

10



# UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

---

---

FACULTAD DE INGENIERIA

CONSTRUCCION DE UN DATA WAREHOUSE  
PARA LOS AGENTES PROMOTORES DE LA  
COMISION NACIONAL DEL SISTEMA  
DE AHORRO PARA EL RETIRO

T E S I S  
QUE PARA OBTENER EL TITULO DE  
INGENIERO EN COMPUTACION  
P R E S E N T A N

ABRAHAM GERARDO AVILA MEJIA  
RICARDO MIRANDA GARDUÑO



DIRECTOR DE TESIS:  
M. I. AURELIO ADOLFO MILLAN NAJERA

MEXICO, D.F.

2002



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

10



# UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE INGENIERIA

CONSTRUCCION DE UN DATA WAREHOUSE  
PARA LOS AGENTES PROMOTORES DE LA  
COMISION NACIONAL DEL SISTEMA  
DE AHORRO PARA EL RETIRO

T E S I S  
QUE PARA OBTENER EL TITULO DE  
INGENIERO EN COMPUTACION  
PRESENTAN

ABRAHAM GERARDO AVILA MEJIA  
RICARDO MIRANDA GARDUÑO



DIRECTOR DE TESIS:  
M. I. AURELIO ADOLFO MILLAN NAJERA

MEXICO, D.F.

2002

PAGINACIÓN

DISCONTINUA

Nuestra mente es sola una pequeña ventana  
hacia el vasto universo, pero cuando  
miramos a través de ella, esa ventana  
enmarca el cielo en el exterior.  
A pesar de que no hay marcos en el cielo,  
para nuestro punto de vista ese marco se  
vuelve el límite de la existencia.

## **AGRADECIMIENTOS**

Como un agradecimiento muy especial a mi esposa Lizeth y a mi beba por su enorme cariño y apoyo incondicional.

A mi papá Bernardino y mi mamá Ma. Teresa, a mis hermanos Claudia, Luis Miguel y Sagrario, a Don Rogelio Jiménez y su esposa Doña Fidelia González.

A mi amigo Abraham que realizó esta tesis conmigo.

**Ricardo Miranda Garduño**

## **AGRADECIMIENTOS**

Primeramente a Dios, que me permitió llegar a este objetivo tan anhelado.

A mis padres Abraham y Esperanza por todas sus enseñanzas y apoyo

(Gracias jefes).

A mi tía Lilia por todo lo que me ha enseñado y apoyado

(Gracias por todo tía).

A mis hermanos Carlos y Sandra

(Va por ustedes)

A mis verdaderos cuates, Erik, Alonso, Carlos, Ricardo y Lizeth con los que se

que cuento en las buenas y en las malas

(Forever in the alcohol).

A Jacqueline Villagrana por su apoyo en todo momento y por ser mi inspiración

(tu sabes lo que significas para mi).

**Abraham Gerardo Ávila Mejía**



## INTRODUCCIÓN

La Comisión Nacional del Sistema de Ahorro para el Retiro (CONSAR) es un órgano desconcentrado de la Secretaría de Hacienda y Crédito Público, creado por Decreto Presidencial, publicado en el Diario Oficial de la Federación el día 22 de Julio de 1994, encargado de llevar a cabo la coordinación, regulación, supervisión y vigilancia de los Sistemas de Ahorro para el Retiro y de los participantes en los mismos, que se rige a través de la Ley de los Sistemas de Ahorro para el Retiro y su Reglamento.

Este ordenamiento legal otorga a la CONSAR autonomía técnica y facultades ejecutivas, así como de competencia funcional propia en los términos de dicho cuerpo normativo.

Su misión es proteger el interés de los trabajadores y de sus beneficiarios, asegurando una administración eficiente de su ahorro para el retiro, desempeñando con responsabilidad y compromiso las funciones que le son asignadas

El objetivo principal de los Sistemas de Ahorro para el Retiro es mejorar las condiciones de vida de los trabajadores retirados, pero también lo es, sin duda, un valioso instrumento para fortalecer el ahorro interno de nuestro país, coadyuvando con ello a su desarrollo económico, pues la mayor parte de los recursos son canalizados a fomentar la actividad productiva nacional, la generación de empleos, la construcción de viviendas, el desarrollo de infraestructura y el desarrollo regional y otra parte son invertidos en diversos instrumentos financieros, que son cuidadosamente supervisados.

Una de las obligaciones que tiene la CONSAR es llevar un registro de los Agentes Promotores de las Administradoras (AFORES) con base en el artículo 36 de la Ley del SAR.

Como Agentes Promotores se entiende que son aquellas personas físicas que teniendo una relación de trabajo con una Administradora o en su caso un contrato con ésta, se encuentren autorizados para realizar actividades de registro de cuentas individuales, así como de comercialización, promoción y atención de solicitudes de traspasos apegándose en todo momento a la normatividad aplicable en dichas actividades y llevándolas a cabo en nombre y por cuenta de la Administradora.

Por esto las Administradoras contratan a sus Agentes Promotores a través de un riguroso proceso de selección, asegurándose de que éstos reúnen las mejores condiciones de aptitud y solvencia moral con respecto a la labor que desempeñan, ya que dichas Administradoras responden directamente, y sin afectar la responsabilidad en que incurren los Agentes Promotores, de los actos realizados por éstos; y asimismo, son estas entidades financieras las que tienen que

---



determinar las acciones civiles y penales que corresponden, cuando la relación laboral entre las administradoras y sus Agentes Promotores o bien, entre las personas morales que estas últimas constituyen o contratan para que les presten servicios administrativos, haya terminado por haberse detectado algún tipo de irregularidad que implique alguna posible responsabilidad en dicha materia.

A los Agentes Promotores que realicen actividades que van en contra de las disposiciones contenidas en la Ley, su reglamento y en disposiciones de carácter general, y cuyo comportamiento sea denunciado ya sea por un trabajador, autoridad o por la propia Administradora a la que esté prestando sus servicios, la Comisión les indicará en base a lo que se derive de su actuación y de los elementos que aporten tanto las Administradoras, los trabajadores u otras autoridades relacionadas con la actuación de dichos Agentes Promotores, la sanción correspondiente con base en el artículo 100 de la Ley o en su caso, se cancelará su registro y serán dados de baja del registro de Agentes Promotores de manera definitiva, sin perjuicio de las responsabilidades civiles, administrativas o penales que pudieran corresponder al agente respectivo y a la Administradora de quien dependa.

Las Administradoras también son responsables en todos los casos, de las actividades de los Agentes Promotores en lo que se refiere al trámite, calidad y legitimidad de los documentos de registro o traspaso de las cuentas individuales de los trabajadores y la difusión de las promociones, incluyendo la responsabilidad civil que pudiera derivarse por los daños ocasionados a los trabajadores en el desarrollo de esta actividad

Es por esto que la inscripción y revalidación de los registros de los Agentes Promotores es requisito indispensable para desempeñar las actividades de comercialización, promoción, registro o traspaso, relacionadas con las cuentas individuales de los trabajadores, operadas por las Administradoras.

