la posibilidad de generar los procesos de Extracción, Transformación y Transporte de los datos se decidió la implementación de los procesos a través del mismo manejador, esto por medio de la creación de procedimientos y funciones que permitan el desarrollo de estos procesos que son cruciales en todo almacén de datos.

- ✓ Las posibles herramientas para la visualización de los reportes fueron: Mondrian (Pentaho) y OpenI, ambas herramientas son web. Uno de los principales requerimientos no funcionales y que son de importancia para la mayoría de los usuarios es una interfaz amigable, Mondrian por su parte no ofrece una interfaz amigable sino más bien orientada a la visualización de información por parte de desarrolladores, mientras que OpenI está orientado más a una interfaz amigable, fácil de usar, configurable lo cual será de mucho más agrado para los usuarios finales, por lo tanto como herramienta para la visualización de los reportes fue seleccionada OpenI.

- ✓ Como herramienta para el diseño y creación de los Cubos OLAP se seleccionó Schema Workbench (Pentaho), esta herramienta posee una interfaz amigable, es muy fácil de usar y existe bastante documentación con la cual se pueden orientar los desarrolladores al momento de crear los Cubos OLAP.

El conjunto de herramientas evaluadas y seleccionadas fueron basadas en el manejador de base de datos seleccionado (MySql), las herramientas seleccionadas ofrecen una buena integración entre ellas y además garantizan el cumplimiento de los requerimientos (tanto funcionales como no funcionales) que fueron determinados en la etapa de concepto de

software. Para más detalle de las herramientas evaluadas consultar el Anexo C.

IV.2.3 Diseño Global.

En esta etapa se detallan los elementos que formarán parte del sistema de apoyo a la toma de decisiones y que en su mayoría son un estándar al momento del desarrollo de este tipo de sistemas:

- *Fuentes de datos para el Datamart.*

Las fuentes de datos son aquellos archivos, bases de datos, etc. que proporcionarán la información que será previamente procesada y validada para posteriormente ser almacenada en las estructuras del Datamart. Para esta implantación utilizaremos archivos de Excel que son proporcionados por distintas fuentes: Coordinaciones y reportes de sistemas (Sistemas de la Escuela de Ingeniería Informática y Sistema de la Universidad Católica Andrés Bello).

- *Procesos de Extracción, Transformación y Transporte (ETT).*

Los procesos de extracción, transformación y transporte fueron implementados como procedimientos y funciones almacenadas en el manejador de base de datos, estos procesos realizarán las validaciones correspondientes para garantizar que la información que sea transportada al Datamart cumpla con las condiciones necesarias que permitan obtener información confiable para la toma de decisiones.

- *El Datamart (almacén de datos).*

Esta estructura de datos permitirá el almacenamiento de la información previamente validada proveniente de las distintas fuentes de datos y la elaboración de estructuras multidimensionales para la manipulación de la información. Esta estructura de datos será la fuente de

datos de los reportes que serán visualizados en el sistema de apoyo a la toma de decisiones, la estructura fue implementada en el manejador de base de datos MySql.

- *Aplicación de Reportes.*

El desarrollo de esta aplicación incluirá el conjunto de reportes que se obtienen a partir de los Cubos OLAP y que serán mostrados al usuario final para ayudar al momento de que los mismos deban tomar alguna decisión.

- *Arquitectura del Sistema.*

Figura 6. Arquitectura del Sistema de Apoyo a la toma de decisiones.
Fuente: Elaboración Propia.

Como se puede observar en la Figura 6 la arquitectura del sistema está conformada por las fuentes de datos (los cuales son archivos básicamente de Excel que contienen los datos necesarios para la obtención de los indicadores establecidos en la etapa de análisis), estos archivos son la entrada para los procesos de extracción transformación y transporte (los cuales son los encargados de obtener los datos, darle el formato necesario, hacer las validaciones correspondientes y finalmente transportar los datos hasta la estructura del datamart), como función final de los procesos ETT tenemos los datos que serán insertados en el almacén de datos los cuales posteriormente a través de estructuras multidimensionales y a través de un sistema de reportes mostrará la información relevante que necesitan los usuarios finales al momento de tomar las decisiones pertinentes.

IV.2.4 Diseño e Implementación del Almacén de Datos.

Esta etapa tal como lo muestra la metodología a utilizar se compone de varias actividades lo que permitirá realizar determinada cantidad de iteraciones hasta lograr tener el almacén de datos implementado. Cada una de estas actividades permite ir perfeccionando y completando el almacén de datos, es iterativa ya que a medida que se vaya implementando se debe ir corrigiendo aquellos errores detectados en iteraciones anteriores y además es incremental ya que permite ir agregando funcionalidades en cada iteración hasta lograr tener implementado el almacén.

El diseño debe contemplar todas aquellas áreas de interés que fueron tratadas en las entrevistas y que realmente serán el objetivo principal de la construcción del almacén. Las iteraciones realizadas fueron básicamente a nivel de coordinaciones, por lo que se necesitaron un total de seis iteraciones para completar la implementación.

En la primera iteración se diseñaron e implementaron los procesos ETT para los indicadores relacionados con la dirección de la escuela, en esta iteración se obtendrían los indicadores relacionados con los profesores, la población estudiantil, los beca-trabajo y los laboratorios de la escuela. Dentro de esta iteración se diseño un proceso ETT para guardar la información de los procesos de carga ejecutados. Estos indicadores se encuentran almacenados en una tabla de hechos que servirá de fuente para los reportes relacionados al área.

En la segunda iteración se diseñaron e implementaron los procesos ETT del área de Preparadurías, estos procesos están relacionados con los indicadores de rendimiento, asistencia y materias dictadas por los estudiantes que cumplen el rol de Preparadores. Estos indicadores forman

parte de un conjunto de indicadores de una tabla de hechos de la cual se obtendrán reportes para necesarios para esta coordinación.

En la tercera iteración se diseñaron e implementaron los procesos ETT para el área de Pasantías, estos procesos están relacionados con los indicadores de rendimiento de estudiantes en las pasantías, de las empresas que ofrecen pasantías y de los profesores que son tutores académicos en las pasantías. Estos indicadores forman parte de un conjunto de indicadores de una tabla de hechos de la cual se obtendrán reportes para necesarios para esta coordinación.

En la cuarta iteración se diseñaron e implementaron los procesos ETT para el área de Electivas, estos procesos están relacionados con los indicadores de oferta de nuevas materias electivas, electivas abiertas por período (tanto las electivas ofertadas por la escuela de ingeniería informática como las ofertadas por otras escuelas y/o facultades). Estos indicadores forman parte de un conjunto de indicadores de una tabla de hechos de la cual se obtendrán reportes para necesarios para esta coordinación.

