



Universidad de San Carlos de Guatemala  
Facultad de Ingeniería  
Escuela de Ingeniería en Ciencias y Sistemas

**INTELIGENCIA DE NEGOCIOS ENFOCADA A LA EDUCACIÓN PÚBLICA  
DE NIVEL PRIMARIO PARA EL SEGUIMIENTO Y APOYO A ALUMNOS  
Y TOMA DE DECISIONES ACADÉMICAS**

**Walter Alejandro Cardona López**

Asesorado por el Ing. Luis Fernando Quiñónez López

Guatemala, julio de 2014

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**INTELIGENCIA DE NEGOCIOS ENFOCADA A LA EDUCACIÓN PÚBLICA  
DE NIVEL PRIMARIO PARA EL SEGUIMIENTO Y APOYO A ALUMNOS  
Y TOMA DE DECISIONES ACADÉMICAS**

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA  
FACULTAD DE INGENIERÍA

POR

**WALTER ALEJANDRO CARDONA LÓPEZ**

ASESORADO POR EL ING. LUIS FERNANDO QUIÑÓNEZ LÓPEZ

AL CONFERÍRSELE EL TÍTULO DE

**INGENIERO EN CIENCIAS Y SISTEMAS**

GUATEMALA, JULIO DE 2014

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA  
FACULTAD DE INGENIERÍA



**NÓMINA DE JUNTA DIRECTIVA**

DECANO	Ing. Murphy Olympto Paiz Recinos
VOCAL I	Ing. Alfredo Enrique Beber Aceituno
VOCAL II	Ing. Pedro Antonio Aguilar Polanco
VOCAL III	Inga. Elvia Miriam Ruballos Samayoa
VOCAL IV	Br. Narda Lucía Pacay Barrientos
VOCAL V	Br. Walter Rafael Véliz Muñoz
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

**TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO**

DECANO	Ing. Murphy Olympto Paiz Recinos
EXAMINADOR	Ing. Luis Fernando Quiñónez López
EXAMINADOR	Ing. Roberto Estuardo Ruiz Cruz
EXAMINADOR	Ing. César Augusto Fernández Caceres
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

## HONORABLE TRIBUNAL EXAMINADOR

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

### **INTELIGENCIA DE NEGOCIOS ENFOCADA A LA EDUCACIÓN PÚBLICA DE NIVEL PRIMARIO PARA EL SEGUIMIENTO Y APOYO A ALUMNOS Y TOMA DE DECISIONES ACADÉMICAS**

Tema que me fuera asignado por la Dirección de la Escuela de Ingeniería en Ciencias y Sistemas, con fecha enero de 2013.

  
Walter Alejandro Cardona López



Universidad San Carlos de Guatemala  
Facultad de Ingeniería  
Escuela de Ingeniería en Ciencias y Sistemas

Guatemala, 05 de Septiembre del 2,013


Ingeniero  
Carlos Alfredo Azurdia

Respetable Ingeniero Azurdia:

Por medio de la presente le informo que he revisado y asesorado el proyecto de tesis del estudiante **Walter Alejandro Cardona López**, con carné **200614901**, titulado **“INTELIGENCIA DE NEGOCIOS ENFOCADA A LA EDUCACION PUBLICA DE NIVEL PRIMARIO PARA EL SEGUIMIENTO Y APOYO A ALUMNOS Y TOMA DE DECISIONES ACADEMICAS”**, y según mi criterio se han alcanzado y cumplido los objetivos propuestos para su desarrollo.

Agradeciendo su atención a la presente,

Atentamente,

  
Ingeniero  
Luis Fernando Quiñonez López  
Colegiado No 7514  
Ing. Luis Fernando Quiñonez López  
Colegiado 7514



Universidad San Carlos de Guatemala  
Facultad de Ingeniería  
Escuela de Ingeniería en Ciencias y Sistemas

Guatemala, 15 de Octubre de 2013

Ingeniero  
**Marlon Antonio Pérez Turk**  
Director de la Escuela de Ingeniería  
En Ciencias y Sistemas

Respetable Ingeniero Pérez:

Por este medio hago de su conocimiento que he revisado el trabajo de graduación del estudiante **WALTER ALEJANDRO CARDONA LÓPEZ**, con carné **2006-14901**, titulado: **"INTELIGENCIA DE NEGOCIOS ENFOCADA A LA EDUCACIÓN PÚBLICA DE NIVEL PRIMARIO PARA EL SEGUIMIENTO Y APOYO A ALUMNOS Y TOMA DE DECISIONES ACADEMICAS"**, y a mi criterio el mismo cumple con los objetivos propuestos para su desarrollo, según el protocolo.

Al agradecer su atención a la presente, aprovecho la oportunidad para suscribirme,

Atentamente,

  
**Ing. Carlos Alfredo Azurdia**  
Coordinador de Privados  
y Revisión de Trabajos de Graduación



E  
S  
C  
U  
E  
L  
A  
  
D  
E  
  
C  
I  
E  
N  
C  
I  
A  
S  
  
Y  
  
S  
I  
S  
T  
E  
M  
A  
S

UNIVERSIDAD DE SAN CARLOS  
DE GUATEMALA



FACULTAD DE INGENIERÍA  
ESCUELA DE CIENCIAS Y SISTEMAS  
TEL: 24767644

*El Director de la Escuela de Ingeniería en Ciencias y Sistemas de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del asesor con el visto bueno del revisor y del Licenciado en Letras, del trabajo de graduación **“INTELIGENCIA DE NEGOCIOS ENFOCADA A LA EDUCACIÓN PÚBLICA DE NIVEL PRIMARIO PARA EL SEGUIMIENTO Y APOYO A ALUMNOS Y TOMA DE DECISIONES ACADÉMICAS”**, realizado por el estudiante WALTER ALEJANDRO CARDONA LÓPEZ, aprueba el presente trabajo y solicita la autorización del mismo.*

**“ID Y ENSEÑAD A TODOS”**

*Ing. Marlon Antonio Pérez Türk  
Director, Escuela de Ingeniería en Ciencias y Sistemas*




*Guatemala, 18 de julio 2014*



DTG. 352.2014

El Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Ingeniería en Ciencias y Sistemas, al Trabajo de Graduación titulado: **INTELIGENCIA DE NEGOCIOS ENFOCADA A LA EDUCACIÓN PÚBLICA DE NIVEL PRIMARIO PARA EL SEGUIMIENTO Y APOYO A ALUMNOS Y TOMA DE DECISIONES ACADÉMICAS**, presentado por el estudiante universitario **Walter Alejandro Cardona López**, y después de haber culminado las revisiones previas bajo la responsabilidad de las instancias correspondientes, se autoriza la impresión del mismo.

IMPRÍMASE:

  
ing. Murphy Olympo Paiz Recinos  
Decano



Guatemala, 24 de julio de 2014

/gdech



## **ACTO QUE DEDICO A:**

- Dios** Por la vida, las oportunidades y un camino pintado de bendiciones.
- Mis padres** Walter Cardona y Miriam López de Cardona. Su amor será siempre mi inspiración y su esfuerzo una motivación para seguir adelante.
- Mis hermanos** Gabriela y Kevin Cardona. Por su eterna amistad y compañía; hemos sido amigos y su presencia me motivará siempre.
- Mis abuelos** Paula Patzán, Félix Cardona, Nicolás López y María Gertrudis Liquez. Su enseñanza, sacrificio y filosofía sigue y seguirá en sus nietos.
- Mi madrina** Ligia Lorena López Liquez. Por su amor incondicional y presencia en cada etapa de mi vida.
- Mi tío** Rudy Cardona. Por sus consejos, ánimos, cariño sincero y amistad, a pesar de la distancia.

**Mis hermanos de tiempo**

Jorge Ávila, Juan Carlos Piedrasanta y Félix Barrios. Por su amistad y hermandad a lo largo de toda mi vida.

## **AGRADECIMIENTOS A:**

<b>La Universidad de San Carlos de Guatemala</b>	Por ser mi casa de estudios y brindarme un lugar donde aprender a ser alguien en la vida.
<b>Facultad de Ingeniería</b>	Por darme la oportunidad de crecer y aprender.
<b>Mis tíos</b>	Ligia, Griselda, Mardoqueo López y Rudy Cardona. Por su apoyo y presencia en cada etapa de mi vida.
<b>Mis primos</b>	Por su amistad, compañía y alegría; son únicos.
<b>Mis amigos de la Facultad</b>	Luis Donis, Luis Fernando López, Marco Guzmán, Luis Felipe Méndez, Jacqueline Hernández, Esteban Marroquín, Otto García, Andrea Grimaldi, Héctor Miranda y Jorge Mario Arriaza. Por su amistad y presencia en diferentes etapas de mi carrera.
<b>Las familias Ávila Enríquez, Piedrasanta Ovalle y Barrios Ordóñez</b>	Por su apoyo y cariño desde siempre.

**Marco G3ngora**

Por ser mi primer maestro en el ambiente laboral y mi amigo desde entonces.

**Inform3tica AGEXPORT**

Por su compa1erismo, ense1anzas, aceptaci3n y sobre todo amistad en mi primer trabajo.

## ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES.....	V
LISTA DE SÍMBOLOS .....	VII
GLOSARIO .....	IX
RESUMEN.....	XI
OBJETIVOS.....	XIII
INTRODUCCIÓN.....	XV
1. SITUACIÓN ACTUAL.....	1
1.1. Ministerio de Educación.....	1
1.1.1. Índice de avance educativo .....	3
1.1.2. Subindicadores .....	4
1.1.3. Índice de avance educativo departamental .....	4
1.2. Escuelas públicas.....	6
1.2.1. Flujo de información .....	6
1.2.2. Aspectos a mejorar.....	8
1.2.3. Tecnología institucional .....	9
2. TÉCNICAS DE INTELIGENCIA DE NEGOCIOS APLICADA AL SEGUIMIENTO EDUCATIVO .....	13
2.1. Inteligencia de negocios .....	13
2.1.1. Nivel operativo.....	14
2.1.2. Nivel táctico .....	15
2.1.3. Nivel estratégico .....	15
2.2. Data Warehouse.....	16
2.3. Data Mining .....	18

2.4.	OLAP.....	19
2.5.	OLTP.....	21
2.6.	Ventajas y beneficios de la inteligencia de negocios en la educación.....	22
3.	MODELO DE DATOS LÓGICO .....	25
3.1.	Modelo de datos.....	29
3.1.1.	Modelo estrella para notas de estudiantes.....	29
3.1.2.	Modelo estrella para asistencia y deserción.....	31
3.1.3.	Modelo estrella para situación socioeconómica del estudiante.....	32
3.1.4.	Modelo estrella para salud estudiantil .....	33
3.2.	Dimensiones del modelo lógico de datos .....	34
3.2.1.	Dimensión estudiante.....	34
3.2.1.1.	Atributos asociados:.....	34
3.2.2.	Dimensión catedrático.....	35
3.2.2.1.	Atributos asociados:.....	35
3.2.3.	Dimensión materia .....	35
3.2.3.1.	Atributos asociados:.....	36
3.2.4.	Dimensión bimestre.....	36
3.2.5.	Dimensión grado .....	36
3.2.6.	Dimensión año .....	37
3.2.7.	Dimensión dirección .....	37
3.2.8.	Dimensión padre .....	38
3.2.9.	Dimensión madre .....	39
3.2.10.	Dimensión encargado .....	40
3.2.11.	Dimensión tipo hogar .....	40
3.2.12.	Dimensión enfermedad .....	41

4.	PROPUESTA DE ARQUITECTURA .....	43
4.1.	Limitaciones.....	43
4.2.	Requerimientos no funcionales .....	44
4.3.	Límites .....	45
4.4.	Casos de uso de alto nivel.....	46
4.4.1.	Casos de uso para el ingreso de información.....	46
4.4.2.	Casos de uso para la visualización de reportes operacionales .....	48
4.4.3.	Casos de uso para la generación de reportes tácticos .....	49
4.4.4.	Casos de uso para el monitoreo estratégico .....	50
4.5.	La estructura básica de la arquitectura Data Warehouse .....	51
4.6.	Arquitectura propuesta .....	53
4.6.1.	Justificación de la propuesta.....	54
4.6.2.	Especificación de la propuesta .....	55
4.7.	Requisitos de hardware .....	57
4.7.1.	Equipo para base de datos transaccional.....	58
4.7.2.	Equipo para el repositorio de datos .....	60
4.7.3.	Equipo de red .....	61
4.7.4.	Ambiente y ubicación del equipo .....	62
5.	DEFINICIÓN DE MÓDULOS.....	65
5.1.	Organigrama institucional .....	66
5.2.	Módulo operativo .....	67
5.2.1.	Reporte de notas por alumno .....	68
5.2.2.	Reporte de asistencia .....	69
5.3.	Módulo táctico .....	70
5.3.1.	Reporte de notas de estudiantes .....	71
5.3.2.	Reporte de asistencia y deserción.....	71

5.3.3.	Reporte socioeconómico .....	72
5.3.4.	Reporte salud estudiantil .....	72
5.4.	Módulo estratégico .....	72
5.4.1.	<i>Dashboard</i> de notas .....	73
5.4.2.	<i>Dashboard</i> de deserción .....	74
5.4.3.	<i>Dashboard</i> socioeconómico .....	74
5.4.4.	<i>Dashboard</i> de salud estudiantil .....	74
CONCLUSIONES .....		75
RECOMENDACIONES .....		77
BIBLIOGRAFÍA .....		79



# ÍNDICE DE ILUSTRACIONES

## FIGURAS

1.	Índice de avance educativo.....	3
2.	Subindicadores de avance educativo.....	4
3.	Índice de avance educativo departamental.....	5
4.	Niveles de toma de decisiones.....	14
5.	Modelo dimensional para notas de estudiantes (esquema en estrella).....	30
6.	Modelo dimensional para asistencia y deserción (esquema en estrella).....	31
7.	Modelo dimensional para situación socioeconómica del estudiante (esquema en estrella).....	32
8.	Modelo dimensional para salud estudiantil (esquema en estrella).....	33
9.	Casos de uso ingreso de información.....	47
10.	Casos de uso visualización de reportes operacionales.....	48
11.	Casos de uso para la generación de reportes tácticos.....	49
12.	Casos de uso para el monitoreo estratégico.....	50
13.	Estructura básica de Data Warehouse.....	52
14.	Arquitectura propuesta.....	53
15.	Organigrama institucional.....	66

## TABLAS

I.	Especificaciones de hardware.....	59
II.	Especificaciones de hardware.....	60

III.	Formato de reporte de notas por alumno.....	68
IV.	Formato de reporte de asistencia .....	69

## LISTA DE SÍMBOLOS

<b>Símbolo</b>	<b>Significado</b>
<b>DMt</b>	Data Mart
<b>DM</b>	Data Mining
<b>DW</b>	Data Warehouse
<b>IAE</b>	Índice de avance educativo



## GLOSARIO

<b>Data Mart</b>	Subconjunto de un Data Warehouse que está orientado a un área específica del negocio. Todas sus métricas y dimensiones están relacionadas con un área de negocio en particular.
<b>Data Warehouse</b>	Sistema de información centralizado que contiene toda la información relevante para una organización, y que permite de una forma ágil y flexible la consulta de información. Está orientado a convertirse en la única fuente de información para todas las áreas del negocio y ser la fuente prioritaria para todas las herramientas de inteligencia de negocios que se deseen utilizar en una organización.
<b>Dimensión</b>	Se refiere a la entidad que se encarga de agrupar, calificar o catalogar cada uno de los hechos contenidos dentro de un Data Mart.
<b>Métrica</b>	Son aquellos datos que implican un valor relacionado con un hecho del negocio. Son siempre valores numéricos propensos a ser utilizados para cálculos o sumatorias, para la obtención de cualquier valor agregado sobre la información del negocio.