En este Registro, son inscritas todas aquellas personas contratadas por las Administradoras para realizar dichas actividades, que hayan cumplido los requisitos para actuar como agente promotor.

Para tal efecto las personas que deseen obtener el registro de Agente Promotor deberán cumplir con lo siguiente

I Gozar de reconocida solvencia moral

II Aprobar los exámenes de conocimientos sobre seguridad social

III Que la Administradora a la cual se desea prestar servicios, presente a la Comisión la solicitud de registro del aspirante a Agente Promotor.

La solicitud de registro para ejercer las actividades de Agente Promotor deberá contener la siguiente información relacionada con el aspirante:



- a) Nombres y apellidos
- b) Domicilio particular (calle, número, colonia, delegación, código postal y entidad federativa)
- c) Teléfono en caso de contar con éste
- d) Fecha de nacimiento
- e) Clave de Registro Federal de Contribuyentes.

El registro que se otorgue a los Agentes Promotores tendrá una vigencia de 3 años contada a partir de la fecha de su expedición, el cual podrá ser revalidado por la Comisión por periodos iguales, siempre y cuando el Agente Promotor de que se trate apruebe el examen de actualización que para tal efecto apliquen las Administradoras.

Para cumplir con lo anterior, la CONRSAR requiere de Recursos Humanos comprometidos y dispuestos a dar su máximo esfuerzo combinado con Recursos Tecnológicos

En la presente tesis proponemos una solución integral con base en tecnologías de vanguardia como **The SAS System™** que permitirá apoyar a los directivos de la CONRSAR en la toma de decisiones sobre el control y monitoreo de los Agentes Promotores a nivel Nacional que participan en el Sistema de Ahorro para el Retiro, pues son parte muy importante del sistema ya que establecen el puente de comunicación entre los trabajadores y las Afores

Nuestra tesis propone, siguiendo la metodología **SAS® Rapid Warehousing**, la creación de un Data Warehouse diseñado específicamente en apoyo de decisiones, proporcionando integración y acceso fácil a los datos desde cualquier explorador de web de los Agentes Promotores de la CONRSAR

En el primer capítulo explicamos los conceptos generales, la necesidad de un Data Warehouse así como sus características y beneficios

En el capítulo número dos realizamos una valoración al proyecto para identificar los objetivos de la organización y entender el proyecto Data Warehouse

En el capítulo tercero recolectamos los requerimientos empresariales.

Reservamos el capítulo número cuatro para realizar el diseño del Data Warehouse. En el capítulo número cinco realizamos pruebas al Data Warehouse para asegurarnos que reúne las especificaciones del capítulo tercero. Revisamos el impacto del Data Warehouse de la organización y para ello desarrollamos el capítulo número seis

En el último capítulo escribimos nuestras conclusiones



**TESIS PARA LA CONSTRUCCIÓN  
DE UN DATA WAREHOUSE  
PARA LOS AGENTES PROMOTORES DE LA CONSAR**

**CONTENIDO****CAPÍTULO I**

<b>ANTECEDENTES</b> .....	1
<b>I.1 INTRODUCCIÓN AL DATA WAREHOUSE</b> .....	
I.1.1 DATA WAREHOUSE .....	2
I.1.2 PROCESOS DE TRANSACCIÓN EN LÍNEA (OLPT) .....	4
I.1.3 SISTEMAS DE APOYO DE DECISIÓN (DSS) .....	4
I.1.4 NECESIDAD DE UN DATA WAREHOUSE .....	5
I.1.5 CARACTERÍSTICAS DE UN DATA WAREHOUSE .....	7
I.1.6 BENEFICIOS DE UN DATA WAREHOUSE .....	7
<b>I.2 APRECIACIÓN GLOBAL DE UN DATA WAREHOUSE</b> .....	
I.2.1 MODELADO DEL DATA WAREHOUSE .....	8
I.2.2 IMPLEMENTACIÓN DEL MODELO .....	13
<b>I.3 TECNOLOGÍA DE INFORMACIÓN (IT)</b> .....	20

**CAPÍTULO II**

<b>VALORACIÓN</b> .....	22
<b>II.1 JUSTIFICACIÓN DE LAS HERRAMIENTAS</b> .....	23
<b>II.2 VALORACIÓN</b> .....	
<b>II.2.1 CUALIFICAR EL PROYECTO</b> .....	
II.2.1.1 DEFINICIÓN DE LA CONSAR .....	24
II.2.1.2 DIRECCIONES ESTRATÉGICAS PARA EL PROYECTO .....	25
<b>II.2.2 CONDUCIR LA VALORACIÓN</b> .....	
II.2.2.1 OBJETIVOS DEL PROYECTO Y DE LA CONSAR .....	26
II.2.2.2 NECESIDADES DE ALTO NIVEL DE LA CONSAR .....	26
II.2.2.3 REQUERIMIENTOS EMPRESARIALES DE ALTO NIVEL .....	26
II.2.2.4 NECESIDADES DE INFORMACIÓN DE ALTO NIVEL .....	27
II.2.2.5 FLUJO DEL PROCESO DE LA CONSAR .....	28
II.2.2.6 PERFILES DE USUARIO .....	30
<b>II.2.3 REQUERIMIENTOS II DE ALTO NIVEL</b> .....	
II.2.3.1 ARQUITECTURA OBJETIVO DE IT .....	31
II.2.3.2 DISPOSICIÓN II DE LA CONSAR .....	31
II.2.3.3 FUENTES DE DATOS DE ALTO NIVEL .....	31
II.2.3.4 MODELO DEL SUBJECT .....	33