En la quinta iteración se diseñaron e implementaron los procesos ETT para el área de Servicio Comunitario, estos procesos están relacionados con los indicadores de rendimiento de los estudiantes durante su realización del seminario de servicio comunitario y las pasantías de servicio comunitario. Estos indicadores forman parte de un conjunto de indicadores de una tabla de hechos de la cual se obtendrán reportes para necesarios para esta coordinación.

En la sexta iteración se diseñaron e implementaron los procesos ETT del área de Trabajo Especial de Grado (TEG), estos procesos están relacionados con los indicadores de rendimiento de los estudiantes durante la

realización de su TEG, con las propuestas entregadas a consejo de escuela por parte de los estudiantes, con la nota promedio de los TEG, y con aquellos trabajos que son abandonados ó cancelados por distintas circunstancias. Estos indicadores forman parte de un conjunto de indicadores de una tabla de hechos de la cual se obtendrán reportes para necesarios para esta coordinación.

IV.2.4.1 *Análisis del origen y destino de datos.*

La principal fuente de datos del almacén son los archivos de Excel proporcionados por cada coordinación así como también los archivos generados del Sistema de Reportes Banner de la Universidad, las otras fuentes son reportes de los sistemas transaccionales. El tener múltiples fuentes puede llegar a representar un problema de consistencia de los datos, es por ellos que uno de los productos de este trabajo de grado fue el diseño de una estructura que permita servir de fuente de datos al almacén, en un futuro próximo cuando la estructura de datos sea implementada y los sistemas cumplan con los procesos de cada coordinación se podrán obtener los archivos de dichos sistemas transaccionales para tener mayor consistencia de los datos y poseer solo una fuente de datos.

El destino de los datos estará basado en la arquitectura ROLAP, en la cual la base de datos relacional maneja los requerimientos de almacenamiento de datos, y el motor ROLAP proporciona la funcionalidad analítica (Para más información consultar el Anexo B.1.2).

IV.2.4.2 *Diseño del Almacén de Datos.*

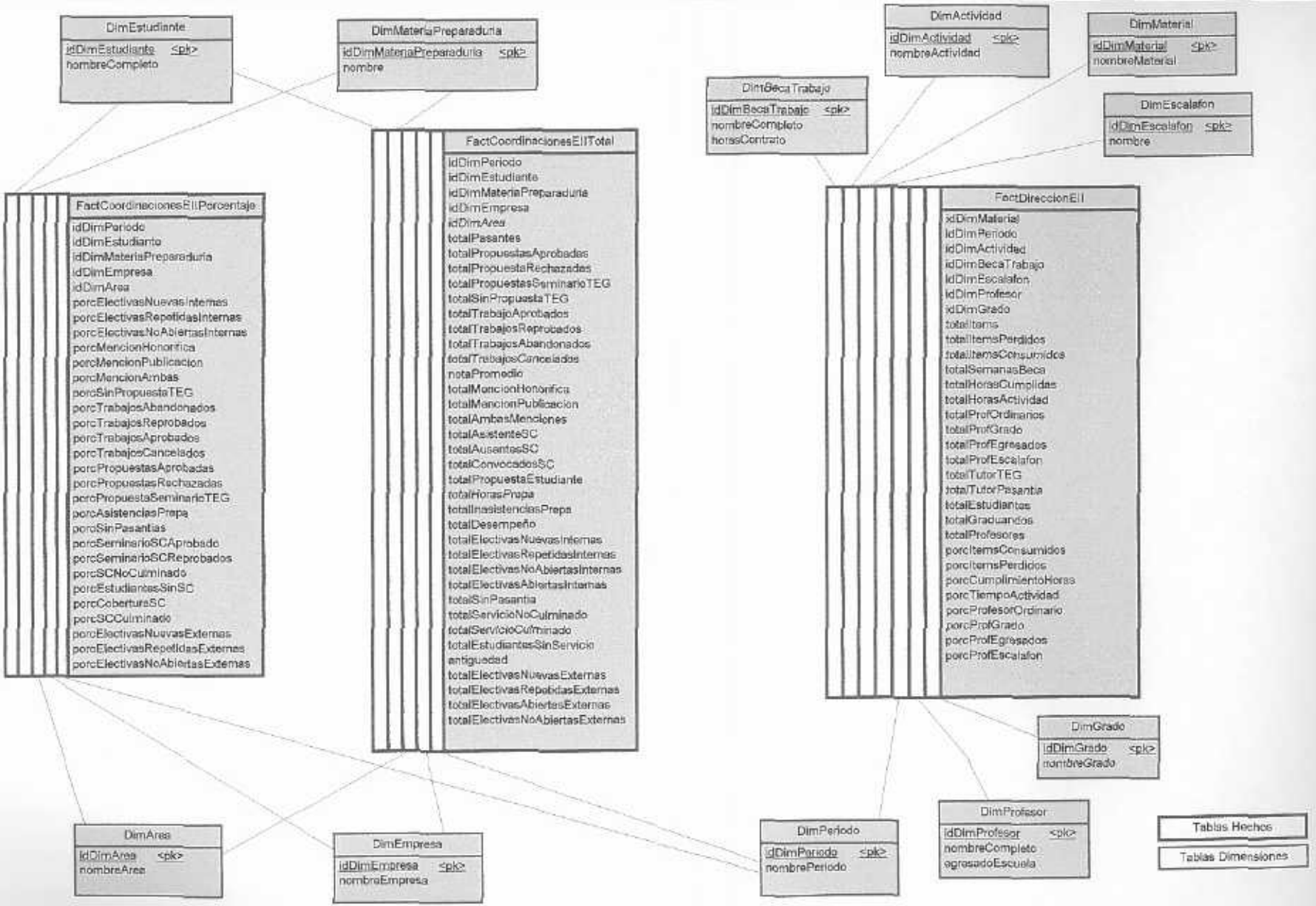
En esta etapa se diseñó la estructura de datos para el almacén de datos, la estructura está compuesta por tres tablas de hechos, once tablas de dimensiones, una tabla de totalización y tres tablas de metadata. Se definieron los atributos y los tipos de datos de los mismos. Después de haber

definido la estructura de datos se estudian los posibles esquemas que pueden ser utilizados en el diseño del almacén de datos (Para información de los esquemas ver el Anexo B.2), se logró determinar que el esquema de constelación es el que mejor se adapta a las características del almacén de datos, este esquema está compuesto por una serie de esquemas en estrella, en donde las tablas se ubican en el centro del modelo y están relacionadas con sus respectivas tablas de dimensiones.

Se realizó una documentación de las tablas que conforman el almacén de datos (Ver Apéndice A).

Se acordó con la dirección de la escuela que el almacén de datos guardará información de los últimos 5 años y la información que esté fuera de estos 5 años, será eliminada al momento de ejecutar los procesos ETT. El tiempo de permanencia de los datos (los 5 años) fue decidido basándose en la cantidad de información que se almacenará (la cual no representa un espacio de almacenamiento amplio sino que al contrario el espacio de almacenamiento es bajo debido a los pocos datos y a que las tablas de hechos almacenarán información sumarizada) y debido a que la necesidad de información para la toma de decisiones no requiere de un alto nivel histórico (gran cantidad de años).