<b>OLAP</b>	Sistema de procesamiento analítico en línea, es decir sistemas orientados al análisis de información, basado en un Data Mart o un Data Warehouse.
<b>OLTP</b>	Sistema de procesamiento de transacciones en línea o sistemas transaccionales, en las cuales residen las operaciones diarias del negocio y que son las fuentes de datos para el Data Warehouse y sus respectivos Data Marts.
<b>Tabla de hechos</b>	Tabla de datos central que contiene cada una de las métricas que se desea establecer y medir dentro de un área específica del negocio.

## RESUMEN

La inteligencia de negocios es el proceso por el cual se obtiene información existente en los datos que posee una organización. En el mundo del avance tecnológico las empresas buscan la ayuda de la tecnología para solventar los problemas con los que se cuenta y mejorar sus procesos; esto lleva a que dichas empresas posean gran cantidad de datos, recolectados en los sistemas transaccionales utilizados en sus procesos comunes y que son esenciales para las operaciones del negocio.

Tener a disposición datos del negocio no es precisamente disponer de información útil para el mismo; para poseer información se requiere de una transformación de los datos que se poseen y así lograr comprender correctamente el funcionamiento del mismo. La inteligencia de negocios brinda una respuesta a tales situaciones, proporcionando a los encargados de tomar las decisiones en cada uno de los niveles de la organización, una mejor respuesta ante los cambios que se presentan en el ambiente en el que la organización compete.

Para poder obtener una solución a un problema de la organización por medio de la inteligencia de negocios es necesario poseer información proveniente de las fuentes transaccionales con las que se posee. Estos datos se consolidan en un repositorio de hechos históricos denominado Data Warehouse, que es una base de datos diseñada para la consulta y análisis de la información.

El objetivo central del presente trabajo consiste en plantear una solución de inteligencia de negocios a las necesidades del manejo de información para el análisis y toma de decisiones académicas de los catedráticos y autoridades de una escuela pública de nivel primario, para el seguimiento y apoyo en el rendimiento estudiantil, estableciendo los factores que afectan el rendimiento académico, la información necesaria para la construcción de un modelo lógico de datos y una arquitectura viable para su posible implementación futura.



## **OBJETIVOS**

### **General**

Plantear una solución de inteligencia de negocios para aportar una ayuda al personal docente y administrativo de las escuelas públicas, para el seguimiento de sus estudiantes, aportando así la capacidad de tomar decisiones en diferentes situaciones futuras.

### **Específicos**

1. Determinar la información necesaria para brindar una solución de inteligencia de negocios para el apoyo en el rendimiento estudiantil.
2. Determinar las herramientas y técnicas de inteligencias de negocios que pueden aportar una ayuda a la toma de decisiones en un ambiente escolar.
3. Diseñar un modelo de datos para determinar el manejo de información en una inteligencia de negocios enfocada a la educación pública a nivel primario.
4. Definir la arquitectura de una solución de inteligencia de negocios con los recursos con los que cuenta una escuela pública guatemalteca.
5. Definir los reportes necesarios para tomar decisiones sobre el rendimiento académico.



## INTRODUCCIÓN

La inteligencia de negocios o BI (*business intelligence* por sus siglas en inglés), es un conjunto de técnicas y herramientas enfocadas al manejo, administración y creación de conocimiento mediante el análisis de datos existentes en una organización o empresa. Dichos datos son transformados en información de valor para la institución, utilizada para la toma de decisiones; esto comprende saber el funcionamiento interno de la empresa para poder ver a futuro qué y cómo se realizará, buscando así poder predecir tendencias y aplicar cambios con base en decisiones tomadas de los datos transformados.

El presente informe describe el planteamiento de una solución de inteligencia de negocios en un ambiente educativo del sector público a nivel primario, tomando en cuenta los factores importantes que influyen en el ambiente educativo, así como las necesidades que este posee para el seguimiento estudiantil.

Se tomaron en cuenta los factores que influyen en el rendimiento académico de los estudiantes y se buscó plantear una solución fiable para la toma de decisiones, evaluando cada uno de las posibilidades que brinda la inteligencia de negocios, para hacer de dicha toma de decisiones un punto de partida en la mejora de la situación estudiantil que sufre el alumnado de una escuela pública de nivel primario.



# **1. SITUACIÓN ACTUAL**

En el presente capítulo se expondrá el ambiente en el que se encuentra el manejo de información de los estudiantes tanto a nivel del Ministerio de Educación como de una institución pública individual, debido a que es necesario saber de dónde se partirá para poder buscar un mejor planteamiento del problema.

La educación en Guatemala es un tema que constantemente se plantea como una llave al crecimiento, tanto económico como competitivo, a nivel internacional. Los temas en campañas políticas exponen a la educación como uno de sus más importantes aspectos a tener en cuenta en su futura agenda de gobierno. Aunque no se discutirá el tema de construcción de nuevos centros educativos, se debe de analizar que esto no es suficiente, puesto que el poder darle una educación íntegra a una población con índices de pobreza, es solo el comienzo para el cambio y el progreso.

El manejo de información a tratar consiste en dos ambientes: el primero es la manera en la cual el Ministerio de Educación administra la información y la segunda se refiere a cómo las escuelas administran su propia información.

## **1.1. Ministerio de Educación**

Actualmente el Ministerio de Educación de Guatemala maneja información respecto de los alumnos, tanto en el ambiente privado como público.

La información es reportada por cada una de las instituciones encargadas de impartir la educación en el territorio nacional; dicha información es mostrada al público por medio de su sitio en internet (<http://www.mineduc.gob.gt/portal/index.asp>). En el área de estadística se puede observar tanto el número de estudiantes inscritos, como el de estudiantes aprobados y el avance académico agrupado por año, localidad y grado.

Estos datos son obtenidos de las boletas de estadística que son recolectadas en cada uno de los establecimientos educativos, ya sea privados o públicos y de todos los niveles educativos, en el periodo al que hacen referencia al final de cada año. Estas estadísticas son utilizadas con la finalidad de proveer al Ministerio de Educación y usuarios en general del sistema, la información confiable y oportuna para la toma de decisiones, entrega de servicios de apoyo, definición de políticas y diseño de planificación de la educación.

Estas finalidades son parte de un tipo de mejora para poder tener el control y así tomar decisiones y brindar ayuda en ciertas áreas a los distintos establecimientos. Podría decirse que es un tipo de inteligencia de negocios por el fin que persigue.

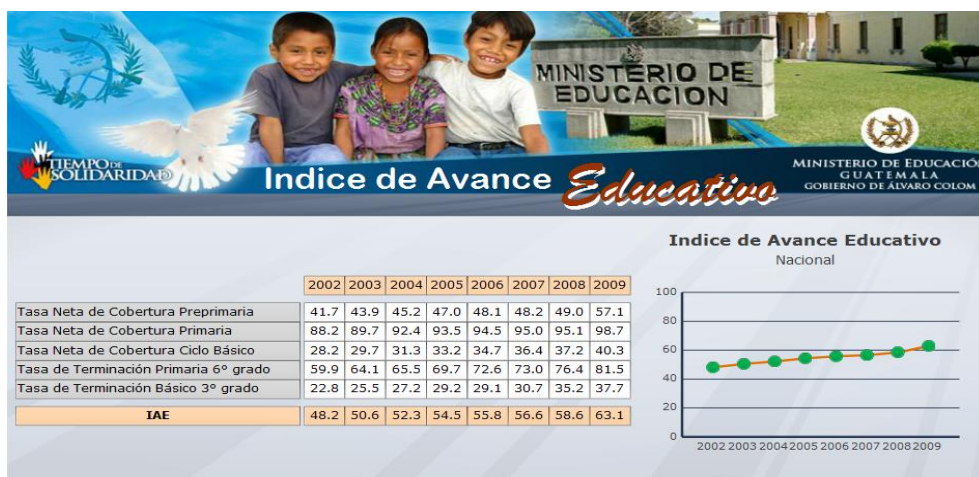
La información que se posee se maneja a través de gráficas que muestran las tendencias de la educación, basándose netamente en el índice de avance educativo, el cual permite saber información acerca de cada departamento y municipio, respecto de si están o no alcanzando las metas para poder focalizar los recursos y esfuerzos sobre aquellos que estén más lejos de alcanzar las metas propuestas.

Esto les permite como Ministerio poder tomar decisiones, viendo un ambiente general de la educación por cada sector de la República, para tomar medida y acciones para avanzar en la educación que se proponen, según sea el nivel educativo que evalúen; es decir primario, básico o diversificado. Dichas gráficas o informes son presentados de la siguiente manera:

### 1.1.1. Índice de avance educativo

En la figura 1 se muestra la gráfica que indica el porcentaje de avance educativo enfocado a la cobertura en cada nivel académico; es decir que indica el porcentaje de lo que se ha avanzado para dar educación en cada uno de los niveles académicos a nivel nacional. En esta área se maneja la información recolectada a nivel nacional, de tal manera que según el avance en cada nivel educativo, se puedan obtener los promedios en dicho avance y tomar decisiones respecto de los niveles más afectados.

Figura 1. Índice de avance educativo



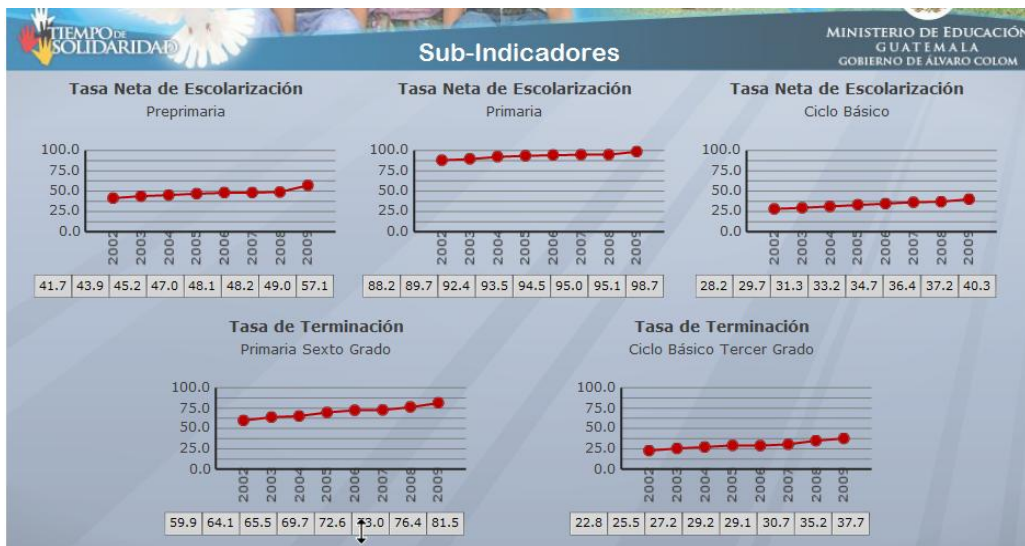
Fuente: <http://www.mineduc.gob.gt/portal/contenido/varios/indiceAvanceEducativo/iae.html>.

Consulta: 1 de mayo de 2012.

### 1.1.2. Subindicadores

Los subindicadores se muestran de tal manera que pueda comprender más detalladamente el avance educativo en cada uno de los niveles académicos, tomando el 100 % como la cobertura educativa deseada o propuesta para dicho periodo.

Figura 2. Subindicadores de avance educativo



Fuente: <http://www.mineduc.gob.gt/portal/contenido/varios/indiceAvanceEducativo/iae.html>.

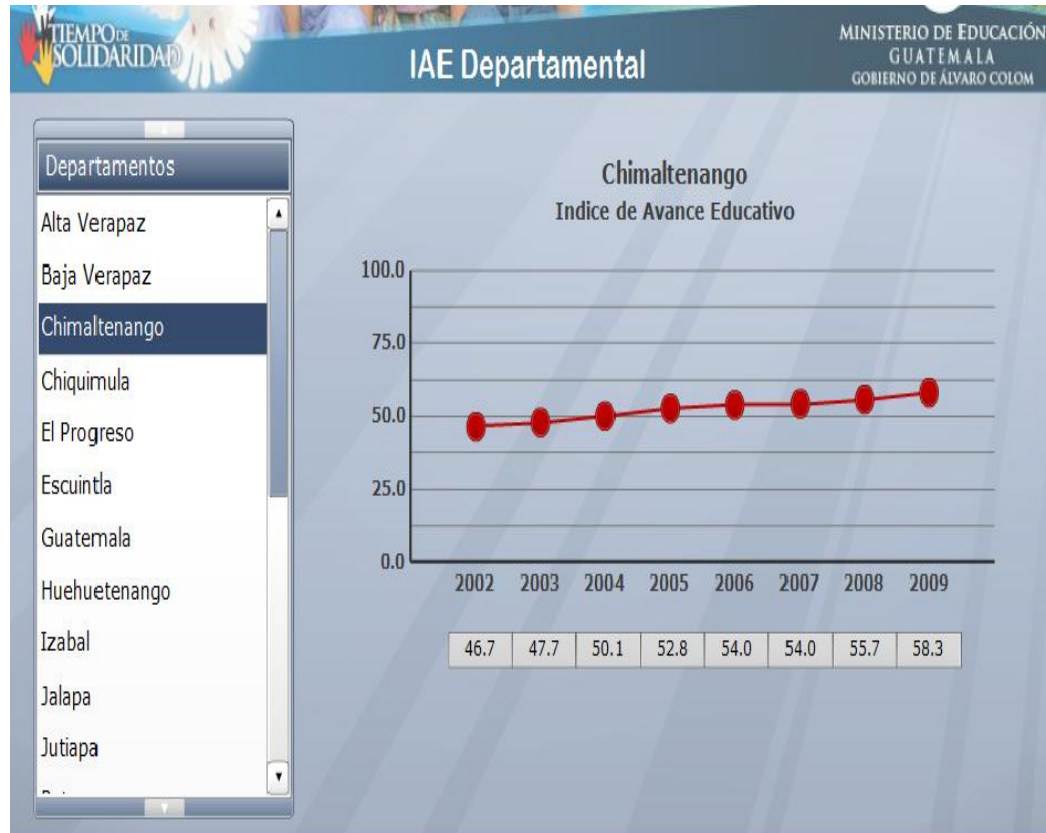
Consulta: 01 de mayo de 2012.