---

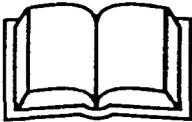
<b>CAPÍTULO III</b>	
<b>REQUERIMIENTOS</b> .....	34
<b>III.1 REQUERIMIENTOS DE LA EMPRESA</b>	
III.1.1 REQUERIMIENTOS DE INFORMACIÓN DETALLADOS .....	35
III.1.2 PERFILES DE ACCESO A LOS USUARIOS .....	38
III.1.3 REQUERIMIENTOS DE SEGURIDAD .....	38
III.1.4 MODELO DIMENSIONAL PRELIMINAR .....	40
<b>III.2 REQUERIMIENTOS DE LA ARQUITECTURA DE INFORMACIÓN</b>	
III.2.1 REQUERIMIENTOS DE LA ARQUITECTURA TÉCNICA DEL DATA WAREHOUSE .....	40
III.2.2 REQUERIMIENTOS DE PLATAFORMA IT .....	41
III.2.3 FUENTES DE DATOS DETALLADAS .....	42
<b>CAPÍTULO IV</b>	
<b>DISEÑO</b> .....	43
<b>IV.1 DISEÑO DEL MODELO FÍSICO DEL DATA WAREHOUSE</b>	
IV.1.1 DISEÑO DEL MODELO FÍSICO DEL DATA WAREHOUSE .....	44
IV.1.2 DISEÑO DE PROCESOS DE METADATOS	
IV.1.2.1 DISEÑO DEL AMBIENTE DE METADATOS .....	48
IV.1.2.2 DISEÑO DEL MODELO DE DATOS PARA EL DEPÓSITO DE METADATOS .....	48
IV.1.2.3 DISEÑO DE LOS PROCESOS DE EXPLOTACIÓN DE METADATOS ....	49
<b>IV.2 DISEÑO DEL MODELO LÓGICO PARA EL DATA WAREHOUSE</b>	
IV.2.1 DISEÑO DEL MODELO LÓGICO PARA EL DATA WAREHOUSE .....	49
IV.2.2 DISEÑO DEL MODELO DE DATOS	
IV.2.2.1 CATEGORIZAR LOS REQUERIMIENTOS EMPRESARIALES .....	50
IV.2.2.2 DISEÑO DE LOS REQUERIMIENTOS DE EXPLOTACIÓN DE INFORMACIÓN .....	50
<b>IV.3 DISEÑO DE LOS PROCESOS DEL DATA WAREHOUSE</b>	
IV.3.1 DISEÑO DEL PROCESO DE EXTRACCIÓN .....	51
IV.3.2 DISEÑO DEL PROCESO DE TRANSFORMACIÓN .....	52
IV.3.3 DISEÑO DEL PROCESO DE CARGA .....	52
IV.3.4 REGISTRO DE METADATOS .....	53
<b>IV.4 DISEÑO DE LA SEGURIDAD</b>	
IV.4.1 DISEÑO DEL ESQUEMA DE SEGURIDAD DE DATOS .....	53
IV.4.2 DISEÑO DE LA INTERFAZ DE SEGURIDAD .....	54
<b>IV.5 DISEÑO DE LA EXPLOTACIÓN DEL DATA WAREHOUSE</b>	
IV.5.1 DISEÑO DE LA INTERFAZ DE EXPLOTACIÓN .....	55
IV.5.2 DISEÑO DE LA INTERFAZ DE APLICACIÓN .....	56

---



---

<b>CAPÍTULO V</b>	
<b>PRUEBAS FINALES</b> .....	57
<b>V.1 PRUEBAS AL SISTEMA</b>	
V.1.1 EJECUTAR PRUEBAS AL SISTEMA .....	58
V.1.2 VERIFICAR LA EJECUCIÓN DEL DATA WAREHOUSE .....	59
<b>CAPÍTULO VI</b>	
<b>REVISIÓN</b> .....	62
<b>VI.1 REVISIÓN DEL IMPACTO EN LA ORGANIZACIÓN</b> .....	63
<b>CONCLUSIONES</b> .....	68
<b>APÉNDICE A : GLOSARIO</b> .....	76



# CAPÍTULO I

## ANTECEDENTES

**CONTENIDO:**

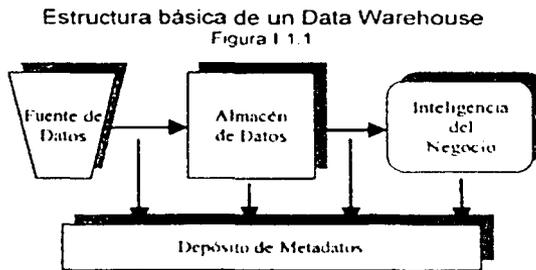
- I.1 INTRODUCCIÓN AL DATA WAREHOUSE**
- I.2 APRECIACIÓN GLOBAL DE UN DATA WAREHOUSE**
- I.3 TECNOLOGÍA DE INFORMACIÓN (IT)**



## CAPITULO I ANTECEDENTES

### I.1 INTRODUCCIÓN AL DATA WAREHOUSE I.1.1 DATA WAREHOUSE

Es un almacén de información de asuntos orientados, diseñado específicamente en apoyo de decisiones (inteligencia de negocios), proporcionando integración y acceso fácil a los datos de una organización.



Estructura básica de un Data Warehouse

Un Data Warehouse es diseñado para reflejar los procesos del negocio (Figura I 1 1)

La inteligencia de negocios esta expresada en reportes y gráficas disponibles en workstations locales

Un Metadata Repository documenta e implementa las reglas y terminología del negocio en cuestion, búsqueda, carga y mantenimiento del almacén de datos.

Las flechas horizontales representan el flujo de datos desde las fuentes de datos al almacén de datos (extracción y carga), y desde el almacén de datos a la inteligencia de negocios (explotación).

Las flechas verticales representan el uso de los Metadatos desde el Repositorio de Metadatos hasta modelar y administrar los flujos de datos, así como también organizar, estructurar el almacén de datos y soportar aplicaciones para la inteligencia de negocios.

Los canales de distribución de inteligencia de negocios son en ocasiones una configuración cliente/servidor.

Esta configuración puede ser implementada usando el modelo de distribución siguiente.

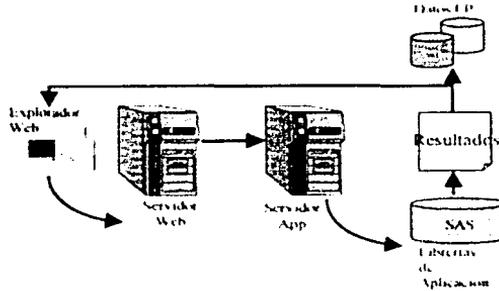


Figura I.1.2  
Distribución basada en Web

#### Distribución basada en web (Figura I.1.2):

- Expande el acceso a la inteligencia de negocios a través de una amplia distribución.
- Decrementa el costo por acceso, el personal sólo necesita buscar en sus workstations.
- Mantiene la centralización, cuando existe:
  - Un almacén detallado de datos
  - Reglas de seguridad
  - Aplicaciones de análisis complejo

Esta distribución puede representar una configuración cliente/servidor de 2, 3 ó 4 capas.

El Web Browser es la primera capa. Está en la Workstation cliente

El servidor Web es la segunda capa. Puede, pero en ocasiones no se ejecuta en la misma plataforma como el servidor de aplicaciones

El servidor de aplicaciones, puede ser la tercera capa, o compartir la plataforma con el servidor web. El servidor de aplicaciones alberga 1 ó más sesiones que puedan responder a las peticiones del cliente canalizadas a través del servidor web.

El Enterprise Data es el Data Warehouse. Puede residir en una cuarta hilera, como una plataforma de datos separada.