El diseño del almacén de datos es mostrado en la Figura 7 :


DimEstudiante
idDimEstudiante <pk>
nombreCompleto
DimMateriaPreparaduria
idDimMateriaPreparaduria <pk>
nombre
DimActividad
idDimActividad <pk>
nombreActividad
DimMaterial
idDimMaterial <pk>
nombreMaterial
DimBecaTrabajo
idDimBecaTrabajo <pk>
nombreCompleto
horasContrato
DimEscalafon
idDimEscalafon <pk>
nombre
FactCoordinacionesEIIPorcentaje
idDimPeriodo
idDimEstudiante
idDimMateriaPreparaduria
idDimEmpresa
idDimArea
porcElectivasNuevasInternas
porcElectivasRepetidasInternas
porcElectivasNoAbiertasInternas
porcMencionHonorifica
porcMencionPublicacion
porcMencionAmbas
porcSinPropuestaTEG
porcTrabajosAbandonados
porcTrabajosReprobados
porcTrabajosAprobados
porcTrabajosCancelados
porcPropuestasAprobadas
porcPropuestasRechazadas
porcPropuestaSeminarioTEG
porcAsistenciasPrepa
porcSinPasantias
porcSeminarioSCAprobado
porcSeminarioSCReprobados
porcSCNoCulminado
porcEstudiantesSinSC
porcCoberturaSC
porcSCCulminado
porcElectivasNuevasExternas
porcElectivasRepetidasExternas
porcElectivasNoAbiertasExternas
FactCoordinacionesEIITotal
idDimPeriodo
idDimEstudiante
idDimMateriaPreparaduria
idDimEmpresa
idDimArea
totalPasantes
totalPropuestasAprobadas
totalPropuestaRechazadas
totalPropuestasSeminarioTEG
totalSinPropuestaTEG
totalTrabajoAprobados
totalTrabajosReprobados
totalTrabajosAbandonados
totalTrabajosCancelados
notaPromedio
totalMencionHonorifica
totalMencionPublicacion
totalAmbasMenciones
totalAsistenteSC
totalAusentesSC
totalConvocadosSC
totalPropuestaEstudiante
totalHorasPrepa
totalInasistenciasPrepa
totalDesempeño
totalElectivasNuevasInternas
totalElectivasRepetidasInternas
totalElectivasNoAbiertasInternas
totalElectivasAbiertasInternas
totalSinPasantia
totalServicioNoCulminado
totalServicioCulminado
totalEstudiantesSinServicio
antiguedad
totalElectivasNuevasExternas
totalElectivasRepetidasExternas
totalElectivasAbiertasExternas
totalElectivasNoAbiertasExternas
FactDireccionEII
idDimMaterial
idDimPeriodo
idDimActividad
idDimBecaTrabajo
idDimEscalafon
idDimProfesor
idDimGrado
totalItems
totalItemsPerdidos
totalItemsConsumidos
totalSemanasBeca
totalHorasCumplidas
totalHorasActividad
totalProfOrdinarios
totalProfGrado
totalProfEgresados
totalProfEscalafon
totalTutorTEG
totalTutorPasantia
totalEstudiantes
totalGraduandos
totalProfesores
porcItemsConsumidos
porcItemsPerdidos
porcCumplimientoHoras
porcTiempoActividad
porcProfesorOrdinario
porcProfGrado
porcProfEgresados
porcProfEscalafon
DimGrado
idDimGrado <pk>
nombreGrado
DimArea
idDimArea <pk>
nombreArea
DimEmpresa
idDimEmpresa <pk>
nombreEmpresa
DimPeriodo
idDimPeriodo <pk>
nombrePeriodo
DimProfesor
idDimProfesor <pk>
nombreCompleto
egresadoEscuela
Tablas Hechos
Tablas Dimensiones

IV.2.4.3 *Implementación del Almacén de Datos.*

Partiendo del diseño del almacén de datos se procedió a la implementación del mismo, como el manejador de bases de datos seleccionado fue MySql entonces las estructuras de datos fueron creadas en dicho manejador. Se crearon todas las tablas de hechos, de dimensiones, de metadata y totalización necesarias junto con el conjunto de atributos para el almacenamiento de los datos. Solo se crearon las estructuras sin realizar ningún tipo de carga de datos en dichas estructuras.

El detalle de las tablas de hechos, dimensiones, metadata y totalización junto con sus respectivos atributos puede ser consultado en el Apéndice H.

IV.2.4.4 *Definición de los procesos ETT.*

En esta etapa se realizó el diseño de los procesos de extracción, transformación y transporte basándonos en las fuentes de datos y en la estructura del almacén de datos. Este diseño permitirá realizar la carga de los datos en el almacén de datos.

Para el diseño de los procesos se tomó en cuenta el tipo de fuente de datos, en este caso archivos de Excel, los cuales debían ser cargados por el manejador de base de datos para su posterior procesamiento. Para realizar esta carga de los archivos se utilizaron tablas temporales que posee el manejador MySql las cuales nos permiten crear estructuras de datos temporales en las cuales poder almacenar información, esto nos facilita que durante la ejecución del proceso se pueda cargar los datos en dichas tablas hacer las validaciones pertinentes, luego insertar dicha información en el almacén de datos y finalmente borrar dichas tablas temporales para que no ocupen espacio de almacenamiento dentro del manejador de base de datos.

Dentro de los procesos se realizaron las validaciones correspondientes que garantizan que la información que será insertada en el almacén de datos será nueva y no generará datos duplicados que puedan dañar la calidad de la información que se mostrará en los reportes.

Se diseñó un proceso que permite almacenar la información de las cargas de datos que han sido ejecutadas, dicha tabla contiene la fecha y hora en que fue ejecutada la carga y el período académico que fue agregado al almacén de datos.

Los datos que fueron extraídos, validados y totalizados poseen un nivel de granularidad bajo, es decir, no se cuenta con mucho detalle de información, sólo en aquellos casos en las cuales es necesario realizar algún nivel de detalle. En su mayoría el detalle llega sólo a nivel de períodos académicos.

Así mismo se necesita contar con un proceso o acción que permita la ejecución de dichos procesos cada cierto tiempo para poder incorporar nueva información dentro del almacén de datos, se programó un proceso de carga semestralmente y en una hora en la cual no interfiere con otros procesos que puedan estarse llevando a cabo en el equipo.

IV.2.4.5 *Implementación de los procesos ETT.*

Ya definidos y diseñados los procesos ETT se procede a la implementación de los mismos, estos procesos estarán basados en funciones y procedimientos que serán creados sobre el manejador de base de datos MySql. Se implementaron tantos procedimientos como fueron necesarios para poder realizar la transformación, limpieza y validación de los datos. Cada uno de estos procedimientos es el encargado de realizar la carga de datos en una dimensión o parte de una tabla de hechos, es decir,

podemos realizar la carga correspondiente a determinadas dimensiones y/o tablas de hechos basándose en la información contenida en los archivos de Excel.