### 1.1.3. Índice de avance educativo departamental

En la figura 3 se reflejan los índices de avance educativo por departamento, el cual se establece también de manera muy generalizada.



Figura 3. Índice de avance educativo departamental



Fuente: <http://www.mineduc.gob.gt/portal/contenido/varios/indiceAvanceEducativo/Dep.swf>.

Consulta: 1 de mayo de 2012.

Tomando este tipo de gráficas como punto de partida, se puede decir que aunque la idea de llevar un control de avance educativo en toda la República es una manera de tomar acción de parte del Ministerio de Educación, puede apreciarse que la información refleja en gran parte las áreas del país a las que se ha podido impartir educación con una información básica proporcionada por las diferentes escuelas o instituciones del país, ya sea educación pública o privada; pero no se posee una información más personalizada respecto de los alumnos que conforman el alumnado de cada una de estas instituciones.

## **1.2. Escuelas públicas**

La propuesta de la inteligencia de negocios aplicada en la educación se basa en la recolección de datos y transformarla en información útil para poder tomar medidas o decisiones, basándose en información más específica de cada estudiante. Esta propuesta en la situación actual no se practica, puesto que en el Ministerio de Educación de Guatemala aún no es posible tomar una decisión por el desempeño de cada uno de los estudiantes. Al enfocarse en las instituciones educativas, esto podría ser una herramienta que ayude al avance y seguimiento del rendimiento de cada estudiante.

Actualmente la tecnología no es una herramienta fundamental en la educación pública; es muy difícil que una escuela tenga acceso a laboratorio de cómputo e inclusive la información no se maneja del todo en una red informática.

Cada escuela cuenta con una boleta por cada uno de los alumnos inscritos y las notas son almacenadas y archivadas físicamente en la institución y luego son enviadas al Ministerio de Educación.

### **1.2.1. Flujo de información**

Como se mencionó anteriormente, cada alumno posee una boleta de inscripción, la cual lo identifica como estudiante de dicha institución; en esta se maneja información básica del futuro estudiante. La información manejada actualmente corresponde a los datos de la fe de edad de la persona, de la cual se extraen los datos específicos como el nombre del estudiante, sexo, partida y folio de la fe de edad. Asimismo se maneja la información de ambos padres tales como nombre completo, edad, y otras identificaciones de ambos padres.

Al ser inscrito por primera vez, a cada estudiante le es asignado un código por el Ministerio de Educación, el cual será su código estudiantil para el traslado de notas o envío de información final.

Respecto de los catedráticos de los establecimientos públicos a nivel primario, se maneja información relativamente básica, específicamente la información contenida en su curriculum vitae; este documento es almacenado por la institución pero no tiene un control específico sobre su contenido, es decir la información no es actualizada a lo largo del tiempo.

Como estrategia, cada uno de los catedráticos prepara el programa y calendarización de cada una de las materias impartidas, de tal manera en que se puedan cumplir los objetivos de enseñanza y cubrir el programa establecido.

Cada catedrático es responsable de entregar un informe al final de cada bimestre con las notas obtenidas por cada uno de sus estudiantes en cada una de las materias impartidas. Dichas notas e informes son almacenados dentro de la institución como registro, pero no son enviados al Ministerio de Educación.

A final de año cada institución presenta en un formato específico la documentación y cuadros de cada estudiante ante el Ministerio de Educación, únicamente con las notas finales en cada materia; esta es la única información que se presenta ante el Ministerio durante el año escolar.

Los cuadros con las notas finales de cada estudiante también son ingresados en una plataforma virtual utilizada por el Ministerio de Educación; dicha información no está a disposición de la institución educativa para la administración de la misma.

### **1.2.2. Aspectos a mejorar**

Con base en las entrevistas realizadas a la licenciada Ligia Lorena López Liques, psicóloga y directora de la Escuela Oficial Urbana Mixta No. 71 Germán Alcántara, se pueden recalcar los puntos a mejorar en el ámbito educativo a nivel de la información manejada por la institución educativa.

Entre los puntos clave que menciona la licenciada López está la necesidad de llevar el control estudiantil mediante una ficha pedagógica con los antecedentes más relevantes de los estudiantes, estableciendo la información social de los mismos como problemas de salud, situación social actual del estudiante e información más detallada respecto de los padres del alumno, para poder tener un control más detallado de los problemas que cada estudiante pueda presentar.

Según lo expuesto por la entrevistada, el catedrático juega un rol de guía respecto del alumno; este rol se manifiesta con la convivencia diaria entre el catedrático y sus estudiantes. El cambio en la conducta de determinado estudiante es notado por el catedrático debido a dicha convivencia. No existe un plan preventivo en relación con el posible cambio en el rendimiento académico del estudiante y se logran detectar problemas en el desempeño, únicamente en los estudiantes problemáticos o cuando el caso ya es muy delicado.

Las conclusiones y medidas respecto de la conducta y el rendimiento del estudiante son tomadas cuando se denota que el estudiante está con problemas más serios y muchas veces solo se tiene una vaga idea de cuál pueda ser la causa en los cambios en el alumno.

Esto se debe a la escasa información que se posee del alumno y a que si se tiene un conocimiento extra acerca del alumno es gracias a la capacidad de interacción del catedrático con sus estudiantes.

También es importante que el Ministerio de Educación provea de tecnología básica a las instituciones a nivel administrativo, puesto que según comenta la entrevistada no se posee acceso a internet en el caso de poseer equipo computacional; la educación en tecnología no está contemplada por el Ministerio de Educación y cuando se posee la capacidad de tener un laboratorio para la enseñanza no existe un maestro con la capacidad de compartir conocimiento tecnológico.

Puntos importantes a mejorar:

- La retroalimentación de parte del Ministerio de Educación respecto de las notas y situación actual de la institución con base en los cuadros finales que se presentan cada año.
- Estadísticas dentro de la misma institución educativa respecto de su estado académico actual.

### **1.2.3. Tecnología institucional**

Las escuelas públicas, especialmente a nivel primario, no cuentan con apoyo gubernamental en el aspecto tecnológico, sus capacidades de obtener material para la enseñanza computacional es un punto muy crítico para las instituciones especialmente en el área urbana.

Cuando una institución pública de enseñanza a nivel primario posee equipo de cómputo en sus instituciones, es debido a donaciones de personas o empresas para fomentar la educación computacional.

A nivel administrativo las instituciones no cuentan con equipo de cómputo para la realización de tareas; la mayor parte de los reportes o cuadros son elaborados en papel y no por medio de alguna herramienta de software, a excepción de los cuadros finales entregados al Ministerio de Educación.

Estos cuadros finales son realizados en su gran mayoría por los directores o catedráticos desde equipos personales y no a través de equipos institucionales; es decir que son elaborados desde sus propios recursos en algunos casos, ya que no todas las veces los miembros del claustro tienen acceso a las herramientas necesarias para la elaboración de estos cuadros finales.

Existen movimientos de organizaciones ajenas al gobierno que proveen equipo de cómputo o ayudan a la elaboración de laboratorios para la instalación del hardware que se posee; esto ocurre en su mayoría en instituciones en el área urbana. Estas organizaciones buscan mejorar la utilización de la tecnología que puede poseer una institución para la enseñanza en el área de computación.

Existe un movimiento emprendido por estudiantes de la Universidad de San Carlos de Guatemala llamado Edulibre, que busca llevar la educación tecnológica a las instituciones públicas, enfocándose en el nivel primario, proveyendo equipo aportado por medio de donaciones, administrándolo e instalándolo en laboratorios para la enseñanza pública.

La misión de las instituciones públicas en el tema de la tecnología va específicamente dirigida a su búsqueda, como herramienta para la enseñanza de computación hacia sus estudiantes y no para el uso administrativo.





## **2. TÉCNICAS DE INTELIGENCIA DE NEGOCIOS APLICADA AL SEGUIMIENTO EDUCATIVO**

En el presente capítulo se especificarán las diferentes técnicas de la inteligencia de negocios capaces de brindar la ayuda necesaria para la correcta administración de información y permitir la toma de decisiones a partir de la misma, con el objetivo primordial de dar un correcto seguimiento al rendimiento de los estudiantes en el ámbito educativo.

Para iniciar con la descripción de las técnicas de inteligencia de negocios es necesario empezar por definir ¿Qué es básicamente la inteligencia de negocios?, ¿qué enfoque se necesita brindar para los fines propuestos? De esa manera se podría determinar cuáles serán las técnicas apropiadas para el seguimiento del rendimiento educativo que se quiere aplicar.

### **2.1. Inteligencia de negocios**

La inteligencia de negocios o *Business Intelligence* (BI) se puede definir como el proceso de analizar los bienes o datos acumulados en una institución o empresa y extraer una cierta “inteligencia” o conocimiento de ellos. Es decir que es el proceso por el cual se puede analizar y tomar decisiones respecto de cierta situación, a partir de un conjunto de datos transformados en información, que brinda conocimiento respecto de la misma.

Dicha información puede ser accedida de manera rápida y sencilla y da la capacidad de analizarla para la toma de decisiones del negocio, tanto a nivel operativo, como táctico y estratégico.

Figura 4. **Niveles de toma de decisiones**



Fuente: elaboración propia.

En la figura anterior se muestra la pirámide de los niveles y sus diferentes herramientas para la toma de decisiones. Como se puede observar, existen diferentes herramientas o maneras de presentar la información según sea el nivel en que se encuentre la necesidad de evaluar la información y el objetivo que se pretende establecer con la decisión que se va a tomar.

### **2.1.1. Nivel operativo**

La inteligencia de negocios a nivel operativo permite tomar decisiones a nivel de empleados, quienes interactúan con la información transaccional cotidiana, es decir con los datos que se manejan frecuentemente en el negocio. Dicha información puede ser presentada en hojas de cálculo con formato fijo o herramientas de reportes.

Esto permite que al nivel más bajo los funcionarios puedan acceder a la información necesaria de manera adecuada y exacta, para la visualización de la misma, de tal manera que permita evaluar las condiciones en las que se encuentran y a partir de ahí tomar las decisiones adecuadas.

### **2.1.2. Nivel táctico**

La inteligencia de negocios a nivel táctico permite que la gerencia media o los analistas de datos de la institución utilicen herramientas de análisis y consulta, con el propósito de poder acceder a los datos directamente y visualizar las tendencias de la información, pudiendo personalizar la misma según sus fines.

### **2.1.3. Nivel estratégico**

La inteligencia de negocios a nivel estratégico le permite a la alta dirección monitorear y analizar las tendencias con base en patrones, respecto de sus metas y objetivos estratégicos como organización. En esta parte se utilizan los *dashboards* y *balanced scorecards*, los cuales permiten medir y evaluar el desempeño para la correcta toma de decisiones, las cuales pueden ser medidas y evaluadas para cualquier parte y nivel de la organización.

Partiendo de estas definiciones, para los fines del presente trabajo, se necesita ver las definiciones “negocio”, “empresa” u “organización” como la institución académica, es decir que las decisiones sobre las acciones a tomar y las evaluaciones sobre las mismas no serán tomadas para un producto o servicio comercial, sino más bien para la mejora en el rendimiento académico de los estudiantes de una escuela primaria.

Teniendo esto claro se podrá empezar a determinar de qué manera las diferentes técnicas o herramientas de la inteligencia de negocios pueden dar la ayuda necesaria para la correcta toma de decisiones en los distintos niveles de la institución.

Cada uno de los niveles para la toma de decisiones viene siendo un factor importante para el objetivo de la implementación de la inteligencia de negocios en un ambiente educativo; está claro que se debe de poseer datos para crear información, esto lleva a la pregunta clave del tema ¿Cómo se integran los datos a un sistema para que provean información? Dicha pregunta es el punto de partida para que se puedan tratar las diferentes técnicas y herramientas de la inteligencia de negocios.

## **2.2. Data Warehouse**

Un Data Warehouse (DW) es básicamente el repositorio en el cual se encuentra la integración de los datos consolidados, provenientes de diferentes fuentes, cuyo propósito es el análisis, y a partir de este tomar decisiones centradas en mejorar la gestión del negocio. En dicho repositorio se almacenan los datos estratégicos, tácticos y operativos para los diferentes niveles de toma de decisión previamente expuestos.

Un Data Warehouse debe de ser diseñado a partir de los fines organizacionales, es decir no es un producto que puede ser comprado y listo para utilizarse. El término de DW tiende a confundirse con una base de datos transaccional, cuando en realidad DW se alimenta de los datos necesarios provenientes de una base de datos transaccional; entre sus principales diferencias están:

- Un DW está orientado a una materia, es decir que organiza y orienta sus datos desde la perspectiva de un usuario final.
- Un DW administra grandes cantidades de información debido a que contiene información histórica que por lo general se retira de los sistemas transaccionales, puesto que ya no puede ser necesaria para las aplicaciones que la utilizan en dichos sistemas.
- Un DW puede ser que guarde la información en diversos medios de almacenamiento debido al volumen que dicha información puede poseer.
- Un DW condensa y agrega información, puesto que con frecuencia es muy alto el nivel de detalle de la información que almacena para poder tomar decisiones de una manera sensata.
- Un DW integra y asocia información de muchas fuentes, ya que puede recibir información de diferentes sistemas transaccionales y es necesario recopilar y organizar en un solo lugar toda la información que dichos sistemas han acumulado durante cierto tiempo.

El "Data Warehousing" se aplica básicamente cuando:

- La gestión de información se efectúa sobre múltiples plataformas, aplicaciones o bases de datos transaccionales, cuya operación y utilización se encuentra de manera aislada sin ningún tipo de centralización.
- Los procesos de consolidación, resumen, clasificación y reconciliación son necesarios para el análisis de información global.

- Existen datos organizacionales cuyo registro histórico es de fundamental importancia para tomar decisiones.

Derivado de lo anterior se puede decir que, un repositorio de datos, es una colección centralizada de información corporativa, histórica y transformada, proveniente de sistemas transaccionales externos, para atender requerimientos que apoyen tecnológicamente el proceso de toma de decisiones administrativas y generales.