### 1.1.2 PROCESOS DE TRANSACCIÓN EN LÍNEA (OLPT)

Es la puerta principal a la información en una organización. Administra las actividades diarias.

Los OLPT coleccionan y editan los datos almacenados.

Typicamente, las aplicaciones que manejan un Sistema OLTP tienen las siguientes características:

- Son de estructura estática con entradas y salidas predecibles
- Son procesos-orientados
  - Tiene un ciclo de desarrollo tradicional
- Están basados en requisitos técnicos bien definidos
- Manejan grandes volúmenes de transacciones
- Se diseñan usando un modelo relacional (ER)
- Guardan datos en Bases de Datos Relacionales
  - Tienen a menudo de 15 a 20 años de vida
  - Apoyan a muchos usuarios

Los datos en un OLTP son dinámicos ya que frecuentemente se presentan actualizaciones debido a las transacciones realizadas, además de ser normalmente exactos y actuales.

### 1.1.3 SISTEMAS DE APOYO DE DECISIÓN (DSS)

Son sistemas que analizan datos de una organización para la toma de decisiones comerciales eficaces y bien informadas, apoyando a las necesidades de los directivos y ejecutivos

Tienen las siguientes características

- Son dinámicos y flexibles con entrada y salida inconstante
- Escritos y mantenidos por analistas comerciales o usuarios
- No tienen un ciclo de desarrollo tradicional
- Traen datos de afuera

Los datos de un DSS son:

- resumidos
- no volátiles
- a menudo denormalizados
- consistentes con respecto a las unidades, nombres, cálculos, etc.,
- de naturaleza histórica
- almacenado en diversas plataformas.

En general, los datos de un DSS no necesitan ser tan actuales como los datos de un OLTP.



### 1.1.4 NECESIDAD DE UN DATA WAREHOUSE

Dentro de una empresa, los sistemas de OLTP recolectan datos y manejan los procesos diarios. Considerando que las aplicaciones de DSS proyectan y manejan decisiones estratégicas ¿Cómo obtener datos externos de sistemas OLTP e internos de las aplicaciones DSS?

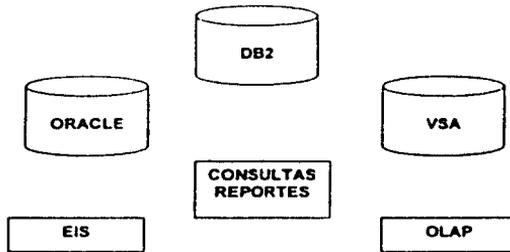


Figura I 1 3  
Sistemas OLTP

Tradicionalmente se tienen dos acercamientos para contestar esta pregunta y son los siguientes:

- El acceso directo a datos del OLTP por los usuarios:  
El acceso directo a datos del OLTP por usuarios comerciales puede ser difícil y sin regularización del acceso pueden ocurrir problemas incluso en credibilidad y consistencia.  
Los datos complejos que brindan soluciones no son a menudo compartidos por otros usuarios o por otras secciones.  
Los usuarios pueden encontrar dificultades al acceder directamente a datos del OLTP debido a que hay acceso limitado de los datos durante horas de trabajo pico y pueden ocurrir crestas de utilidad antes de tener acceso exitoso.  
Navegar por caminos hacia los datos necesarios son a menudo complejos y embarazosos, produciendo resultados cuestionables o incoherentes.  
Las actividades de DSS son a menudo difíciles de poner en palabras via SQL.  
No hay ningún nivel de servicio garantizado y los datos históricos no están típicamente disponibles.
- Extracción de programas

Estos programas son subconjuntos de Bases de Datos o archivos basados en criterios de selección predeterminados



El proceso de extracción puede involucrar redes complejas, creando sincronización y problemas de credibilidad en los datos exportados. Es improbable o inalcanzable que los datos extraídos sean exactos y fiables.

Los problemas de credibilidad pueden unirse a

- ninguna base de datos a tiempo
- diferencias en constrains
- diferentes niveles de extracción
- diversas fuentes externas de datos
- nombres incoherentes, unidades, cálculos, etc.,
- falta de fuentes de datos comunes

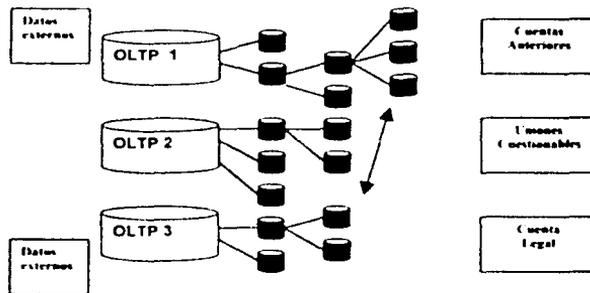


Figura 1 1 4  
Proceso de extracción en sistemas OLTP

Un Data Warehouse permite obtener datos fuera de los sistemas OLTP y dentro de las aplicaciones DSS, mientras se minimizan o incluso se evitan los problemas asociados con el acceso directo de datos OLTP o con la extracción de programas.

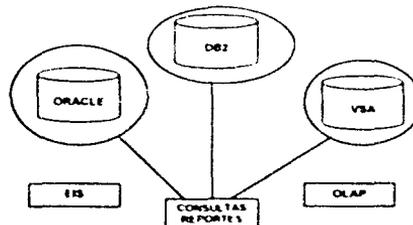


Figura 1 1 5  
Acceso a datos en sistemas OLTP



### I.1.5 CARACTERÍSTICAS DE UN DATA WAREHOUSE

Características como atributos requeridos de un Data Warehouse:

Tiene temas orientados ya que está organizado por asuntos o temas. Un solo tema exige a menudo encausar los datos en varias Bases de Datos de OLTP.

Es Variante en el Tiempo. Los datos históricos son importantes en un ambiente de DSS. La información en un Warehouse es una colección instantánea de datos tomada a intervalos regulares de tiempo.

Es No volátil debido a que la información en un Warehouse es histórica y no cambia. Es principalmente lectura de datos.

Permite la Integración de los datos, frecuentemente se reúnen en numerosos y diversos OLTP.

Es portable al ser implementado en cualquier plataforma.

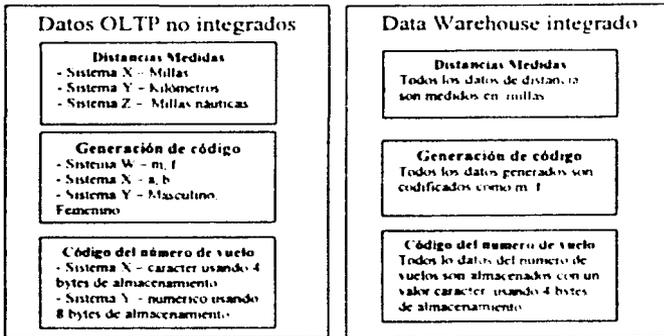


Figura I 1 6  
Ejemplo de las características de un Data Warehouse

### I.1.6 BENEFICIOS DE UN DATA WAREHOUSE

Una Data Warehouse puede ofrecer muchos beneficios a una organización

Permite tener un conocimiento amplio de la ejecución y producción. No sufren problemas debido a constantes consultas y reportes. El Mantenimiento de datos se limita a los datos de la producción.