Por ejemplo, se tiene la carga de datos de los indicadores de los profesores, estos indicadores forman parte del conjunto de indicadores que serán utilizados por la dirección de escuela. En este caso se realiza la carga de los datos en tablas temporales, luego se realizan un conjunto de transformaciones, validaciones, totalizaciones y cálculo de porcentajes que finalmente serán transportados hasta la tabla de hechos correspondiente. Por cada tabla de hecho existe un conjunto de procedimientos que permiten realizar la carga completa de todos los atributos de la misma. En la Figura 8 se muestra como es la secuencia de pasos que deben realizarse para poder realizar la carga de los datos relacionados a los profesores de la escuela:

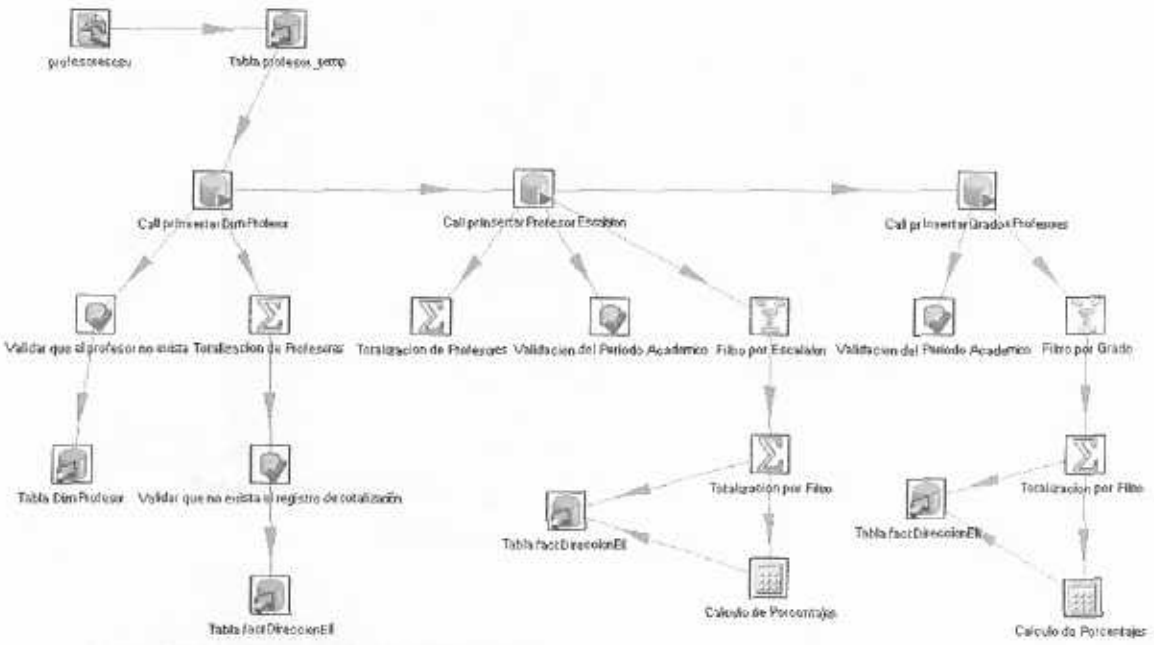


Figura 8. Diagrama ETT para los indicadores relacionados a los profesores.
Fuente: Elaboración Propia.

Los diagramas completos de todos los procedimientos implementados pueden ser consultados en el Apéndice I.

IV.2.4.6 *Elaboración de las pruebas ETT y carga del almacén de datos.*

Una vez implementados todos los procesos ETT es necesario verificar que su funcionamiento sea el correcto, motivo por el cual se realizaron unas pruebas en las cargas de los datos para verificar la consistencia de los mismos. Las pruebas consistieron básicamente en verificar que los datos, las totalizaciones y los porcentajes insertados en el almacén sean los que realmente deberían estar. En la Tabla 8 se muestra un ejemplo de los resultados obtenidos de la prueba al ejecutar el procedimiento de inserción de datos de los distintos períodos académicos:

Período para realizar la carga	Estado	Tiempo de Ejecución
200422	Completado.	0:12.530 Seg.
200423	Completado.	0:12.282 Seg.
200521	Completado.	0:15.173 Seg.
200522	Completado.	0:19.750 Seg.
200523	Completado.	0:20.187 Seg.
200621	Completado.	0:20.250 Seg.
200622	Completado.	0:24.156 Seg.
200623	Completado.	0:24.437 Seg.
200721	Completado.	0:28.953 Seg.
200722	Completado.	0:33.320 Seg.
200723	Completado.	0:33.672 Seg.
200821	Completado.	0:39.765 Seg.
200822	Completado.	0:46.281 Seg.
200823	Completado.	0:46.816 Seg.
200921	Completado.	0:54.640 Seg.

Tabla 8. Resultado de la ejecución de carga de los períodos.

Fuente: Elaboración Propia.

Una vez concluida la ejecución de los procedimientos se procedió a realizar los cálculos correspondientes para validar que los datos insertados fueran válidos. Se validaron los cálculos contra los registros insertados en el

almacén de datos y de esta forma se confirmó que los procedimientos estaban funcionando correctamente.

Luego de confirmar que todo funcionaba correctamente se crearon tres tareas en el administrador de tareas de Windows que permite la ejecución de todo el conjunto de procedimientos que son necesarios para la carga de los datos en el almacén de datos de cada período en específico, estas tareas fueron programadas para ejecutarse el segundo miércoles del mes de mayo, noviembre y diciembre. Para más detalle de estas tareas consultar el Apéndice J.

IV.2.5 Diseño de la aplicación de reportes.

Esta etapa de la metodología no se realizó debido a que uno de los objetivos de nuestro TEG era evaluar y seleccionar una herramienta de usuario final.

IV.2.6 Implementación de la aplicación de reportes

Para esta etapa no se realizó la implementación de la aplicación de reportes por lo mencionado anteriormente, sino que se realizó el diseño e implementación de los reportes en la herramienta seleccionada.

IV.2.6.1 *Diseño e Implementación de los cubos OLAP.*

Para el diseño e implementación de los cubos OLAP se utilizó la herramienta Schema Workbench de Pentaho, esta herramienta permite la creación de los cubos de forma gráfica, generando un archivo XML que sirve de fuente para el sistema de reportes. Los cubos implementados contienen una tabla de hechos y un conjunto de tablas de dimensiones que permiten obtener la información de los indicadores que forman parte del sistema de apoyo a la toma de decisiones.