Teniendo un DW se podrán encontrar diferentes Data Marts (DMt) que vienen siendo un subconjunto del Data Warehouse utilizado normalmente para el análisis parcial de datos; su razón de ser viene a ser la simplificación de la complejidad que podría dar el querer realizar el análisis global de todas las dimensiones que puede tener un DW completo; si este posee demasiadas subdivisiones, cada una con un fin determinado dentro de la organización, un DMt ayuda de esa manera a la subdivisión de la información para la toma concreta de decisiones sobre cierta área de la organización, ayudando también a la rapidez de respuesta al solicitar los datos de la misma.

### **2.3. Data Mining**

Es el descubrimiento del conocimiento o inteligencia oculta en la información almacenada, analiza datos desde diferentes perspectivas, con el objetivo de segmentar los datos en información útil para los fines de la organización. Este proceso ayuda a que sea posible evidenciar las relaciones ocultas entre diferentes tipos de sucesos históricos que son difíciles de palpar con simple estadística dentro de una base de datos transaccional.

DM permite trabajar con grandes cantidades de observaciones y tratar una gran cantidad de variables predictivas; esto último es de gran utilidad para la selección de los patrones y variables útiles dentro de un gran conjunto de información. El proceso de DM ayuda a la toma de decisiones, evalúa los procesos y mejora la calidad de los datos, puesto que extrae lo que realmente es útil para el fin de la organización.

## **2.4. OLAP**

El término OLAP proviene de *Online Analytical Processing* o proceso analítico en línea; es un conjunto de tecnologías diseñadas para mantener específicamente el análisis multidimensional de datos y acceso a los mismos, para permitirle al usuario tener una visión más rápida e interactiva con dichos datos. OLAP viene siendo un sinónimo con vistas multidimensionales de los datos del negocio; dichas vistas proporcionan la base técnica para cálculos y análisis requeridos por las aplicaciones de inteligencia de negocios.

Las aplicaciones OLAP son utilizadas por analistas y gerentes que frecuentemente necesitan visualizar la información directamente y a nivel superior. Estas aplicaciones son actualizadas por lotes, es decir por conjuntos de muchos registros, que generalmente suelen venir de diferentes fuentes.

Muchos de los problemas que se pretenden resolver con la tecnología OLAP son multidimensionales por naturaleza, es decir no existe una relación de uno a uno entre los registros, sino más bien un registro puede estar orientado directamente hacia varias posibilidades, las cuales brindan la posibilidad de analizar tendencias, puesto que una dimensión puede relacionarse con otra y esta a su vez puede estar a su vez relacionada con una nueva.

Dichas dimensiones son creadas estratégicamente para alimentar la posibilidad de acceder a la información dentro de un Data Mart, cuyo fin es específico para un área de la gerencia de la institución interesada.

Es posible utilizar 3 tipos de OLAP, cuya variación radica en la implementación del DW, a saber:

- ROLAP: es un tipo de OLAP que se ejecuta sobre Data Warehouse implementados como tablas relacionales (filas y columnas), utilizando ya sea el modelo de “estrella” o el de “copo de nieve”.
- MOLAP: es un tipo de OLAP que se utiliza sobre Data Warehouse, implementados sobre bases de datos multidimensionales.
- HOLAP: es un enfoque híbrido de ROLAP y MOLAP.

En función de los propósitos para los cuales se diseña el DW y los cubos multidimensionales, se puede realizar operaciones específicas propias de OLAP entre las cuales estan:

- "Drill down" y "Roll up". Entrar en detalle y abstraer.
- "Slice & dice". Rebanar y generar nuevos cubos cambiando naturaleza de dimensiones.
- "Rotate". Cambiar dimensiones unas con otras.



## 2.5. OLTP

Procesamiento de transacciones en línea, por sus siglas en inglés “*OnLine Transaction Processing*”, se refiere a un tipo de sistema que facilita y administra las aplicaciones transaccionales; son utilizados con frecuencia para la entrada, procesamiento y recuperación de transacciones; dichos sistemas de información son enfocados con la arquitectura cliente servidor, puesto que son generalmente pertenecientes a un sistema distribuido.

Los OLTP son considerados bases de datos orientadas específicamente al procesamiento de transacciones; dichas bases de datos entran en este contexto, siendo estas las fuentes de alimentación para un Data Warehouse, dado a que una de las características principales de los sistemas OLTP es que almacena básicamente datos actuales y dichos datos son dinámicos, es decir actualizables; esto brinda los datos que utiliza frecuentemente la organización y los lleva al Data Warehouse para llevar un control general e histórico de la información transaccional.

Con estos conceptos se podría decir que metodológica y genéricamente, el proceso de "Data Warehousing" se compone de:

- Alimentación del Data Warehouse y de sus respectivos Data Marts (réplica, transformación, adecuación, resumen).
- Consultas, requerimientos, análisis multidimensional y minería de datos. Visualización y representación.

## **2.6. Ventajas y beneficios de la inteligencia de negocios en la educación**

En el ambiente educativo en el sector privado se carece de información disponible respecto del alumnado y del personal educativo de la institución; esto promueve la falta de seguimiento hacia los alumnos y la toma de decisiones tardías o erróneas. La inteligencia de negocios promueve la utilización de los datos que se poseen para la interpretación útil hacia el objetivo de la información.

En un ambiente con una buena práctica de inteligencia de negocios se plasma un mejor seguimiento de los acontecimientos ocurridos para la toma de decisiones futuras, basándose en un repositorio histórico de hechos; esto genera un mejor desarrollo en el ambiente operativo, puesto que se pueden tomar decisiones inmediatas tomando en cuenta situaciones similares ocurridas en el pasado, llevando un mejor control del flujo de trabajo institucional.

Con un enfoque específicamente en el ambiente educativo, una solución de inteligencia de negocios puede brindar las siguientes ventajas y beneficios:

- Poseer un conjunto de datos útiles para la información a disposición de la institución, por medio de una base de datos transaccional.
- Tener el control de los datos útiles y promover al catedrático a organizarlos según se requieran.
- Poseer un repositorio de datos históricos para el control de los avances y del rendimiento estudiantil en los diferentes aspectos de relevancia para la correcta enseñanza.

- Proveer datos actuales transformados en información de utilidad para la toma de decisiones operativas inmediatas, respecto de las notas obtenidas en los diferentes bimestres escolares.
- Proporcionar al catedrático información detallada de cada uno de los alumnos y las materias que ellos mismos imparten.
- Proporcionar a la dirección del establecimiento un enfoque personalizado del rendimiento estudiantil, basándose en las notas obtenidas a lo largo de los diferentes ciclos estudiantiles.
- Brindarle al claustro de maestros una perspectiva social del ambiente en que sus alumnos se manejan.
- Lograr asociar datos que por sí solos no brindan beneficios y convertirlos en información de utilidad para la toma de decisiones académicas.
- Brindar estadísticas generales y personalizadas respecto de los factores que afectan el rendimiento educativo.
- Brindar estadísticas generales y personalizadas respecto de las notas obtenidas tanto por los alumnos individualmente como en un grado en específico.

El beneficio más relevante es la capacidad de obtener, almacenar, administrar e interpretar datos, para la toma de decisiones respecto del comportamiento y rendimiento estudiantil.



### **3. MODELO DE DATOS LÓGICO**

Teniendo claros los conceptos acerca de las técnicas y herramientas de inteligencia de negocios que pueden ser útiles para los fines de este trabajo, se procederá a darle a dichas técnicas y herramientas el enfoque hacia su utilización a nivel educativo, visualizando la manera en que estas pueden tener un impacto beneficioso en dicho enfoque y la manera en que pueden ser utilizadas y posiblemente implementadas.

Para una correcta realización de una administración de negocios se debe de entender cuál es el objetivo primordial de la organización; en este caso dicha organización será un centro educativo público de nivel primario, y se definirá el objetivo del mismo como: “Dar seguimiento al rendimiento estudiantil y apoyo a los alumnos”.

Para lograrlo, el sistema propuesto deberá de proveer precisamente un conjunto de herramientas para poder tomar decisiones con base en sucesos y tendencias, las cuales serán proporcionadas por cada una de las técnicas de administración de negocios.

Se empezará recalcando la necesidad de un sistema OLTP para el ingreso de la información transaccional; la transacción en este caso se refiere a cada uno de los datos de interés que suceden comúnmente en el ámbito docente y estudiantil.

Primero se iniciará por definir cuáles son esos posibles datos de interés, según se ha evaluado mediante entrevistas a la psicóloga y directora de la Escuela Oficial Urbana Mixta No. 71 Germán Alcántara, Lic. Ligia Lorena López Liques; lo primero que se debe definir son los datos estudiantiles de importancia, los cuales son:

- Nombre del estudiante
- Código del estudiante
- Dirección
- Fecha de nacimiento del estudiante
- Grado
- Nombre del padre
- Fecha de nacimiento del padre
- Nombre de la madre
- Fecha de nacimiento de la madre
- Divorciados
- Cantidad de hermanos

Como se podrá observar, dicha información es una parte de la ficha de inscripción del estudiante, buscando tener un nivel de control tanto del mismo alumno como del entorno familiar que lo rodea.

Por otra parte, es importante también tener un determinado control del personal docente que labora en la institución educativa, por lo que la propuesta de los datos a manejar es la siguiente:

- Nombre del docente
- Código del docente
- Fecha de nacimiento del docente

- Sexo
- Estado civil
- Dirección
- Cantidad de hijos

Esta información es importante sobre todo porque psicológicamente existe una etapa en la niñez determinada de 6 a 12 años denominada “Fase de relativa serenidad” y abarca de los 8 a los 10 años, en la cual en el ambiente escolar la clase no es un grupo unido, es el profesor el que da cohesión al grupo como “líder” y “jefe”.

Pero los niños se comienzan a unir por la necesidad de participar en un juego en común, aunque estos grupos no son fijos y son relativamente inconscientes de ello.

Este rol de “líder” tiene un impacto en el desempeño académico de una manera directa, puesto que influye en la capacidad de atención y aprendizaje que el estudiante puede tener y es por ello que al administrar y contar con información del personal docente se puede llegar a tomar medidas, si se llegase a determinar que es el docente un factor a mejorar para el rendimiento académico.

Como se puede observar, esta información en realidad no es netamente transaccional puesto que son datos que por regularidad se toman una sola vez en un ciclo escolar, tomando como premisa el establecer el personal tanto docente como estudiantil de un año nuevo de escuela.

Para completar la definición de datos básicos que alimentarán la base de datos transaccional, es necesaria la información de cada una de las materias impartidas y estos datos deben ser:

- Nombre de la materia
- Id de la materia
- Así también los datos del grado impartido:
  - Identificador de grado
  - Descripción del grado

Este tipo de información es la que comúnmente se puede pedir en una base de datos relacional y transaccional y según la licenciada López es lo primordial que un sistema administrador de información debe de manejar para el control de todos los miembros de una escuela primaria.

Tomando esto como premisa se puede evidenciar que cada uno de los elementos propuestos puede fácilmente visualizarse como un conjunto de tablas con relaciones entre ellas, para poder realizar consultas y extraer información de utilidad para el personal, tanto docente como administrativo de la escuela pública.

Esto puede ser así para llevar temas en el nivel operacional de la institución educativa, mas no así en los niveles tanto estratégico como táctico de la organización. Por eso a continuación se procederá a definir el modelado lógico de lo que deberá ser el Data Warehouse y sus Data Marts para los objetivos de la organización.



### **3.1. Modelo de datos**

Para este modelado cabe recalcar que se utilizarán como premisas los requerimientos que posee una escuela pública del nivel primario, como se mencionó en la sección anterior, se poseen datos sobre ciertas entidades de interés, las cuales serán elaboradas sobre el modelo en estrella para los Data Marts que conformarán el Data Warehouse.

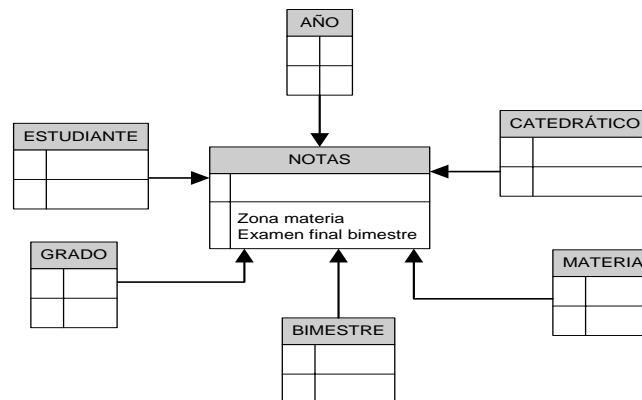
Se seleccionó el modelo en estrella, dado que los datos con los que se dispone pueden ser transformados de tal manera que puedan formar las diferentes dimensiones y se utilizará una tabla de hechos dependiendo del objetivo del Data Mart; dicha tabla de hechos contiene todo lo relevante del negocio; en este caso la institución educativa, en determinada área. Para empezar con el modelado de datos se determinaron cuáles serán los distintos Data Marts que formarán parte de la solución dentro del Data Warehouse:

- Notas de estudiantes
- Asistencia y deserción
- Situación socioeconómica del estudiante
- Salud estudiantil

#### **3.1.1. Modelo estrella para notas de estudiantes**

El siguiente modelo en estrella permite obtener la información sobre las notas de estudiante en cada una de sus materias. Este esquema es la base de la solución propuesta, dado que se presenta la información de las notas recibidas por cada estudiante, en cada materia evaluada y el tiempo en que han sido ingresadas las notas al sistema.

Figura 5. **Modelo dimensional para notas de estudiantes (esquema en estrella)**



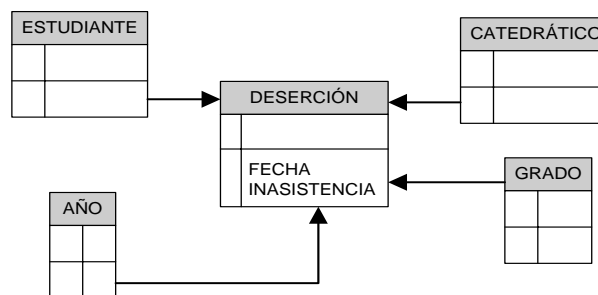
Fuente: elaboración propia.

- Métricas asociadas: zona materia, zona del estudiante al momento en que es ingresada la información y examen final bimestre, esta métrica se refiere a la nota del estudiante en el examen final de bimestre.
- Dimensiones asociadas: estudiante, catedrático, materia, grado y tiempo.
- Cálculos: nota final, la cual corresponde a la sumatoria entre la zona de la materia y el examen final de bimestre.
- Periodicidad de carga: este esquema en estrella debe de ser refrescado al finalizar cada bimestre, es decir cuando las notas finales estén ya tabuladas.
- Granularidad de datos: el nivel de detalle de los datos en este esquema en estrella es el de cada nota obtenida por el estudiante en cada una de sus materias, para cada ciclo de fin de bimestre.