El acceso a los datos se realiza de una manera oportuna ya que los usuarios requieren la información crucial fácilmente.

Los datos son estructurados y ofrecen una sola versión de los datos.

La calidad de los datos es fiable pues son integrados y consistentes.

Con la exploración y descubrimiento de los datos se abren nuevas posibilidades para valorar tendencias, modelos, o relaciones que puedan conducir a mejoras en los procesos existentes o decisiones en beneficio directo de la organización.



## 1.2 APRECIACIÓN GLOBAL DE UN DATA WAREHOUSE

### 1.2.1 MODELADO DEL DATA WAREHOUSE

#### MODELO LÓGICO

El modelo lógico define las entidades de datos y las relaciones entre estas entidades requeridas para resolver el problema en cuestión.

Consiste en definir la funcionalidad deseada del Data Warehouse.

- confirmando los temas de los datos
- identificando la variación en el tiempo
- identificando la granularidad del tema
  - descomponiendo los temas en entidades de datos compuestas de hechos (facts) y dimensiones.

#### GRANULARIDAD

La Granularidad (o grano) describe el nivel de detalle almacenado en el Data Warehouse físicamente.

Baja Granularidad indica demasiado detalle, mientras que una alta granularidad indica lo contrario.

La Granularidad afecta

- volumen de datos
- mantenimiento de los datos
- clasificación
- nivel de exploración de los datos
- restricciones en consultas y reportes
  - resúmenes adicionales
- recursos de hardware

#### TRADEOFFS

Un grano más pequeño, es decir más detalle, permite un análisis más cercano del proceso.

Un grano más grande requiere menos recursos de almacenamiento.

Con un grano más grande, es decir con tablas más pequeñas, el tiempo de respuesta puede mejorar

#### PROCESOS Y SUBJECTS

Son tareas enfocadas a los procesos de negocios no a transacciones individuales. Revelando términos importantes para los negocios con Subjects particulares.



Figura I 2 1  
Procesos y Subjects

## DIMENSIONES Y FACTS

Casi cada consulta, reporte, análisis, o aplicación divide a los datos en dos categorías:

- facts
- dimensiones

Los Facts son valores que contabilizan personas, lugares y cosas. Los Facts son:

- numéricos
- generalmente continuos
- usados para calcular estadísticas, como medias o sumas

Ejemplos: distancias viajadas, unidades vendidas, pesos, costo en dólares

Las Dimensiones son valores que categorizan a otros valores. Las dimensiones son:

- numéricos o caracter
- discretos para valores numéricos
- usados para identificar clases, particiones, o categorías sobre los cuales se han hecho cálculos
- usados para subconjuntos de datos

Ejemplos: tiempo, producto, región, industria

Un Subject frecuentemente tiene más de un Fact. El número de valores únicos para una Dimensión determina el número de subgrupos. Un Subject casi siempre tiene más de una dimensión.

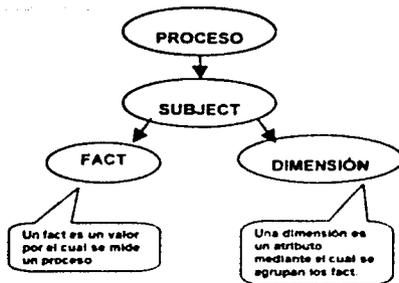


Figura 1 2 2 Dimensiones y Facts

### MODELADO DE DIMENSIONES

Para responder a preguntas específicas se deben ligar los facts con las dimensiones. El modelado dimensional es el proceso de establecer apropiadamente facts y dimensiones y ordenarlos en uno o más esquemas de estrella.

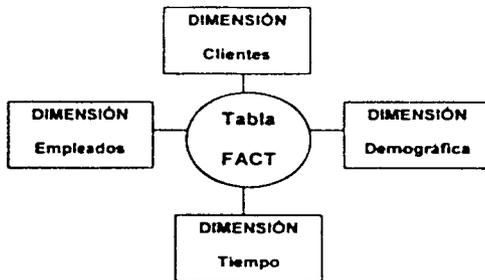


Figura 1 2 3 Modelado Dimensional

### MODELO FÍSICO

El modelo físico define la arquitectura actual de almacenamiento del Data Warehouse.

En el modelo físico se definen las tablas del Data Warehouse.

Los modelos físicos pueden trazarse sobre varios diseños de construcciones, como:



- modelos entidad - relación
- tablas resumidas
- bases de datos multidimensionales
- esquema de estrella
- esquema copo de nieve

Dentro de la implantación de un Warehouse, varias de estas construcciones pueden ser integradas para formar un diseño óptimo.

### MODELO ENTIDAD – RELACIÓN

Tradicionalmente, los modelos entidad - relación han sido usados para diseñar la mayoría de los sistemas OLTP

Dentro de este contexto, dicho modelo está basado en la teoría de conjuntos y en SQL. Altamente normalizado, es decir tiene redundancia de información pequeña o nula.

Optimizado para cambios rápidos de actualización y transacción

No está diseñado para consultarse en un ambiente de Data Warehouse

Este modelo no está bien adaptado para la vasta consulta y la orientación a temas del ambiente Warehouse

Muchas consultas en contra de un Data Warehouse diseñado sobre este modelo requerirán varias uniones posiblemente muy complejas

### NORMALIZACIÓN

La redundancia de los datos es minimizada requiriendo de todas las tablas en el modelo para seguir un conjunto de reglas de normalización

- Cada tabla tiene una llave primaria con valores únicos
- Todas las columnas en una tabla se relacionan directamente a la llave primaria de esa tabla
- Una columna puede contener una llave foránea a otra tabla

### TABLAS SUMARIZADAS

Las tablas resumidas almacenan facts en categorías por dimensiones, como tiempo y región. Un tema (subject) puede tener un gran número de tablas resumidas asociadas

Frecuentemente estas tablas resumidas siguen jerarquias naturales, prestando los datos a las aplicaciones que penetran arriba y debajo de esta información resumida

Hay que considerar cuidadosamente cuando estamos determinando cuales tablas resumidas almacenaremos en el Data Warehouse



## **BASES DE DATOS MULTIDIMENSIONALES**

Una base multidimensional (MDDB) es un diseño que almacena registros en ejes de n-dimensiones, formando una matriz o cubo de dimensiones.

Cada dimensión tiene un número fijo de niveles.

La combinación única del nivel de cada dimensión forma una celda. Cada celda contiene un vector de valores para esa coordenada particular.