Se generaron básicamente tres estructuras multidimensionales (cubos) para posteriormente ser incluidas en el sistema de reportes y poder visualizar la información tanto en forma de tablas como de gráficos. La Figura 9 muestra el diseño del cubo que contiene los indicadores que son de interés para la dirección de escuela.

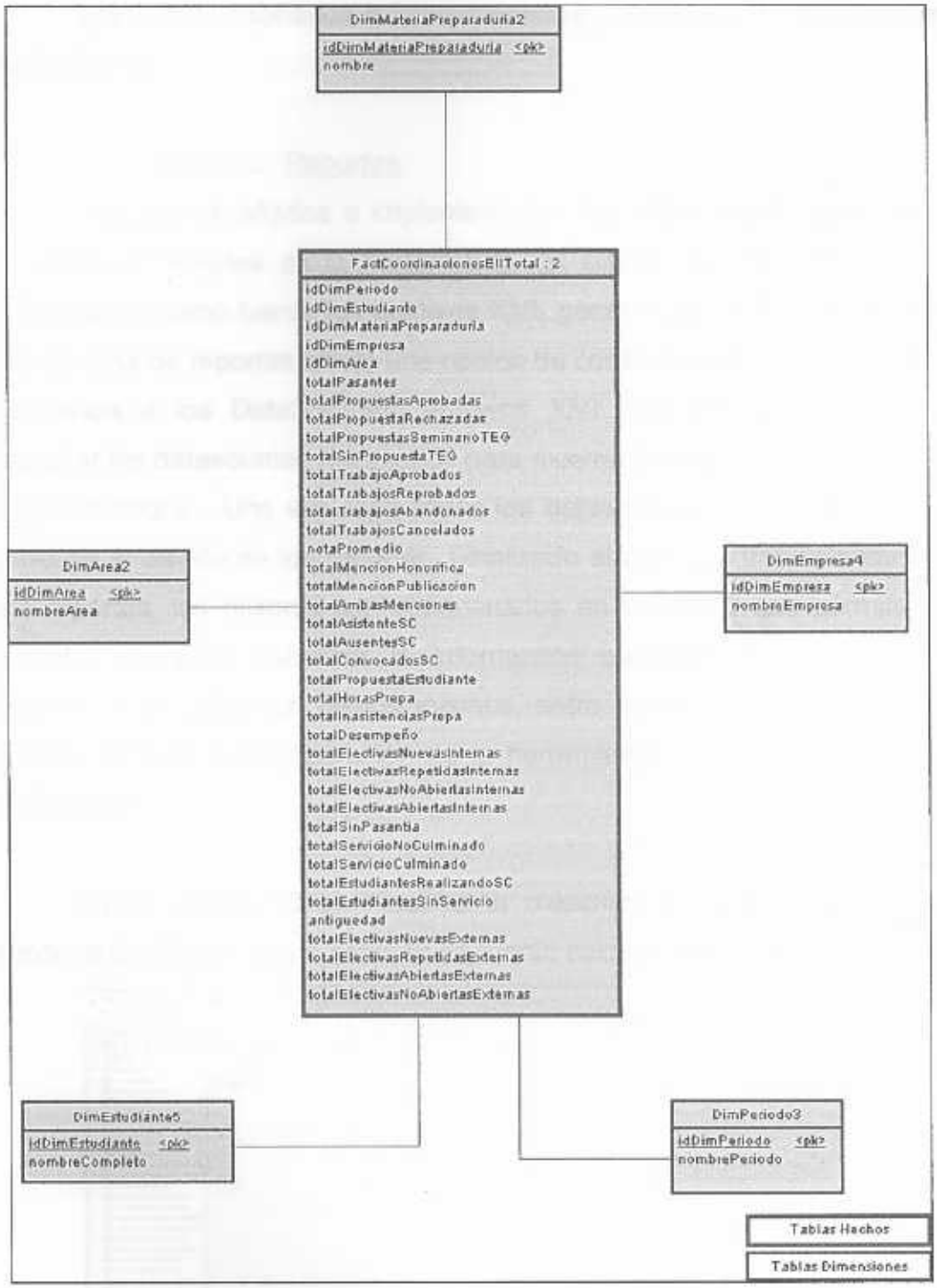


Figura 9. Diseño del Cubo para la tabla de factDireccionEII.

Fuente: Elaboración Propia.

Los cubos diseñados e implementados pueden ser consultados en el Apéndice H.2.

IV.2.6.2 *Reportes.*

Una vez diseñados e implementados los cubos OLAP se procedió a generar los reportes en la herramienta del sistema de reportes, para ello utilizaremos como fuente los archivos XML generados con los cubos. Dentro del sistema de reportes existe una opción de configuración que permite hacer referencia a los DataSources (archivos XML con los cubos), se deben registrar los datasources necesarios para mostrar la información de los cubos implementados. Una vez registrados los datasources en la herramienta se procede al diseño de los reportes. Finalizado el diseño e implementación de los reportes, los mismos son almacenados en un menú que permite a los distintos usuarios visualizar la información contenida en ellos, también permite la modificación de los mismos, entre otras opciones. Para mayor detalle de toda la configuración de la herramienta de reportes consultar el Apéndice J.

En la Figura 10 se muestra la creación del archivo XML con los reportes diseñados y que servirán de fuente para la aplicación de reportes.

Figura 10. Datasource con los reportes creados.

Fuente: Elaboración Propia.

En la Figura 11 se muestra la pantalla inicial de la aplicación de reportes donde el usuario debe iniciar sesión.

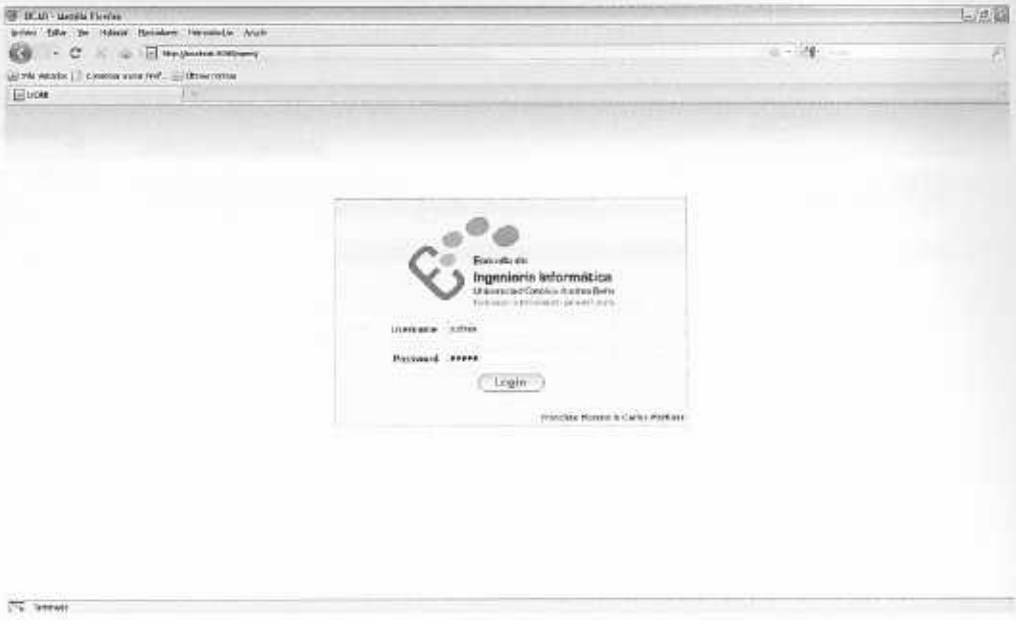


Figura 11. Pantalla de inicio de la aplicación de reportes.