### 3.1.2. Modelo estrella para asistencia y deserción

El siguiente modelo en estrella presenta la información relacionada con la asistencia del estudiante durante el transcurso de la semana, tomando como punto de partida la lista tomada por cada docente diariamente, sobre la inasistencia de sus estudiantes.

Figura 6. **Modelo dimensional para asistencia y deserción (esquema en estrella)**



Fuente: elaboración propia.

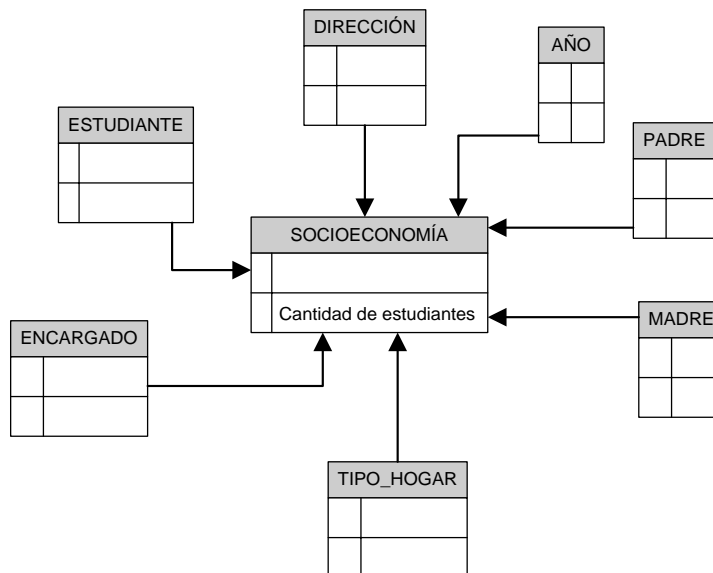
- Métricas asociadas: inasistencia, la cual corresponde a la cantidad de inasistencias del alumno durante la fecha establecida.
- Dimensiones asociadas: estudiante, catedrático, grado y fecha.
- Cálculos: inasistencia, métrica generada a partir de un conteo de los estudiantes que han faltado a clases, según la combinación de las dimensiones utilizadas en la consulta.
- Granularidad de los datos: el nivel de detalle del esquema de información será de cada estudiante al cual se le haya reportado una inasistencia.

- Periodicidad de carga: este esquema en estrella debe ser refrescado semanalmente, para tener los datos actuales de las inasistencias ocurridas durante la semana.

### 3.1.3. Modelo estrella para situación socioeconómica del estudiante

Este esquema en estrella presenta la información socioeconómica de estudiantes inscritos en el ciclo estudiantil, tomando en cuenta la ubicación de su hogar, su encargado y la información correspondiente a los padres.

Figura 7. **Modelo dimensional para situación socioeconómica del estudiante (esquema en estrella)**



Fuente: elaboración propia.

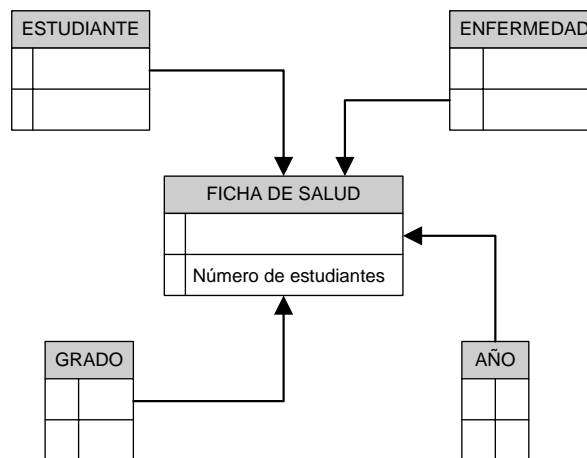
- Dimensiones asociadas: estudiante, encargado, dirección, padre, madre y tipo de hogar.

- Periodicidad de la carga: se deberá de realizar la actualización del sistema de información cada inicio del ciclo escolar, con el fin de saber la situación de cada estudiante inscrito en cada ciclo escolar.
- Granularidad de los datos: el nivel de detalle es el de cada estudiante inscrito en cada ciclo escolar, definiendo el tipo de hogar que integra, los datos de los padres y del encargado.

### 3.1.4. Modelo estrella para salud estudiantil

En este modelo en estrella se presenta la información de cada estudiante respecto de su salud a lo largo del ciclo estudiantil, para determinar las enfermedades a las que está propenso. Este esquema contiene información acerca de las enfermedades que se presentan en el ambiente estudiantil y su asociación con ciertos estudiantes del establecimiento.

Figura 8. **Modelo dimensional para salud estudiantil (esquema en estrella)**



Fuente: elaboración propia.

- Dimensiones asociadas: estudiante, enfermedad, grado y año.
- Cálculos: número de estudiantes, el cual corresponde a la métrica que se genera del conteo de estudiantes que padecen cierta enfermedad según las combinaciones de dimensiones para los filtros de la consulta.
- Periodicidad de carga: este esquema deberá de ser actualizado semanalmente, debido a la posibilidad de asociación de un estudiante con cierta enfermedad durante el transcurso de la semana.
- Granularidad de los datos: el nivel de detalle del esquema de información de cada estudiante que padezca de una enfermedad determinada, mostrando el año y grado en el que la padeció.

### **3.2. Dimensiones del modelo lógico de datos**

Para el modelo de datos lógico se establecen las dimensiones detalladas a continuación.

#### **3.2.1. Dimensión estudiante**

Representa a un estudiante inscrito en la institución educativa; contiene información muy general de cada estudiante con datos tales como: nombre, apellido, código del Ministerio de Educación, fecha de nacimiento y sexo.

##### **3.2.1.1. Atributos asociados**

El código Ministerio de Educación es el número que se le ha asignado a un estudiante al ser inscrito por primera vez.

- Nombres del estudiante
- Apellidos del estudiante
- Fecha de nacimiento del estudiante
- Genero, sexo del estudiante

### **3.2.2. Dimensión catedrático**

Representa al maestro que imparte su cátedra en la institución educativa. Inicialmente se cuenta con los catedráticos que conforman el claustro de maestros de la institución.

#### **3.2.2.1. Atributos asociados**

- El ID personal es un número con el que se le identifica dentro de la institución
- Nombres del catedrático
- Apellidos del catedrático
- Fecha de nacimiento, fecha en que nació el catedrático
- Género, sexo del catedrático
- Estado civil del catedrático
- Número de hijos del catedrático, si no tiene hijos, se asigna 0

### **3.2.3. Dimensión materia**

Representa a cada materia impartida en la escuela, dado que en cada grado se repiten las materias con diferente contenido; se manejará la información de cada una de ellas por separado, es decir un código para la matemática de primero primaria y un código distinto para la matemática de segundo primaria, independientemente que la descripción sea la misma.

### **3.2.3.1. Atributos asociados**

- Identificador de materia y código numérico distintivo
- Descripción, nombre de la materia

### **3.2.4. Dimensión bimestre**

Representa a cada uno de los bimestres dentro del ciclo escolar, que son tomados en cuenta desde enero hasta octubre. Mediante esta dimensión se pueden asociar los eventos ocurridos en determinado bimestre escolar.

Los atributos asociados son:

- Identificador de bimestre, código numérico para hacer distinción de cada bimestre del año escolar.
- Descripción, cadena que en la que se concatenan ambos meses que forman el bimestre con el formato enero-febrero, marzo-abril, etc.

### **3.2.5. Dimensión grado**

Representa cada uno de los grados impartidos en la institución educativa; en este caso equivale a los grados desde primero hasta sexto primaria. Mediante esta dimensión se puede determinar en qué grado sucedió el hecho registrado por cada estudiante.



Los atributos asociados son:

- Identificador de grado y código numérico único
- Descripción, cadena que contiene nombre del grado.

### **3.2.6. Dimensión año**

Representa a cada uno de los años en los que pudo suceder un evento de interés para la institución educativa, ya sea la evaluación de salud, la inscripción de un estudiante o el conjunto de notas de dicho estudiante. Permite diferenciar los hechos ocurridos por un mismo estudiante, incluso si este repite el año académico. El atributo asociado es el año, número entero que representa el año del calendario gregoriano.

Para la carga de dicha dimensión se deberá de crear automáticamente al menos 50 años a partir del año en que se reporten los primeros datos provenientes de la base de datos transaccional; esto con el fin de tener un listado disponible, tomando en cuenta el tiempo desde que se inicia la primera carga de datos reales para no crear inconsistencia en la información.

### **3.2.7. Dimensión dirección**

Representa la ubicación de la vivienda del estudiante; dicha dirección será clasificada en dirección (calle y avenida), zona, colonia, municipio, departamento, para un mejor control de posicionamiento del área en la que el estudiante habita.

Los atributos asociados son los siguientes:

- Identificador de dirección: dato numérico único para la dirección del estudiante.
- Dirección: cadena en la cual se especifica el número de casa, calle y avenida (si aplican).
- Zona: descripción o nombre de la zona en la que se encuentra la vivienda.
- Colonia: nombre de la colonia en la que se encuentra la vivienda (si aplica).
- Municipio: nombre del municipio en el que se encuentra la vivienda.
- Departamento: nombre del departamento donde se encuentra la vivienda.

Al momento de generar la carga de datos, se creará un número automático para el identificador de dirección del estudiante, debido a que la información de la dirección puede variar demasiado y no encontrarse un identificador único y distinguible para la relación su respectiva tabla de hechos.

### **3.2.8. Dimensión padre**

Representa los datos del padre del alumno que están disponibles y son necesarios para la inscripción del mismo dentro del establecimiento.

Mediante esta dimensión se puede tener un mayor conocimiento de los datos referentes al padre del alumno y así obtener información respecto del estilo de vida del mismo.

Los atributos asociados son:

- Identificador padre: código numérico generado para identificar la información del padre del estudiante
- Nombres del padre
- Apellidos del padre
- Estado civil del padre
- Fecha de nacimiento del padre

### **3.2.9. Dimensión madre**

Representa los datos de la madre del alumno que están disponibles y que son necesarios para la inscripción del mismo dentro del establecimiento. Mediante esta dimensión se puede tener un mayor conocimiento de los datos referentes a la madre del alumno y así obtener información respecto del estilo de vida del mismo.

Son atributos asociados los siguientes:

- Identificador madre: código numérico generado para identificar la información de la madre del estudiante
- Nombres de la madre
- Apellidos de la madre
- Estado civil de la madre
- Fecha de nacimiento de la madre

### **3.2.10. Dimensión encargado**

Representa a los datos que se poseen sobre el encargado del estudiante, como encargado se hace referencia al contacto primordial para comunicarse y velar por el rendimiento, salud y seguridad del estudiante. Esta dimensión nos servirá como identificador si el estudiante vive con su familia o no.

Son atributos asociados los siguientes:

- Identificador encargado: código numérico generado automáticamente para identificar al encargado del estudiante
- Nombres del encargado
- Apellidos del encargado
- Fecha de nacimiento del encargado
- Género, sexo del encargado
- Estado civil del encargado

### **3.2.11. Dimensión tipo hogar**

Representa una descripción que servirá de medidor para identificar si el estudiante vive solo con su padre, con su madre, con ambos o con ninguno.

Son atributos asociados los siguientes:

- Identificador tipo hogar, código entero que identificará el tipo de descripción.
- Descripción, la descripción del tipo de hogar en que el estudiante habita.

Para el objetivo de esta dimensión se deberá de especificar concretamente cuál será la descripción correcta para cada uno de los tipos de hogar que el sistema tendrá en cuenta.

### **3.2.12. Dimensión enfermedad**

Representa la descripción de las enfermedades que puede padecer un estudiante durante el ciclo escolar. Esto permitirá identificar cuáles son las enfermedades más comunes en la institución educativa, para asociarlas a los estudiantes que las padecen.

Son atributos asociados los siguientes:

- Identificador enfermedad, código numérico único para cada enfermedad.
- Descripción, nombre de la enfermedad.

Para efectos de no redundar en datos, se utilizarán términos generales para no entrar en los pequeños detalles distintivos en cuanto a la descripción se refiere.



## **4. PROPUESTA DE ARQUITECTURA**

El modelo de arquitectura de un sistema consiste en trazar un mapa o plano, que documenta cómo todos los componentes del sistema trabajan juntos para hacer alcanzar el objetivo final de dicho sistema.

En el presente capítulo se propone una arquitectura fiable y práctica para una solución de inteligencia de negocios, haciendo énfasis en la recolección, interpretación, almacenamiento y presentación de la información de utilidad para la institución educativa.

Asimismo, se toman en cuenta las limitaciones que puede poseer una institución educativa a nivel primario, siendo esta parte del sector público de la educación guatemalteca. Se definen los requerimientos no funcionales que deberán satisfacer la aplicación, así como también se especifica el uso que debe de dársele al sistema según los usuarios que lo utilicen.

### **4.1. Limitaciones**

Debido a que el sistema pretenderá ser implementado sobre los recursos con los que puede contar una escuela pública, las limitantes están muy marcadas en cuanto a costos e infraestructura. La tecnología en los establecimientos públicos no está apoyada por el Ministerio de Educación, por lo que en los establecimientos en los que se cuenta con algún tipo de infraestructura computacional, es porque fueron ayudas o donaciones de terceros.

Partiendo de ese punto, se tiene la premisa de qué recursos en la mayoría de los casos no existen o son escasos, por lo que es necesario realizar una propuesta de arquitectura, cubriendo los requerimientos funcionales y no funcionales necesarios, para que se pueda practicar la inteligencia de negocios, con un límite de recursos. Por consiguiente se pueden definir las limitaciones con la que se contará, para proponer un sistema como:

- Limitaciones en costos: dado que no existe apoyo gubernamental y en la educación pública no se dispone de un cobro hacia los estudiantes para la compra de equipo.
- Limitaciones en infraestructura: debido a que no necesariamente se cuenta con el equipo necesario y el apoyo tecnológico, no existe en las instituciones educativas públicas la capacidad de obtención de equipo para realizar una infraestructura adecuada; es la limitación más marcada para esta solución.

#### **4.2. Requerimientos no funcionales**

Para dar inicio con la propuesta de la arquitectura es necesario identificar los requerimientos no funcionales, con los que debe cumplirse para que el planteamiento de sistema cumpla su objetivo.

Los requerimientos no funcionales que se identificaron son:

- Rendimiento: el sistema tiene que proveer un rendimiento óptimo para que la transformación y transferencia de información sea eficiente.



- Usabilidad: el usuario final debe de sentirse familiarizado con el sistema para utilizarlo con mayor fluidez, tomando en cuenta que la mayoría de usuarios finales en esta solución carecen de conocimiento tecnológico y computacional.
- Costo: el sistema no puede tener un costo excesivo debido a que no se cuenta con el recurso económico para la implementación con un coste excesivo.
- Operatividad: el sistema debe de operar correctamente en el procesamiento de la información para que la misma sea consistente en el tiempo.