A veces, no hay facts que correspondan a una combinación particular de n dimensiones, ocasionando que la celda esté vacía. La dispersión se refiere a la proporción de celdas vacías. Una MDDB dispersa tiene muchas celdas vacías.

Una MDDB cuidadosamente construida puede mejorar su rendimiento significativamente, así como permitir a los usuarios rebanar y cortar en cuadritos los datos fácilmente.

Construir una MDDB requiere de cuidado y prudencia.

## **ESQUEMA DE ESTRELLA**

El esquema de estrella está basado en una tabla central de facts rodeada por varias tablas de dimensiones.

La tabla central de facts es larga y estrecha con muchos renglones y pocas columnas. Alberga los facts medidos y tiene una conexión (llave foránea) para cada dimensión.

Las tablas de dimensión son cortas y amplias con pocos renglones y muchas columnas. Almacenan los atributos de una dimensión.

Usando esquemas de estrella es la manera natural de presentar los resultados del análisis dimensional.

## **ESQUEMA COPO DE NIEVE**

El esquema copo de nieve es una versión más normalizada del esquema de estrella, con la adición de tablas de dimensiones normalizadas.

La normalización de las tablas de dimensiones reduce el sobrealmacenamiento, eliminando valores redundantes en la tabla de dimensión.

Sin embargo, esta normalización puede impactar la ejecución de forma adversa y mitigar su uso.

## **DENORMALIZACIÓN COMPLETA**

Máxima redundancia y no requiere de uniones.

- Unir los hechos y las tablas de dimensión genera una sola tabla de-normalizada.
- Pre-unir los datos de esta manera puede ser útil para datos muy grandes almacenados en un ambiente de procesamiento en paralelo.



## MODELOS ALTERNATIVOS DE DATOS



Figura 1.2.4  
Modelos de datos

La opción de elegir el modelo lógico de datos es un tradeoff entre el tamaño del Data Warehouse y el tiempo de respuesta a las consultas. El tiempo de respuesta decrece cuando nosotros descendemos del esquema de copo de nieve al nivel de sumarización. Por otro lado, se incrementa el tamaño si seguimos en esa dirección.

### 1.2.2 IMPLEMENTACIÓN DEL MODELO

Las características de un Data Warehouse proporcionan una estructura conceptual. A continuación describimos las tareas requeridas para implementar un Data Warehouse.

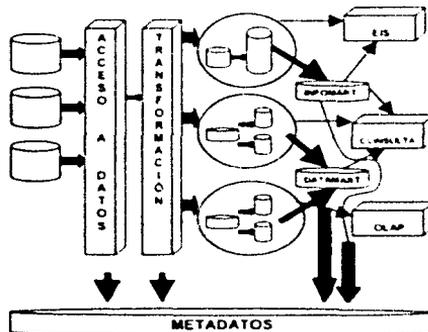


Figura 1.2.5  
Modelo Data Warehouse



## ACCESO Y EXTRACCIÓN DE DATOS

La primera tarea es acceder y extraer los datos de todas las fuentes relevantes de datos como sistemas OLTP, fuentes externas de datos o sistemas heredados.

La consideración debe ser dada de acuerdo a la variedad de estructuras de archivos de estas fuentes de datos:

- bases de datos relacionales
- archivos jerárquicos
- archivos planos
- conjuntos de datos SAS
- otros archivos del propietario.

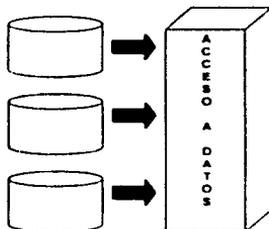


Figura 1 2 6  
Proceso de Acceso y Extracción de datos

## TRANSFORMACIÓN

La transformación es el término de cobertura para todas las áreas que preparan a los datos para ser cargados dentro del Data Warehouse.

Las tareas de transformación incluyen:

Validación – proyección de los datos para identificar valores potencialmente inválidos, fuera de rango, perdidos o duplicados

Depuración – la limpieza de datos grabando o removiendo valores inválidos, duplicados, fuera de rango, perdidos o duplicados.

Integración – conciliando diversos atributos de los datos así como sus valores, en una forma estándar y consistente.

Esto puede ser ejecutado como una tarea por separado, pero a veces es ejecutada simultáneamente con otras tareas de transformación o durante la fase de extracción.

Estructuración – creando nuevos (derivados) valores de los datos y modificando los valores ya existentes



Denormalización – liberando los datos de su estructura OLTP. Este proceso busca optimizar la ejecución del Data Warehouse, y frecuentemente resulta en algo de redundancia en los datos.

Sumarización – agregando datos de los resúmenes estadísticos por computadora sobre datos continuos o creando cuentas sobre datos discretos

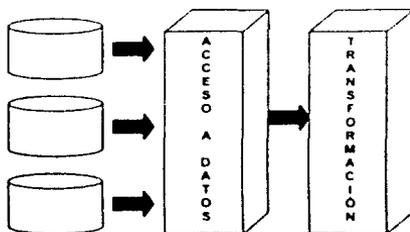


Figura I 2 8  
Proceso de Transformación de datos

### CARGA DEL DATA WAREHOUSE

Después de que los datos externos son transformados, están listos para ser cargados dentro del medio físico de almacenamiento.

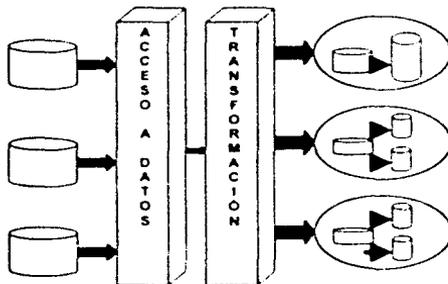


Figura I 2 9  
Proceso de Carga de datos

Después de la carga inicial de datos, hay dos aproximaciones posibles para refrescar el Data Warehouse:



- ciclo de refresco completo: en cada intervalo fijo, el Data Warehouse físico es remplazado en su integridad. Esta estrategia puede ser impráctica porque de incrementarse el volumen de los datos es probable de examinar los datos de la fuente de manera compleja.
- cambio de adquisición de datos: cambiar los datos se refiere a cambios en sumas, eliminaciones, o actualizaciones para la fuente de datos desde la última carga. Sólo los datos cambiados se reemplazan en el Data Warehouse.

### CARGA DE LOS MERCADOS DE DATOS Y LOS MERCADOS DE INFORMACIÓN

Una comunidad de usuarios finales puede muy raramente acceder a los datos en el Warehouse de manera directa.

En lugar de eso, los usuarios finales van en contra de:

- datos localizados en mercados de datos
- reportes, gráficas o programas localizados en mercados de información.