Fuente: Elaboración Propia.

En la Figura 12 se muestra un ejemplo de un reporte del sistema de apoyo a la toma de decisiones.

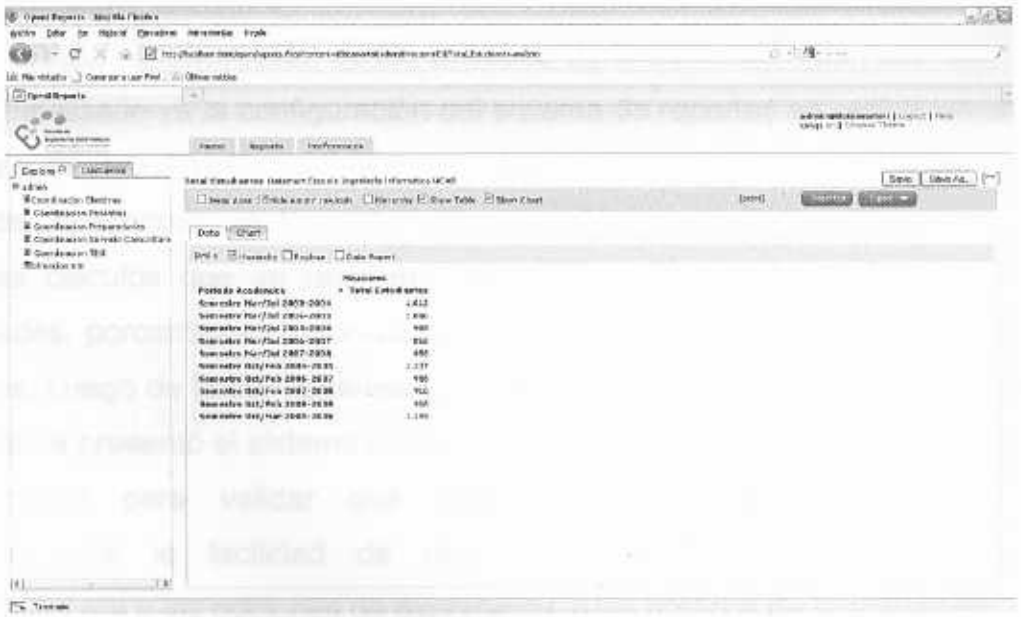


Figura 12. Reporte del Total de Estudiantes por Período Académico.

Fuente: Elaboración Propia.

En la Figura 13 se muestra el reporte gráfico del total de estudiantes por período académico.

Figura 13. Pantalla con reporte gráfico del total de estudiantes.

Fuente: Elaboración Propia.

Para mayor detalle del uso del sistema de reportes consultar el Apéndice J.

IV.2.7 Elaboración de las pruebas de la aplicación de reportes.

Finalizado ya la configuración del sistema de reportes se realizaron un conjunto de pruebas para determinar que la información visualizada era correcta, basándose en los datos almacenados en el Datamart y en los posibles cálculos que se realicen sobre dichos datos. Se verificaron las cantidades, porcentajes e información de las dimensiones mostradas en los reportes. Luego de haber confirmado que toda la información visualizada era correcta se presentó el sistema final a uno de los usuarios finales (director de la escuela) para validar que todo estuviese correcto, se evaluó principalmente la facilidad de uso y la amigabilidad del sistema, adicionalmente a las opciones de exportación y los gráficos de la aplicación.

CAPÍTULO V. RESULTADOS

Los resultados obtenidos durante la realización de este Trabajo Especial de grado son los siguientes:

1. Con la información proporcionada por los coordinadores en las distintas entrevistas se pudo definir cuáles eran los procesos que se realizaban a nivel de cada coordinación, la lista de procesos se puede consultar en el Capítulo IV.1.1.

2. Se realizaron el conjunto de diagramas de procesos y la documentación de los mismos, en dichos procesos se establece la secuencia de pasos que deben realizarse en cada coordinación, esto permite que futuros coordinadores no modifiquen los procesos según sus necesidades, los mismos deberán regirse por los procesos levantados y documentados.

3. Se logró la integración de todas las bases de datos de los sistemas transaccionales de la escuela, así como también se realizó una documentación detallada de las tablas de bases de datos y de la funcionalidad que deben cumplir los sistemas transaccionales para poder cubrir la mayoría de las etapas de cada uno de los procesos de cada coordinación.

4. Se definió una nomenclatura de base de datos lo cual permitirá tener un estándar en todas las bases de datos futuras de los sistemas que se desarrollen.

5. Se determinó que las fuentes de datos para el Datamart no podrían ser las bases de datos transaccionales ya que las mismas no poseen

actualmente la información necesaria para el sistema de apoyo a la toma de decisiones, por lo tanto, se hizo uso de archivos proporcionados por las coordinaciones y por el sistema de Reportes Banner de la Universidad.

6. Como parte de este trabajo se realizó el diseño de una integración de las aplicaciones transaccionales, al momento de ser implementado se podrá prescindir de las fuentes de datos inicialmente establecidas y se podrá obtener la información que alimentará al Datamart de estos sistemas transaccionales.

7. Gracias a la información suministrada en las entrevistas y a un trabajo [22] previamente realizado por un coordinador se determinaron los indicadores más importantes que debían ser tomados en cuenta en este Trabajo Especial de Grado.

8. Con los indicadores determinados se diseñó e implementó una estructura de datos que permitió el almacenamiento de toda la información necesaria para el sistema de apoyo a la toma de decisiones.

9. Con toda la estructura de datos y la información de los indicadores se realizó la evaluación de los distintos manejadores de bases de datos para seleccionar aquel que mejor se adaptara a las necesidades del proyecto.

10. Se realizó la implementación de los procesos de extracción, transformación y transporte (ETT) como procedimientos almacenados en el manejador de bases de datos MySql, los mismos son alimentados por los archivos fuentes en un formato .CSV (Archivo delimitado por comas).

11. Como resultado de la evaluación de las herramientas que serían utilizadas para el desarrollo del sistema de apoyo a la toma de decisiones se seleccionaron las siguientes herramientas: MySql como manejador de base de datos, OpenI como aplicación de reportes, Tomcat como servidor de aplicación J2EE para la aplicación de reportes, Schema Workbench para la creación de los cubos que sirven de fuente a la aplicación de reportes.