#### **4.3. Límites**

Para establecer la propuesta de la arquitectura se necesitará definir los límites del sistema, entre los cuales se identifican los siguientes:

- El sistema será utilizado únicamente dentro del establecimiento educativo, es decir las bases de datos transaccionales serán alimentadas desde un mismo servidor, debido tanto a la falta de infraestructura como al poco acceso del personal a utilidades tecnológicas.
- La información transaccional será ingresada con base en un programa local, el cual mediante pantallas de ingreso de datos permitirá a los catedráticos y personal administrativo alimentar de información al sistema, dando un estándar para el ingreso de dichos datos.

- El sistema será administrado con base en roles de usuario, es decir un usuario no necesariamente tendrá acceso a todo el sistema de información y a la información que este brinda como salida.
- No se cuenta con grandes servidores de datos, por lo que se utilizarán computadoras sencillas para el manejo y administración de los mismos.
- Las instituciones públicas educativas a nivel primario no cuentan con una base de datos transaccional, por lo que la implementación de la misma debe de iniciar siendo únicamente para el ingreso de datos y no para consultas pequeñas o transacciones que emitan algún tipo de cálculo complejo. Es decir que la base de datos transaccional que se implementará, únicamente tendrá como objetivo alimentar al Data Warehouse y sus respectivos Data Marts.
- El nivel de toma de decisiones estratégicas será utilizado por la mayor autoridad de la institución sin dar algún tipo de registro al sistema del Ministerio de Educación directamente.

#### **4.4. Casos de uso de alto nivel**

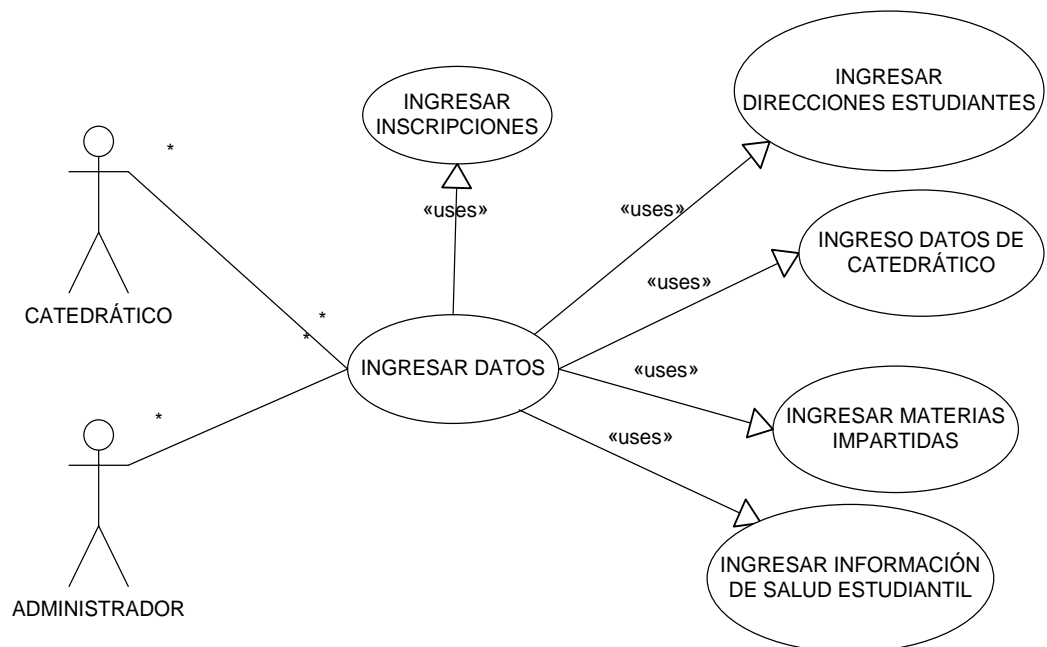
A continuación se detallan los casos de uso de alto nivel para el planteamiento de las necesidades actuales.

##### **4.4.1. Casos de uso para el ingreso de información**

En el siguiente diagrama se identifican los casos de uso necesarios para especificar lo que será el ingreso de datos al sistema transaccional; este es el punto de partida para la ejecución del sistema de inteligencia de negocios.

Este diagrama se basa en las operaciones realizadas sobre una aplicación de ingreso de datos, la cual será la responsable de alimentar la base de datos transaccional de la institución.

Figura 9. **Casos de uso ingreso de información**



Fuente: elaboración propia.

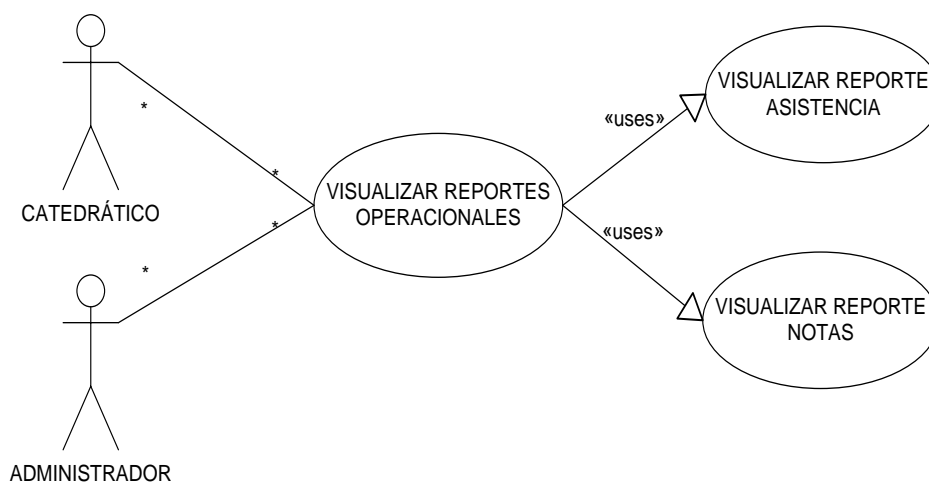
- **Actores**
  - **Catedrático:** persona que labora en la institución como catedrático y encargada de ingresar el listado de sus alumnos inscritos, así como las diferentes categorías expuestas en el diagrama.
  - **Administrador:** persona con privilegios a nivel de base de datos encargada de la administración de los datos y la corrección de errores en los ingresos.

- Especificaciones: el ingreso de información transaccional comprende el ingreso de datos obtenidos, ya sea en variaciones en el avance del ciclo educativo, las notas obtenidas, deserciones de los estudiantes e información fundamental para la organización, tal como el ingreso de datos del personal docente o datos respecto de los alumnos inscritos en la institución.

#### 4.4.2. Casos de uso para la visualización de reportes operacionales

En el siguiente diagrama se especifican los actores que interactúan con el nivel operacional de los reportes propuestos; se toman en cuenta dos reportes necesarios, debido a que con base en ellos pueden tomarse decisiones a nivel operacional, ya se refiera a notas de estudiantes, o a su inasistencia o deserción.

Figura 10. Casos de uso visualización de reportes operacionales



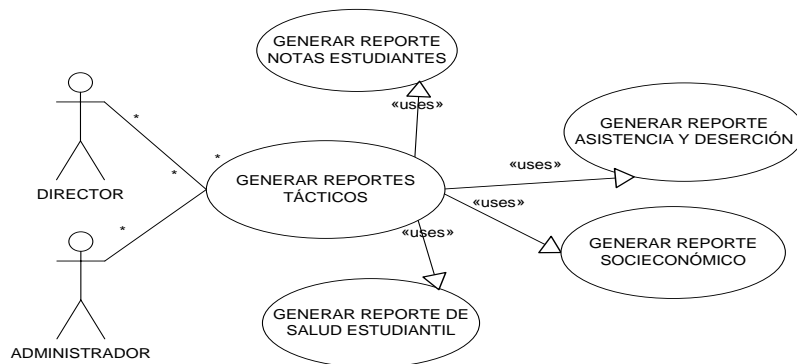
Fuente: elaboración propia.

- Actores
  - Catedrático: persona encargada de impartir clases en la institución educativa; es el que tiene contacto frecuente con las funciones operacionales de la organización, que en este caso se refieren a brindar una educación de calidad a los estudiantes.
  - Administrador: persona encargada de administrar el área de reportes, verificando que cada uno cumpla con el cometido de mostrar claramente los datos y evitar o resolver errores.

#### 4.4.3. Casos de uso para la generación de reportes tácticos

En el siguiente diagrama se especifica la interacción entre los actores y el área de reportes tácticos, los cuales son visualizados a nivel de gerencia de área; en este caso, a nivel de dirección escolar se identifican los reportes que pueden ser generados en este nivel a través de los distintos Data Marts existentes en el Data Warehouse.

Figura 11. Casos de uso para la generación de reportes tácticos



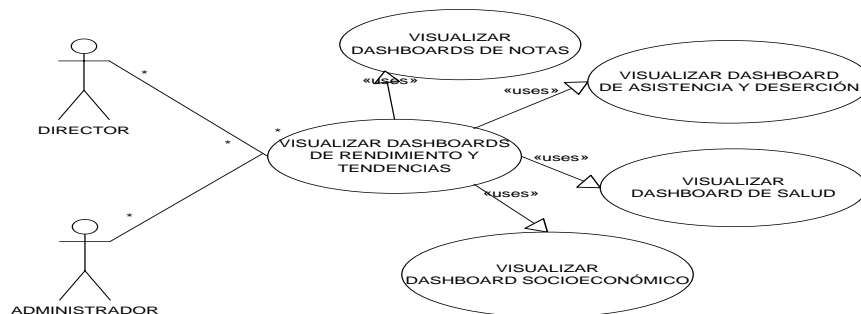
Fuente: elaboración propia.

- Actores
  - Director: persona encargada de dirigir la institución estudiantil, velando por que se cumplan los programas educativos y coordinando al claustro de maestros; este actor es el encargado de tomar decisiones específicas desde una perspectiva global y es el encargado de visualizar la información de los distintos Data Marts.
  - Administrador: encargado de velar porque la información en los Data Marts sea consistente de acuerdo con lo que se necesita extraer de los mismos.

#### 4.4.4. Casos de uso para el monitoreo estratégico

En el siguiente diagrama se muestran los actores y su interacción con los diferentes *dashboards* o tableros de mando, en los cuales se muestran de manera gráfica y estadística las tendencias en rendimiento institucional, según los fines de la organización.

Figura 12. Casos de uso para el monitoreo estratégico



Fuente: elaboración propia.

- Actores
  - Director: persona encargada de tomar las decisiones a nivel global en la institución, y responsable de que la misma cumpla con los objetivos que en este caso es brindar una educación integral a los estudiantes.
  - Administrador: persona encargada de administrar el sistema de *dashboards* y velar por su correcto funcionamiento e integridad en los datos.

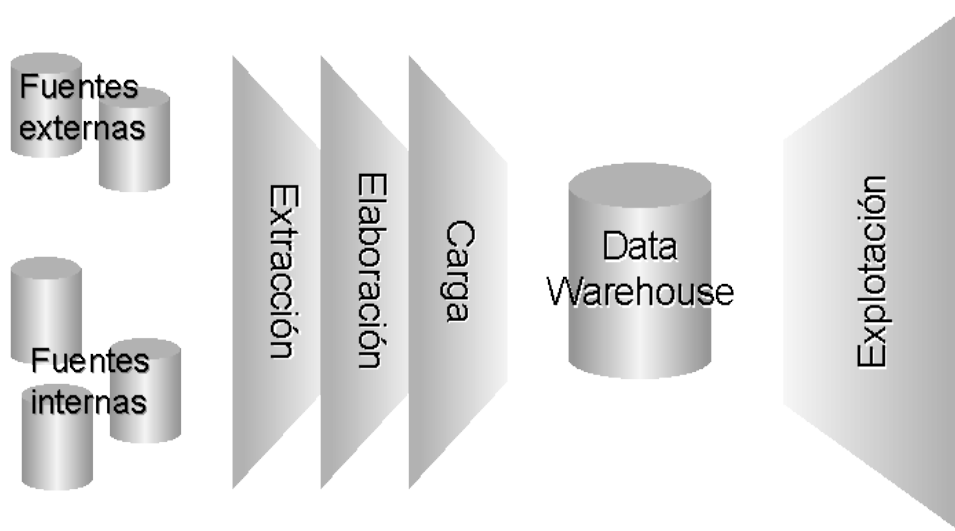
#### **4.5. La estructura básica de la arquitectura Data Warehouse**

La estructura de la arquitectura de un Data Warehouse es el punto de partida para definir una arquitectura propia para la solución; dicha estructura de la arquitectura incluye de manera básica:

- Datos operacionales: es necesario contar con un origen de datos que permitan poblar el repositorio del Data Warehouse; dichos datos provienen tanto de las bases de datos transaccionales de la institución como de fuentes de datos externas a la misma.
- Extracción de datos: comprende la selección de los datos que se poseen en las diferentes fuentes para poblar la estructura del Data Warehouse.
- Transformación de datos: es el proceso por el cual se realizan cambios operativos en los datos extraídos para cumplir con los objetivos de la organización.

- Carga de datos: se refiere básicamente a la inserción estratégica de los datos extraídos hacia el almacenamiento del Data Warehouse.
- Data Warehouse: área de almacenamiento físico de la arquitectura Data Warehouse.
- Herramientas de acceso al área física del Data Warehouse: estas son las que proveen a los usuarios finales el acceso a los datos; dichas herramientas pueden ser comerciales, ya que incidirán en una visualización de las bases multidimensionales o Data Marts almacenadas en el Data Warehouse; así también pueden ser herramientas que provean un soporte para las decisiones, extracción de información para *dashboards* de rendimiento u operaciones básicas sobre la información en el Data Warehouse.

Figura 13. **Estructura básica de Data Warehouse**



Fuente: <http://www.dataprix.com/que-es-un-datawarehouse>.