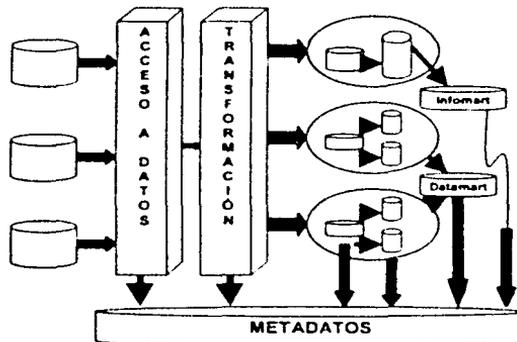


Figura 1 2 10

Proceso de Carga de Mercados de datos y de Mercados de información

### METADATOS

La información es requerida para ayudar al mantenimiento y uso del Data Warehouse. Los metadatos son datos sobre los datos.

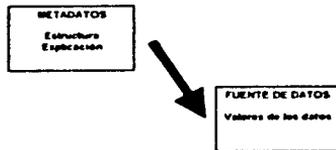


Figura I 2 11  
Metadatos

Los Metadatos son esenciales ya que los problemas significativos pueden incrementarse, como:

Las personas guarden detalles importantes o reglas comerciales que están fortalecidas por el Data Warehouse, actualización de datos a mano o problemas de mantenimiento.

El Metadato es un componente necesario de un Data Warehouse, debe ser colectado, organizado, mantenido y actualizado.

Existen dos puntos de vista de los Metadatos:

Empresarial: son atributos y propiedades del Warehouse para uso de los usuarios empresariales.

Técnico: describe el flujo de los datos desde los sistemas operativos dentro del Data Warehouse.

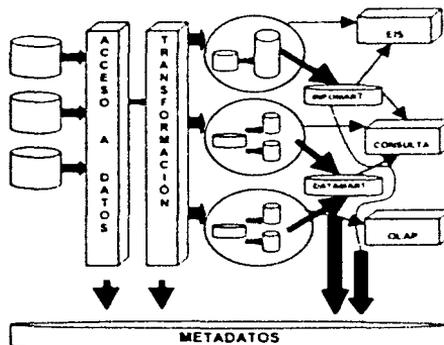


Figura I 2 13  
Modelo Data Warehouse

Un Data Warehouse captura un proceso en uno o más subjects que tienen:

- o Un grupo de facts con uno o más facts
- o Un grupo de dimensiones con varias dimensiones.

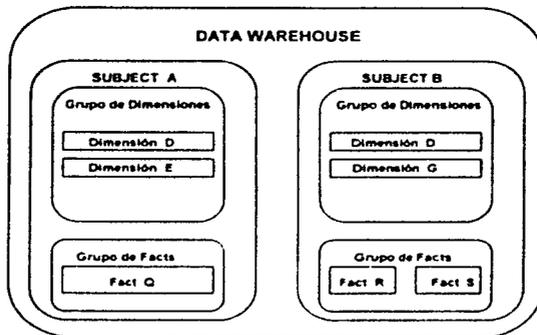


Figura 12 14  
Ambiente Data Warehouse

## EXPLOTACIÓN

Las tareas de extracción, transformación, y carga estructuran y organizan datos, pero estas tareas son una medida del final, no el final mismo.

Un Data Warehouse esta incompleto hasta que se le proveen herramientas para explotar la información dentro del Warehouse

Durante las etapas de diseño, se debe prestar especial atención de como se va a utilizar el Warehouse

Sin esto un Warehouse no tendrá éxito alguno

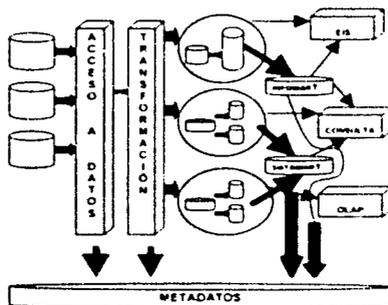


Figura 12 15  
Proceso de explotación de datos

## ORGANIZACIÓN DE UN DATA WAREHOUSE

### Estadísticos Descriptivos

Los facts no son almacenados como un solo simple valor. El valor siempre tiene asociado con él una descripción estadística. Frecuentemente más de un estadístico dado para una fact es almacenado. Cada par de fact-estadístico es almacenado en una tabla.

### Datos sumarizados

- Usar inteligencia comercial para continuar la mejora del proceso implica que algunas consultas se repitan en una base regular
- Los elementos de datos requeridos para las consultas más frecuentes pueden guardarse en forma pre-resumida
- La necesidad de varios tipos y niveles de sumarización cambia a medida que el Data Warehouse se involucra.

Un Data Warehouse esta formado por las siguientes capas

**Fundation Layer:** Tiene el máximo nivel de detalle en un Data Warehouse. Esta capa está incluida en el almacén de datos.

**Tablas Sumarizadas** con un nivel puede contener cualquier estadístico descriptivo.

Otras funciones que proporciona la capa de Tablas Sumarizadas en un Data Warehouse incluye:

- tablas históricas creadas, como la edad del Data Warehouse
- tablas preunidas que se crearon para mejorar la ejecución de las consultas
- mercados de información que contienen colecciones de datos que no pueden guardarse en estructuras tradicionales

**MDDB:** es una colección de sumarizaciones de múltiples niveles. El diseño básico de un MDDB esta orientado a almacenar estadísticos.

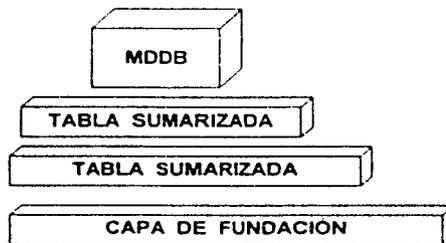


Figura 1.2.16  
Organización de un Data Warehouse

### Estadísticas aditivas

- Count y sum se usan normalmente ya que son totalmente aditivas.



- Algunas estadísticas derivadas, como la media, pueden ser calculadas rápidamente de las aditivas.
- Otras estadísticas no-aditivas, como la razón de dos hechos, puede presentar un desafío al diseñador del DW.
- Las estadísticas Aditivas hacen más fácil el diseño e implementación con capacidad de consultas y penetración ad-hoc.