12. Se implementaron el conjunto de reportes en la herramienta seleccionada, esto le permite a los usuarios finales poder consultar toda la información a través de una interfaz de usuario amigable que permite el manejo de los reportes de forma dinámica.

CONCLUSIONES

Partiendo de los resultados obtenidos durante la realización de este Trabajo Especial de Grado se puede llegar a las siguientes conclusiones:

1. Es vital para la Escuela de Ingeniería Informática tener bien definidos y documentados los procesos que son llevados a cabo por las distintas unidades académicas que hacen vida en ella, esto permitirá que el proceso sea llevado de la misma forma sin importar la persona responsable que esté realizando el proceso en determinado momento.

2. Los sistemas transaccionales con los que cuenta la Escuela de Ingeniería Informática son grandes fuentes de datos por lo tanto se debe garantizar que los mismos manipulen información correcta, precisa y actualizada, ya que la misma en la mayoría de los casos sirve de fuente al momento de realizar alguna toma de decisión dentro de la institución.

3. La aplicación de estándares en la Escuela de Ingeniería Informática debe ser un requisito necesario para que la misma mejore su competitividad, eficiencia e innovación, por lo tanto se debe considerar aplicar aquellos estándares que mejor se adapten a las aplicaciones.

4. La Escuela de Ingeniería Informática debe realizar evaluaciones periódicas de las herramientas de software que salen al mercado y que posiblemente le ayudarán a automatizar la mayoría de los procesos lo cual creará mayor rapidez y precisión en el manejo de la información.

5. Los sistemas de apoyo a la toma de decisiones son requisito indispensable en el mundo globalizado en el cual nos encontramos, se debe considerar implementar dichos sistemas al momento de formar una empresa, esto traerá muchos beneficios tanto a nivel estratégico como táctico y operativo. Debido a la importancia que tienen los datos que arrojan los sistemas transaccionales es necesario contar con herramientas que permitan el manejo y la interpretación de los datos de forma tal que pueda ser fácilmente comprensible por los usuarios finales.

6. Los sistemas de apoyo a la toma de decisiones basados en web representan un beneficio sumamente importante en las empresas ya que el actual crecimiento del internet posibilita que los directivos de las empresas puedan hacer uso de sus sistemas en el momento que lo necesiten y sin necesidad de estar en su lugar de trabajo sino que simplemente pueden acceder a dichos sistemas gracias a las redes de computadoras que actualmente existen.

7. Los sistemas de reportes permiten a los usuarios finales ver la información lo más resumida posible y dan la opción de generación de gráficos lo cual permite detectar patrones de conducta que pueden ayudar a dichos usuarios a la hora de la toma de decisiones.

8. Para que los sistemas de apoyo a la toma de decisiones sean exitosos los mismos deben contener los indicadores que representan datos críticos para las empresas, es por ello que es necesario realizar un buen levantamiento de información y determinar indicadores concretos que realmente le ofrezcan información a los usuarios al momento de la toma de decisión.

RECOMENDACIONES

- Se recomienda el diseño e implementación de un instrumento de evaluación para los preparadores y ayudantes de cátedra que permita obtener la opinión de los estudiantes, del profesor de la materia y del coordinador de preparadurías.

- Se recomienda la aplicación del estándar de nomenclaturas de base de datos para futuros desarrollos que se lleven a cabo dentro de la Escuela de Ingeniería Informática.

- Se recomienda la revisión de los sistemas transaccionales existentes para verificar si poseen algún patrón de diseño, en caso de no poseerlo es recomendable modificar la aplicación para aplicarle un patrón de diseño y los estándares de codificación.

- Se recomienda el diseño e implementación de un sistema orientado a mantener contacto directo con los egresados de la escuela, con la finalidad de obtener información que puede ser necesaria en determinada situación.

- Se recomienda que los archivos que sean proporcionados por las distintas coordinaciones de la escuela sean archivos confiables, es decir, que la información proporcionada sea la información definitiva y se encuentre en estado consistente.

- Se recomienda concientizar a los usuarios finales de los sistemas transaccionales de la importancia que tienen los datos que estén manipulando ya que estos servirán de fuente para futuras decisiones que tome la institución.

REFERENCIAS BIBLIOGRÁFICAS

LIBROS

[1] Cardoso, L (2006). *Sistemas de Bases de Datos II: teoria aplicada para profesores y estudiantes.* (1ª. Ed.). Venezuela: Publicaciones Ucab.

[2] Gray, P (1998). *"Decision support in the data warehouse". (2ª. Ed.).* Upper Saddle River, NJ : Prentice Hall.

[3] Groth, R (2000). *"Data mining : building competitive advantage". (2ª. Ed.).* Upper Saddle River, NJ : Prentice Hall.

[4] Inmon, W (1996). *"Building the Datawarehouse". (2ª. Ed.).* New York : Wiley Computer Pub.

[5] McConnell Steve (1996). *Desarrollo y gestión de proyectos informáticos.* (1ª. Ed.). Madrid: McGraw-Hill.

[6] *Microsoft SQL Server 7.0 Data Warehousing Training Kit. (2000).* Washington: Microsoft Press.

[7] Navathe, S.B. y Elmasri, R.A.(2002). *Fundamentos de Sistemas de Bases de Datos.* (3ª. Ed.). España: Pearson Educación.

[8] O'Brien, J.A. y Marakas, G.M. (2006). *Sistemas de Información Gerencial.* (7ª. Ed.). México: McGraw-Hill.

[9] Silberschatz, A., Korth ,H. y Sudarshan, S. (2002). *Fundamentos de Bases de Datos.* (4ª. Ed.). Madrid: McGraw-Hill.

PÁGINAS WEB

[10] Bernabeu Ricardo (2007). *DATA WAREHOUSING: Investigación y Sistematización de Conceptos, Arquitectura del Data Warehouse* [Página Web en línea].
Disponible en:
http://www.dataprix.com/es/arquitectura-del-data-warehouse#x1-710003.6.2
[Consulta: 2009, Marzo 31].

[11] Bernabeu Ricardo (2007). *DATA WAREHOUSING: Investigación y Sistematización de Conceptos, Otros Conceptos a Tener en Cuenta* [Página Web en línea].
Disponible en:
http://www.dataprix.com/es/otros-conceptos-tener-cuenta
[Consulta: 2009, Marzo 21].

[12] Escamilla Claudia (Fecha no disponible). *Sistemas de Soporte de Decisiones en Pro de la Educación* [Página Web en línea]
Disponible en:
http://www.deltaasesores.com/articulos/autores-invitados/otros/3016-sistemas-de-soporte-de-decisiones-en-pro-de-la-educacion
[Consulta: 2009, Agosto 08]

[13] Instituto Nacional de Seguridad e Higiene en el Trabajo (Fecha no disponible). *Indicadores* [Página Web en línea]
Disponible:
http://www.oect.es/portal/site/Observatorio/
[Consulta: 2008, Diciembre 01].