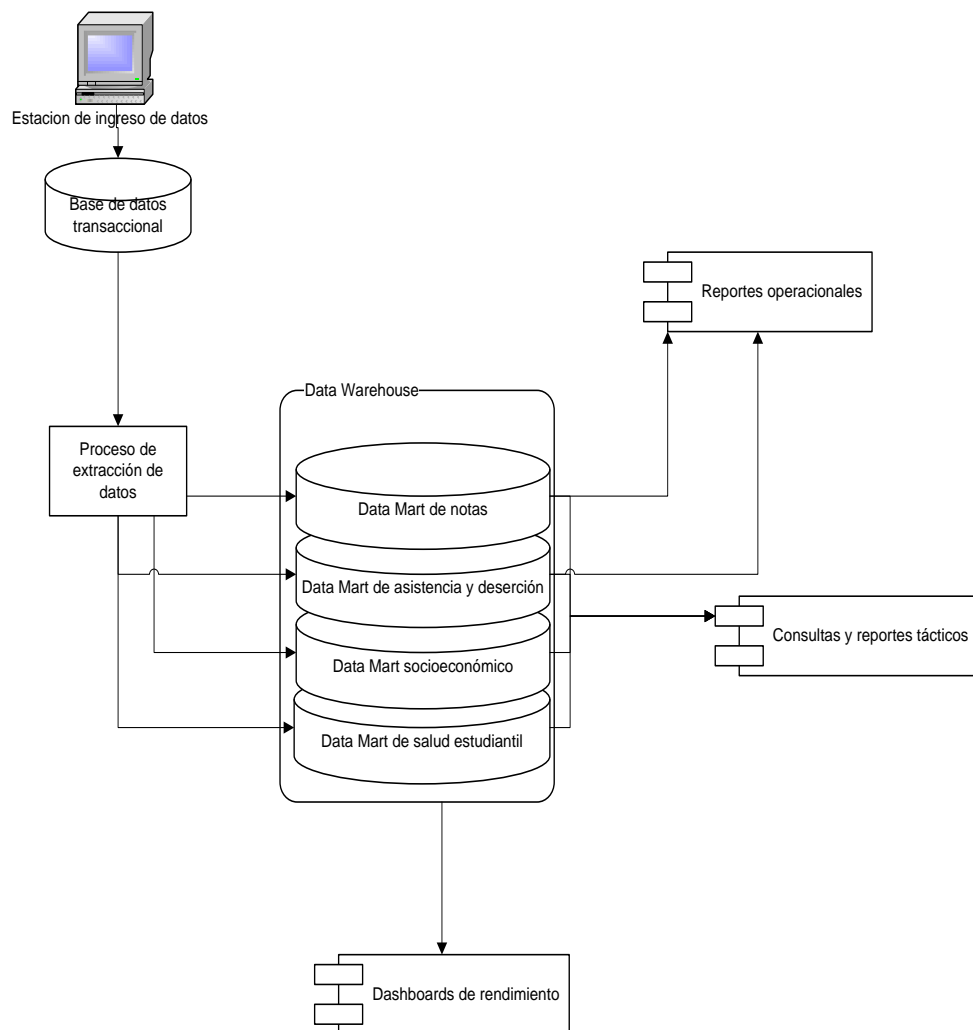
Consulta: 30 de junio de 2012.



#### 4.6. Arquitectura propuesta

Con base en lo expuesto, en los requerimientos no funcionales a las limitaciones y suposiciones, así como al evaluar los casos de uso, se puede determinar la arquitectura más acorde a las capacidades y objetivos de la institución.

Figura 14. **Arquitectura propuesta**



Fuente: elaboración propia.

#### **4.6.1. Justificación de la propuesta**

La arquitectura propuesta es una modificación de la estructura de la arquitectura básica de un Data Warehouse, en la cual se tomaron en cuenta las siguientes consideraciones:

- La escasez de recursos de infraestructura en las escuelas públicas pone como condición el uso mínimo de dichos recursos, por lo que se plantea una máquina local en la cual se ingresarán toda los datos disponibles sobre una base de datos transaccional ubicada en la misma computadora.
- Partiendo del punto anterior se estableció que los datos externos o de múltiples fuentes quedan fuera del modelo de la arquitectura, debido precisamente a la escasez de recursos y a la falta de un sistema transaccional antiguo o previamente establecido en las escuelas públicas.
- El proceso de extracción y carga de datos se unificó, eliminando así la transformación de los datos; debido a que en las instituciones públicas no se cuenta con una base de datos dedicada al proceso de ingreso de información académica con continuidad, se puede definir una estructura en la cual los datos sean ingresados de tal manera que lleven un formato específico para el repositorio de datos (DW) y de esta manera no será necesaria dicha transformación.

- El Data Warehouse está compuesto únicamente por los Data Marts modelados anteriormente, con un modelo estrella, haciendo así el tiempo de respuesta mucho más rápido y menos robusto para el ordenador, que como se mencionó anteriormente, será de recursos limitados.
- La información se extrae directamente del repositorio de datos hacia los reportes, informes y estadísticas finales que serán utilizados con el objetivo de la toma de decisiones académicas y control del rendimiento estudiantil.

#### **4.6.2. Especificación de la propuesta**

En el modelo propuesto para la arquitectura de software se especifican 3 áreas importantes y bien marcadas. Debido a lo poco variable que es una arquitectura básica para la elaboración de un Data Warehouse, es necesario especificar cada una de las áreas que influirán en este para sus diferentes propósitos.

En relación con el área operacional, primeramente se denota el área operacional de la aplicación; la misma está planteada para que sea la fuente de alimentación al sistema de información, es decir que en ella se realicen la carga e ingreso de datos útiles para la institución educativa.

Se plantea poseer un equipo de cómputo con las características básicas para poseer una base de datos instalada, la cual será la base de datos transaccional de la solución. Debido a los escasos de recursos, se plantea que en el mismo equipo en el cual se encuentre la base de datos transaccional encontrar la aplicación que alimentará dicha base de datos.

Esto brinda una mayor capacidad de utilización de recursos físicos así como la facilidad de implementación de la aplicación encargada de ingresar los datos al sistema.

Este equipo tiene la función de carga de datos iniciales, que luego serán extraídos para alimentar el repositorio de datos históricos.

En el área de repositorio se encuentra físicamente el repositorio de datos; este consta de un equipo de cómputo más potente que el anterior debido al volumen de información que almacenará, comparando el equipo con la base de datos transaccional.

Este equipo será el encargado de resguardar la información extraída de la base de datos transaccional y administrarla en cada uno de los diferentes Data Marts que se especificaron en el capítulo anterior.

En esta área se encontrará el corazón de la solución propuesta, puesto que será la encargada de interpretar los datos extraídos y administrarlos dentro del Data Warehouse; el equipo dedicado al almacenamiento de estos datos históricos deberá ser el encargado de utilizar las herramientas de software para la administración de los cubos OLAP que son representados en cada uno de los Data Marts establecidos.

El área de reportes es contemplada con la posibilidad de mostrar la información solicitada por el usuario, ya sea en un tercer equipo o en el equipo local que contiene el repositorio de datos.

Esta área será la encargada de mostrar los resultados de la manipulación de los datos convertidos en información de utilidad; esta parte debe de ser capaz de mostrar los informes en cada una de las áreas organizacionales.

Para el correcto funcionamiento de esta área es necesario contemplar los requerimientos mínimos de hardware para verificar que la capacidad de visualización de los datos es clara para la toma de decisiones institucionales.

En esta área es en donde se plasman las ventajas y beneficios de la implementación de la inteligencia de negocios, puesto que en ella se puede observar la información personalizada y objetiva del tema que va a evaluar la institución educativa.

#### **4.7. Requisitos de hardware**

Con base en la arquitectura propuesta anteriormente, se debe de establecer cuáles deben de ser los requisitos básicos para que dicha arquitectura se cumpla.

Hay que tener en cuenta cuando se habla de requisitos de hardware en un Data Warehouse el nivel de concurrencia en transacciones y la cantidad de datos históricos que se almacenarán, dependiendo del planteamiento lógico de los datos a manejar.

Teniendo esto como premisa se puede establecer un conjunto de necesidades básicas que debe de cumplir el equipo necesario para la implementación de un sistema de Data Warehouse para una institución educativa del sector público.

Dado que el sistema planteado no es un sistema con operaciones transaccionales diarias y concurrentes, se puede establecer que los recursos necesarios para una implementación de Data Warehouse en un establecimiento educativo público a nivel primario son los que a continuación se describen.

#### **4.7.1. Equipo para base de datos transaccional**

Se necesita un equipo dedicado al ingreso de los datos transaccionales, que en este caso son las notas e información del personal tanto administrativo como de los estudiantes miembros del establecimiento. En este equipo debe de realizarse el ingreso de la información periódica, tanto a final de semestre como a final de año, o cada vez que se presente un cambio en la impartición de clases o ingreso de nuevos miembros en el establecimiento.

Partiendo de aquí se puede determinar que el equipo necesario debe de cumplir con las características de rendimiento para la utilización de una aplicación desarrollada en un lenguaje de programación, que servirá de entorno gráfico para el ingreso de la información; también debe de cumplir con las características necesarias para la utilización y administración de una base de datos sencilla y ligera.

Con los requisitos de hardware establecidos en la siguiente tabla se busca la capacidad de realizar transacciones sobre una base de datos relacional por medio de una conexión desde una aplicación de escritorio desarrollada sobre un lenguaje de programación; estos requisitos permiten soportar aplicaciones Java, Visual Basic o C#, para brindar la facilidad de la creación de dichas aplicaciones sobre un lenguaje estable y conocido.

Tabla I. **Especificaciones de hardware**

<b>Requisito</b>	<b>Mínimo</b>
Memoria física (RAM)	1 Gb
Memoria virtual	El doble del tamaño de la memoria RAM (2 Gb)
Disco duro	8 Gb
Procesador	550 Mhz
Tarjeta de red	1

Fuente: elaboración propia.

Se toman en cuenta estos requerimientos, ya que proporcionan la capacidad de poseer un sistema operativo Linux en cualquiera de sus versiones o un sistema operativo Windows XP.

Estos requisitos de hardware también brindan la posibilidad de la utilización de sistemas de administración de base de datos sencillos como Access, hasta un sistema más completo como Oracle 10 g.

Se propuso un disco duro de 8 Gb puesto que en él se pueden tener instalados correctamente un sistema operativo y una base de datos relacional y dado que la información transaccional que manejará no es de gran tamaño, durante el año escolar puede almacenar la información operaria del año en curso sin ningún problema. Este requisito se evaluó con base en la cantidad y periodicidad de las transacciones.

#### 4.7.2. Equipo para el repositorio de datos

Partiendo del objetivo de poseer un repositorio de datos históricos, se puede determinar que es necesario un equipo con más capacidad que el utilizado para la base de datos transaccional; teniendo en cuenta las limitaciones económicas y la capacidad de conseguir equipo de calidad, se proponen los requisitos mínimos que debe de cumplir el equipo para la correcta administración, almacenamiento y presentación de la información histórica que debe contener el repositorio.

Este equipo cumplirá con la función de almacenamiento, recolección, carga, administración y presentación de la información que es de utilidad para la institución educativa. En él se encontrarán almacenados los datos históricos que la organización maneja a lo largo de todos los ciclos escolares. Los requerimientos mínimos de hardware propuestos tras la evaluación del objetivo del sistema propuesto y la capacidad de obtención del equipo por parte de una institución educativa serán expuestos a continuación.

Tabla II. **Especificaciones de hardware**

<b>Requisito</b>	<b>Mínimo</b>
Memoria física (RAM)	2 Gb
Memoria virtual	El doble del tamaño de la memoria RAM (4 Gb)
Disco duro	32 Gb
Procesador	800 Mhz
Tarjeta de red	2

Fuente: elaboración propia.



Como fue expuesto anteriormente, debido a que el repositorio de datos contiene una cantidad relativamente mayor que la base de datos transaccional, debe poseer un disco duro mayor para la capacidad de tener un sistema operativo, ya sea Linux en cualquiera de sus versiones o un Windows XP para una correcta administración del sistema.

Se tomó en cuenta que aunque la información es histórica, esta representa transacciones que no poseen una concurrencia diaria, lo cual permite determinar que aunque la información tiene un volumen mayor al de la base de datos transaccional, no abarca un espacio demasiado grande en disco; esto va de la mano con la capacidad del software del repositorio para el almacenamiento de los datos.

Se requiere una mayor capacidad de memoria RAM puesto que se realizarán operaciones de carga y administración más complejas que en una base de datos transaccional y esto permite también la utilización de un software ya sea libre o comercial, con una buena capacidad de almacenamiento y presentación de los datos. Se especifican 2 tarjetas de red tomando en cuenta la posibilidad de no tener la necesidad de presentar los datos localmente, es decir presentar los distintos informes y *dashboards* en un tercer equipo de cómputo y no solo la necesidad implícita de la conectividad con la base de datos transaccional.

#### **4.7.3. Equipo de red**

Para establecer la conectividad entre los distintos dispositivos necesarios para la correcta implementación de la arquitectura propuesta, se plantea establecer una red cableada, la cual fue seleccionada por su facilidad de implementación y bajo coste.

La red propuesta deberá de ser capaz de conectar como mínimo el equipo donde se aloja la base de datos transaccional y el equipo que contiene el repositorio de datos.

Para ello se evaluó la posibilidad de implementar una red con cables cruzados y un *switch* de 8 puertos que permita contemplar la posibilidad de expandir la presentación de los diferentes reportes a más computadoras, o poseer en un futuro una base de datos transaccional extra para alimentar el repositorio de datos según como el sistema pueda ir evolucionando u obteniendo recursos para una mayor facilidad en el ingreso y presentación de la información.

#### **4.7.4. Ambiente y ubicación del equipo**

Es bien conocido que para el buen mantenimiento del equipo de cómputo especialmente si este es utilizado como servidor de datos, es necesario contemplar el factor temperatura para el no sobrecalentamiento de los equipos; debido a la escasez de recursos se recomienda cumplir con las siguientes especificaciones:

- La solución propuesta está planteada para poseer los equipos que contienen tanto a la base de datos transaccional como el que posee el repositorio de datos bajo una misma ubicación, es decir dentro de un mismo sitio dentro del establecimiento educativo.
- Mantener los equipos en un sitio en el que no existan filtraciones de agua ni tuberías.

- Mantener el equipo aislado de posibles amenazas como golpes, humedad o polvo.
- Brindarle limpieza periódica al hardware que se posee.
- Utilizar el hardware estrictamente para el ingreso y análisis de la información.
- Establecer reglas de uso y tomar en cuenta que el equipo utilizado para la implementación de la solución no tiene como objetivo la enseñanza de computación hacia el alumnado.



## 5. DEFINICIÓN DE MÓDULOS

Como parte de la solución de inteligencia de negocios, y siendo el objetivo de la misma, es necesario proceder a identificar y definir los diferentes medios por los cuales se logrará obtener la información necesaria proveniente del Data Warehouse definido anteriormente, para la correcta toma de decisiones.

Para la solución a este sistema y con base en la teoría de inteligencia de negocios, se han planteado los 3 niveles de toma de decisión y las diferentes herramientas para la visualización y obtención de información; dichos niveles como ha sido descrito con anterioridad serán:

- Nivel operativo
- Nivel táctico
- Nivel estratégico

Como solución al problema planteado se distingue cada una de las áreas según la información a la cual se necesita acceder y la presentación de la misma para su correcta interpretación.