[14] Rada Gabriel (2007). *Atributos de un buen indicador* [Página Web en línea].

Disponible en:
http://escuela.med.puc.cl/Recursos/recepidem/insIntrod9d.htm
[Consulta: 2008, Diciembre 01].

[15] Sinnexus (2007). *Datawarehouse* [Página Web en línea].
Disponible:
http://www.sinnexus.com/business_intelligence/datawarehouse.aspx
[Consulta: 2008, Diciembre 01].

[16] Universidad de Belgrano (2008). *Datawarehouse* [Página Web en línea].
Disponible:
http://www.ub.edu.ar/catedras/ingenieria/ing_software/ubftecwwwdfd/ora-data/oradata.htm
[Consulta: 2008, Noviembre 29].

DOCUMENTOS ELECTRÓNICOS

[17] Garza Ernesto (2002). *Nivel de implementación de la tecnología Da Warehouse en grandes empresas.* [Documento en línea]
Disponible en:
http://copernico.mty.itesm.mx/phronesis/graduados/scripts/get_text.cgi?nombre_file=ITESMMTY2002215.pdf
[Consulta: 2008, Diciembre 02]

[18] Pérez Erith (2008). *Data Warehouse, Modelo, Conceptos e Implementación orientada a SQL Server* [Documento en Línea].
Disponible en:
http://www.fileden.com/files/2007/10/3/1478941/data-warehouse-sql.zip
[Consulta: 2008, Noviembre 30].

TRABAJOS ESPECIALES DE GRADO

[19] Berroterán, D y Gamboa, E. (2007). *Sistema de Información de apoyo para la toma de decisiones relacionadas con la facturación, pedidos y necesidades de recursos humanos de la empresa Flamberge.* Tesis de Grado. Ing. Informático. Escuela de Ingeniería Informática, Facultad de Ingeniería, Universidad Católica Andrés Bello, Caracas, Venezuela.

PUBLICACIONES PERIÓDICAS ELECTRÓNICAS

[20] Chaudhuri, S. y Dayal, U. (1997, Marzo). An Overview of Data Warehousing and OLAP Technology. *SIGMOD Record, vol. 26* [EN LINEA], Nº 1.

Disponible en:

http://www.cs.sfu.ca/CC/459/han/papers/chaudhuri97.pdf

[Consulta: 2009, Abril 15]

[21] DELGADO, Mª del C.(2002, Marzo). Los indicadores educativos. Estado de la cuestión y uso en geografía. *Biblio 3W, Revista Bibliográfica de Geografía y Ciencias Sociales* [EN LINEA], Nº 354.

Disponible en:

http://www.ub.es/geocrit/b3w-354.htm

[Consulta: 2009, Marzo 30]

TRABAJOS NO PUBLICADOS

[22] Cardoso, L. (2009). *Plan de la Calidad para la Escuela de Ingeniería Informática de la Universidad Católica Andrés Bello.* Trabajo no publicado. Trabajo de Grado, Especialista en Sistema de la Calidad, Universidad Católica Andrés Bello, Caracas.

GLOSARIO

C

Cubo: Un cubo OLAP es una vista lógica de los datos y representan un conjunto de datos relacionados con un tema particular como Ventas, Operaciones, etc., y están a disposición de los clientes a quienes les puedan interesar.

D

Data Cleansing: Proceso de asegurar que todos los valores en un conjunto de datos sean consistentes y correctamente registrados.

Datamart: Un Datamart es una solución que, compartiendo tecnología con el Data Warehouse (pero con contenidos específicos, volumen de datos más limitado y un alcance histórico menor), permita dar soporte a una empresa pequeña, o un departamento o área de negocio de una empresa grande. El Datamart cubre las necesidades de informes bien porque no es conveniente efectuar consultas sobre el sistema transaccional, bien porque hay que integrar datos de varios sistemas transaccionales.

Data Warehouse: Es una colección de datos orientado a temas, integrado, no volátil, de tiempo variante, que se usa para el soporte del proceso de toma de decisiones gerenciales. Un data warehouse es una colección de datos en la cual se encuentra integrada la información de la Institución y que se usa como soporte para el proceso de toma de decisiones gerenciales.

Desnormalización: Conjunto de técnicas empleadas en el diseño de una Base de Datos Relacional para mejorar el rendimiento de las consultas, creando redundancia de datos.

Dimensión: Son catálogos de información complementaria necesaria para la

presentación de los datos a los usuarios. Por ejemplo: tiempo, zonas, etc.

E

Esquema Constelación: Para cada esquema estrella o esquema de copo de nieve en almacén de datos es posible construir un esquema de constelación de hechos. Este esquema es más complejo que las otras arquitecturas debido al factor de que contiene múltiples tablas de hechos. Con esta solución, las tablas de dimensiones pueden estar compartidas entre más de una tabla de los hechos.

Extracción, Transformación y Transporte (ETT): Los procesos de Extracción, Transformación y Transporte constan de múltiples pasos, cuyo objetivo es transferir datos desde las aplicaciones de producción a los sistemas de Inteligencia de negocio.

G

Granularidad: La granularidad es el nivel de detalle de los datos a manipular.

M

Medida: Una Medida es una propiedad de un Hecho (casi siempre numérica), que es usada para su análisis.

N

Normalización: Es el proceso mediante el cual se transforman datos complejos a un conjunto de estructuras de datos más pequeñas, que además de ser más simples y más estables, son más fáciles de mantener.

O

OLAP: El término OLAP, que proviene de Online Analitical Processing

(Procesamiento Analítico en Línea), define a una tecnología que se basa en el análisis multidimensional de los datos y que le permite al usuario tener una visión más rápida e interactiva de los mismos.

OLTP: Es un tipo de proceso especialmente rápido en el que las solicitudes de los usuarios son resueltas de inmediato; naturalmente, ello implica la concurrencia de un «mecanismo» que permite el procesamiento de varias transacciones a la vez.

Open Source: término con el que se conoce al software distribuido y desarrollado libremente.

R

ROLAP (Relational OLAP): Tanto los datos precalculados y agregados como los datos fuente residen en la misma base de datos relacional. Si el DataWarehouse es muy grande o se necesita rapidez por parte de los usuarios puede ser un problema.

S

Sistema de Soporte a la toma de decisiones (SSD): Un conjunto de Procedimientos basados en un modelo que permiten procesar datos y generar juicios para facilitar la toma de decisiones de los administradores

SQL: El Lenguaje de Consulta Estructurado (Structured Query Language) es un lenguaje declarativo de acceso a bases de datos relacionales que permite especificar diversos tipos de operaciones sobre las mismas.

T

Tabla de Hechos: tabla central de un esquema dimensional (en estrella o en copo de nieve) y contiene los valores de las medidas de negocio.