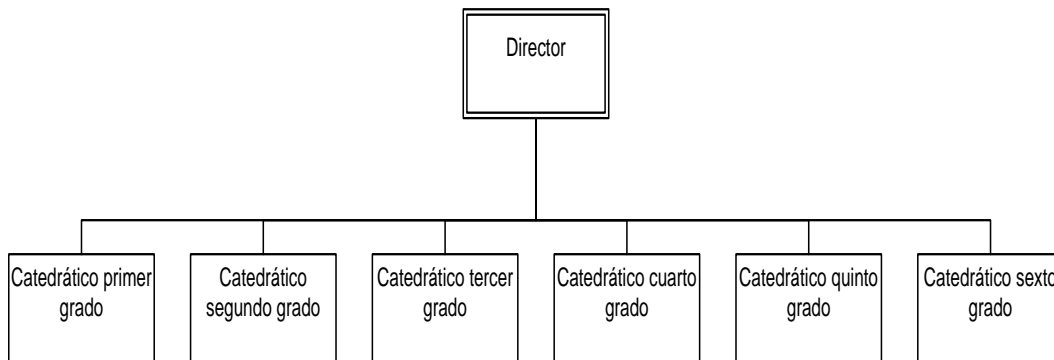
Dichas áreas serán especificadas a continuación, determinando su función, utilidad, visualización de información y roles para el acceso a dicha visualización de información y de esa manera hacer uso de ella.

## 5.1. Organigrama institucional

Para empezar con la definición de los módulos y los roles que tendrán acceso a dichos módulos, es necesario conocer el esquema institucional para definir qué usuarios tendrán los roles y la utilización de los módulos para la toma de decisiones, según su posición dentro de la organización.

La educación pública, en especial a nivel primario, tiene una estructura organizacional básica a diferencia de la educación privada, puesto que no se cuenta con coordinadores de área específicos para determinados grados; la organización institucional radica prácticamente de la siguiente manera:

Figura 15. **Organigrama institucional**



Fuente: elaboración propia.

Este organigrama puede variar según la cantidad de maestros que impartan cierto grado, dependiendo de la escuela y la cantidad de secciones abiertas. Como se puede observar, la máxima autoridad o por así decirlo la alta gerencia del negocio es el director de la institución y el área de operadores es dada por el claustro de maestros.

Partiendo de aquí se puede empezar a distinguir con claridad los roles que cada usuario del sistema tendrá para la visualización de la información y las diferentes maneras de toma de decisiones organizacionales que se definieron anteriormente.

## **5.2. Módulo operativo**

Este módulo está planteado de tal manera que sea utilizado específicamente por los usuarios que interactúan constantemente con los objetivos operativos del negocio; en este caso, dichos objetivos operativos se centran en brindar educación a los estudiantes, es decir que será utilizado directamente por el personal docente para el control de sus estudiantes.

El rol docente será asignado a toda aquella persona que esté involucrada con la impartición de cátedra al alumnado de la institución educativa y a todo aquel usuario que brinde apoyo en la docencia.

En este módulo se plantearon dos tipos de reportes a visualizar; dichos reportes cumplen con las siguientes características:

- Formato de tabla: los reportes generados a nivel operativo son mostrados como tablas, desplegando por filas los datos devueltos por las consultas, y por columnas, la información especificada.
- Formato fijo: esto quiere decir que los datos a desplegar están estandarizados para que, dependiendo los filtros de búsqueda, se generen mostrando siempre las mismas columnas.

- Capacidad de exportar a hoja de cálculo: se debe de tomar muy en cuenta que los reportes por sí solos únicamente muestran un conjunto de datos extraídos con base en una consulta específica a través de un conjunto de filtros, pero para una correcta apreciación e indagación en la información, se requiere que sean capaces de ser exportados a una hoja de cálculo para poder realizar operaciones sobre las filas devueltas por la consulta.

Para este módulo se planificaron dos tipos de reportes, los cuales se explicarán a continuación.

### 5.2.1. Reporte de notas por alumno

En este reporte se pretende mostrar en una tabla los datos básicos de un alumno como: género, código y nombre, así también la zona obtenida y la nota del examen final correspondiente a una materia determinada; especificando también el bimestre y el año en que dicha nota fue obtenida. Este reporte pretende a ayudar al funcionario operativo de la institución a evaluar las notas obtenidas por sus estudiantes y así verificar si está cumpliendo sus objetivos como catedrático.

Tabla III. Formato de reporte de notas por alumno

Código	Nombres	Apellidos	Género	Materia	Bimestre	Año	Zona	Examen final	Nota final
001	Juan Carlos	Pérez López	Masculino	Música	Marzo- abril	2012	50	20	70
005	María	Navas Juárez	Femenino	Idioma Español	Marzo- abril	2012	50	25	75

Fuente: elaboración propia.



Los filtros a especificar serán los siguientes:

- Código
- Género
- Materia
- Bimestre
- Año

### 5.2.2. Reporte de asistencia

En este reporte se pretende llevar el control de la asistencia e inasistencia de los alumnos en cierto periodo de tiempo; esto brindará al funcionario operativo un control sobre los alumnos que se presentan a clase e identificará quiénes faltan demasiado con tendencia a desertar.

Tabla IV. **Formato de reporte de asistencia**

Código	Nombres	Apellidos	Genero	Materia	Bimestre	Inasistencia
001	Juan Carlos	Pérez López	Masculino	Música	Marzo-abril	0
005	María	Navas Juárez	Femenino	Idioma español	Marzo-abril	4

Fuente: elaboración propia.

Los filtros a especificar serán los siguientes:

- Código
- Género
- Materia

- Bimestre
- Año

### **5.3. Módulo táctico**

En este módulo se pretende desglosar la información “inteligente” para el análisis de las diferentes áreas que afectan al desarrollo y rendimiento estudiantil. Este módulo está diseñado con el propósito de que los usuarios con el rol de analista puedan indagar en la información directamente, “jugando” con las diferentes combinaciones que esta presenta, para encontrar patrones y estadísticas de utilidad para la toma de decisiones con base en lo que la información vaya marcando históricamente.

El rol del analista rol será otorgado a los funcionarios con la capacidad de tomar decisiones sobre cierta área educacional, es decir son todos aquellos usuarios cuyo cargo tiene la potestad de tomar decisiones más generales respecto de las situaciones que se dan dentro de la institución. En el caso de la escuela primaria se consideró como usuario principal al director del establecimiento y tomando en cuenta también si en la organización de la escuela existen jefes de áreas debido a que no en todas se da este rango. Los jefes de áreas son los encargados de cierta cantidad de grados sin ser catedráticos en cada uno de ellos.

Para este módulo se crearon específicamente cada uno de los Data Marts modelados en el capítulo 3 del presente trabajo, los cuales tienen como función brindar información histórica y combinable, según lo que se desee evaluar.

En este módulo los reportes se generan de manera personalizada, es decir los usuarios del sistema con el rol analista pueden visualizar los diferentes Data Marts y hacer combinaciones con cada una de las dimensiones especificadas. En este módulo los reportes se visualizan como una tabla dinámica en la cual se pueden especificar los campos con cada una de las condiciones que se quieren, ya sea evaluar, totalizar o subdividir.

### **5.3.1. Reporte de notas de estudiantes**

Este reporte hace uso del Data Mart de notas, el cual tiene como propósito brindar la información de cada uno de los estudiantes y sus notas, teniendo como dimensiones el año, grado y catedrático, mostrando como métricas la zona obtenida y la nota del examen final; esto brindará al análisis del rendimiento académico de cada uno de los estudiantes, y buscará las combinaciones que puedan dar una pauta de cómo se están desarrollando en el ámbito académico.

### **5.3.2. Reporte de asistencia y deserción**

Este reporte pretende llevar el control de las inasistencias a clases por parte de los alumnos, tomando como posibles combinaciones el grado, bimestre, año y profesor que imparte las clases; con esto se da una gran ayuda a la administración de la educación, puesto que se puede determinar qué estudiantes tienden a la deserción y buscar alguna razón externa que pueda estar afectando el hecho de asistir a clases.

### **5.3.3. Reporte socioeconómico**

Este reporte tiene como fin brindar información de relevancia respecto del ambiente en que los estudiantes permanecen, tomando en cuenta la ubicación de su hogar y la situación del mismo debido a posibles problemas familiares, falta de atención, madres o padres solteros, llevan así un control extraestudiantil del alumno para ver la posibilidad de una ayuda psicológica para que su rendimiento académico no disminuya por este tipo de factores que en la educación nacional son la causa de la deserción en la educación, vandalismo y delincuencia.

### **5.3.4. Reporte salud estudiantil**

Este reporte está enfocado a la posibilidad de hacer un análisis más profundo de otra de las causas que afectan negativamente al rendimiento estudiantil de los estudiantes en las escuelas primarias, puesto que sí es bien conocido que se tienen los recursos para brindarle ayuda médica a jóvenes estudiantes, porque no es el objetivo de la organización ni está contemplado en el presupuesto brindado por el gobierno; se puede hacer el análisis de los estudiantes, cuyo rendimiento realmente es afectado por su estado de salud, y de esta manera buscar una solución a ese problema ya sea informando a los padres o encargados de la situación y poder prevenir algún problema futuro.

## **5.4. Módulo estratégico**

Este módulo está pensado específicamente para brindar a la alta dirección de la organización el poder monitorear y analizar las tendencias con base en los patrones que la información inteligente brinda.

En esta parte se plantea la utilización de *dashboards* o tableros de mando, los cuales permiten medir y evaluar el desempeño estudiantil para la correcta toma de decisiones globales en la institución.

El rol del director para este módulo plantea la utilización de las herramientas de la toma de decisiones estratégicas únicamente para la alta gerencia, es decir al director estudiantil, puesto que en él recae la responsabilidad de la toma de decisiones drásticas que pueden afectar a la organización entera.

Este módulo pretende extraer información de cada Data Mart para la visualización de las gráficas estadísticas que proporcionan los *dashboards* y cuyo propósito es brindar la información concisa y apreciable para una toma de decisiones más efectiva y rápida.

Con base en el diseño de cada uno de los Data Marts se especificaron los diferentes tipos de *dashboards* necesarios para este tipo de toma de decisiones.

#### **5.4.1. *Dashboard* de notas**

En esta presentación de información se deben mostrar los datos globales y por grado de las notas obtenidas en cada bimestre y en cada materia. Esto ayudará a facilitar la identificación de problemas con los contenidos o metodologías de enseñanza aplicadas, si el alumnado falla de manera significativa en cierta materia.

#### **5.4.2. *Dashboard* de deserción**

En esta presentación de la información se deben mostrar los datos globales de la institución por año y grado, en los cuales existe una significativa inasistencia para tomar decisiones respecto de brindar algún tipo de incentivo material o psicológico para que los alumnos permanezcan en la escuela y sigan su formación académica durante el año escolar.

#### **5.4.3. *Dashboard* socioeconómico**

En este *dashboard* es necesario presentar las gráficas estadísticas de las tendencias que se tienen dentro del alumnado respecto de su situación familiar; de esta manera se puede detectar en qué grado existe una cantidad considerable de problemas intrafamiliares que pueden repercutir en el ambiente estudiantil puesto que si es un gran número de alumnos, el rendimiento baja a nivel global.

#### **5.4.4. *Dashboard* de salud estudiantil**

En este *dashboard* es necesario representar la información, de tal manera que se denoten estadísticamente los problemas de salud que atentan contra el alumnado, puesto que teniendo clara una tendencia que esté ocurriendo, se puede prevenir tanto el contagio como el decaimiento académico de los estudiantes; este *dashboard* está enfocado a un asunto más estadístico y preventivo, puesto que en la salud no se puede intervenir económicamente para la ayuda a tratamientos o enfermedades crónicas que se presentan dentro del grupo de estudiantes.

## CONCLUSIONES

1. La implementación de un proceso de inteligencia de negocio en una organización permite que la información fluya de una forma natural y controlada desde donde se producen las transacciones del día a día de la organización, hasta convertirlas en información y conocimiento que permiten a los usuarios finales tomar mejores y efectivas decisiones.
2. La implementación de una solución de inteligencia de negocios en el sistema educativo a nivel primario, especialmente en el área pública, beneficiaría al estudiante y al docente, puesto que relaciona estrechamente el rendimiento de ambas partes para lograr un fin común que es dar y recibir educación.
3. Una implementación de inteligencia de negocios actualmente no puede realizarse en cualquier escuela pública a nivel primario, debido a las limitaciones económicas, tecnológicas y de infraestructuras que dichas escuelas padecen.
4. Para establecer una solución de inteligencia de negocios en un entorno educativo, es de suma importancia el enfoque en el rendimiento estudiantil de los alumnos y para llevar un control del mismo, es necesario velar por los factores que le afectan y en las escuelas públicas predomina la falta de un hogar bien integrado, la economía y la salud del estudiante.

5. Las técnicas de inteligencia de negocios muestran una solución innovadora para el tema educacional, puesto que facilitan el control y el análisis de los sucesos que ocurren dentro de la institución, debido al control histórico llevado a través de las diferentes tablas de hechos dentro de un Data Warehouse.



## RECOMENDACIONES

1. Empezar a hacer énfasis en la recolección y almacenamiento de información en los establecimientos educativos a nivel público, puesto que la información que se maneja actualmente es demasiado elemental.
2. El presente trabajo ha servido para el planteamiento de una solución de inteligencia de negocios sobre el ambiente de una escuela pública a nivel primario; por lo que se recomienda que se busque realizar una implementación buscando brindar ayuda al sector educativo.
3. Promover la inteligencia de negocios en establecimientos para un mejor control del rendimiento estudiantil, independientemente del nivel académico que se maneje ya se trate de una institución pública o privada.
4. Que en las instituciones públicas se vea la manera de empezar manteniendo una base de datos para el control propio dentro de la institución, siendo este un comienzo para el cambio en el paradigma gubernamental.



## BIBLIOGRAFÍA

1. DÓMINA, María Cecilia. *Data Warehousing. Relevamiento y aplicación de técnicas de modelado dimensional*. Tesis de Licenciatura en Ciencias de la Computación. Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur, Argentina: 2008. 140 p.
2. KIMBALL, Ralph. *The Data Warehouse Toolkit*. 2da ed. Estados Unidos: John Wiley & Sons, 2002. 421 p.
3. PERALTA, Verónica. *Diseño lógico de Data Warehouses a partir de esquemas conceptuales multidimensionales*. Tesis de maestría. Facultad de Ingeniería, Universidad de la República Pedeciba Informática, Uruguay: 2001. 153 p.
4. PONNIAH, Paulraj. *Data warehousing fundamentals*. Estados Unidos: John Wiley & Sons, 2001. 516 p.
5. REYES, Mario Roberto. *Desarrollo de un Datamart de información académica de estudiantes de la Escuela de Ciencias y Sistemas de la Facultad de Ingeniería de la USAC*. Trabajo de graduación de Ing. en Ciencias y Sistemas, Universidad de San Carlos de Guatemala, Facultad de Ingeniería, 2007. 83 p.