### LOS DATA WAREHOUSE DISTRIBUIDOS

Muchos Data Warehouse son implementados en un ambiente multiplataforma. Típicamente los diferentes componentes de los datos residen en plataformas diferentes:

- Los OLTP y datos legados residen en mainframes
- El Data Warehouse físico reside en mainframes o UNIX
- Los datos e información de mercados residen en servidores o en desktop de usuario (local)

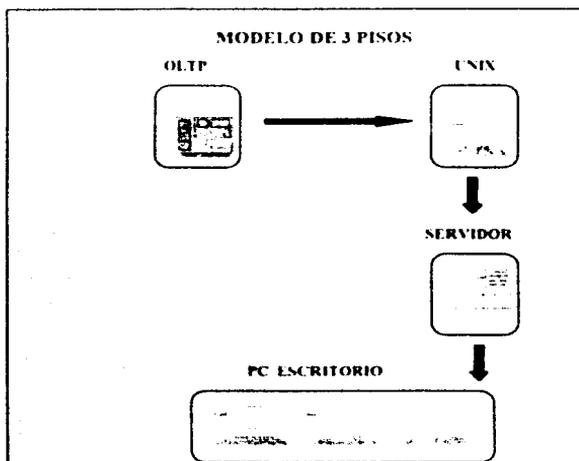


Figura 1 2 1  
Data Warehouse Distribuido

### 1.3 TECNOLOGÍA DE INFORMACIÓN (IT)

Se entiende por tecnología de la información, una técnica utilizada para tratar la información. Comprende toda la tecnología informática y de las



telecomunicaciones, junto con partes muy importantes de electrónica de consumo y difusión.

Actualmente la IT juega un papel importante dentro de los procesos empresariales, ya que crea nuevas necesidades, desarrollos y procedimientos. La implementación de IT en una organización, produce cambios internos que conduce a la mejora continua de todos los procesos de una organización. La tecnología se ha hecho estratégica en cuanto a que es un componente necesario para la ejecución de una estrategia "Business Intelligence". Permite modelar la tecnología que se implementa en las organizaciones, controlar su uso y determinar el efecto sobre su propio trabajo; es decir, el uso efectivo de la tecnología determina el éxito personal y organizacional.

En la tabla siguiente se muestra algunos reportes del impacto de la tecnología de información en los procesos

CAPACIDAD	IMPACTO Y BENEFICIO
Automática	La IT puede reemplazar o reducir la labor humana en un proceso
Analítica	IT puede mejorar el análisis de información y la toma de decisiones
Desintermediación	IT puede ser usada para conectar dos partes dentro de un proceso y eliminar a los agentes intermediarios de un proceso
Geográfica	IT puede transferir y coordinar información con rapidez y facilidad, a través de grandes distancias, creando procesos independientes geográficamente
Informacional	IT puede capturar grandes cantidades de información detallada del proceso con el único objetivo de entenderlo
Integrativa	IT puede coordinar tareas y procesos
Manejo del conocimiento	IT permite la captura y diseminación del conocimiento para mejorar el proceso
Secuencial	IT puede habilitar cambios en la secuencia de tareas de un proceso, a menudo permitiendo paralelismo
Rastreo	IT permite monitorear el detalle de un proceso, sus entradas y salidas
Transaccional	IT puede transformar procesos no estructurados dentro de rutinas de transacción.



# CAPÍTULO II

# VALORACIÓN

## CONTENIDO:

- II.1 JUSTIFICACIÓN DE LAS HERRAMIENTAS
- II.2 VALORACIÓN
  - II.2.1 CUALIFICAR EL PROYECTO
  - II.2.2 CONDUCIR LA VALORACIÓN
  - II.2.3 REQUERIMIENTOS IT DE ALTO NIVEL



---

## CAPÍTULO II

### II.1 JUSTIFICACIÓN DE LAS HERRAMIENTAS

The SAS System™ Software es una herramienta integrada. El software conjunta todos los elementos críticos para administrar, organizar, y explotar la información. Provee datos optimizados para apoyar la toma de decisiones estratégicas y herramientas altamente flexibles que pueden ser:

- usadas para proveer soluciones completas
- integradas dentro de los sistemas, metodologías y herramientas ya existentes.

Podemos tener un enfoque del Data Warehouse en Administración, Organización y Explotación de la información. Es el depósito ideal del Data Warehouse. Está al mismo nivel que cualquier RDBMS comercialmente disponible, con respecto a los tipos de servicio que ofrece, incluyendo:

- Administración de archivos para crear, poblar, borrar y respaldar datos.
- Inventarios de datos
- Procesamiento de consultas para reparar, filtrar, organizar, presentar y desplegar datos.
- Actualizar los procesos para cambiar los datos existentes o agregar nuevos datos

Soporta clasificación y comprensión para ajustar y mantener los datos de manera eficiente. Tablas de seguridad, incluyendo la seguridad ya establecida dentro de los componentes de la infraestructura existente.

Permite que las aplicaciones corran de forma transparente a través de todo el hardware y software de la mayoría de las plataformas. Provee una integración flexible dentro de los sistemas, Warehouses, y aplicaciones existentes. Soporta todos los modelos o configuraciones de procesamiento cliente/servidor

Dentro de The SAS System™ encontramos una herramienta que nos va a servir para hacer la explotación de los datos

Esta herramienta se llama SAS® Web Tools: Accessing MDDB Data Using webEIS® Software.

El webEIS® Software es una solución basada en Java que permite:

- Accesar a MDDBs o tablas de SAS desde un explorador de web
- Crea tablas multidimensionales y gráficas para ver los datos vivos desde un servidor de SAS
- Crea un documento que es mostrado como una Applet de Java con un explorador de Web.



Esta diseñado para análisis de negocio:

- para entender MDDBs
- El conocimiento del lenguaje de Java no es necesario para el uso de webEIS® Software

Los usuarios pueden consultar y ver reportes además de gráficos de datos que están almacenados en una base de datos multidimensional (MDDB) sin ejecutar una sesión de SAS

Una MDDB es un almacén de datos especializado que contiene datos sumarizados para un fácil y rápido acceso

Con esta herramienta los usuarios pueden ver rápidamente grandes cantidades de información como un solo valor en cualquier sección de cruce entre las diferentes dimensiones.

Una dimensión puede ser cualquier vista de los datos que crean un sentido, como el tiempo, geografía, o un producto. El webEIS® Software permite a los usuarios que no tienen acceso al software SAS ver datos en una MDDB. Esta capacidad elimina la necesidad de tener que estar ejecutando SAS en todas las máquinas de los usuarios y provee acceso a reportes y gráficas MDDB en un ambiente Web

Otras características que incluye esta herramienta son las siguientes

- La habilidad para desplegar estadísticas para variables de análisis
- La habilidad de ordenar un gran número de variables.
- La habilidad para ejecutar el **MDDB Report Viewer™** sin frames HTML
- La habilidad para modificar y configurar la salida HTML usando nuevas opciones

## II.2 VALORACIÓN

Objetivo: Valorar la disponibilidad general para sobrellevar el proyecto de Data Warehouse propuesto.

### II.2.1 CUALIFICAR EL PROYECTO

#### II.2.1.1 DEFINICIÓN DE LA CONSAR

Las Afores envían a la CONSAR, la información y movimientos de sus Agentes Promotores, del día de la operación. La CONSAR procesa la información enviada por la Afore. Estos datos son enviados a la Empresa Procesadora el día siguiente a su recepción. Por lo que la información sobre los Agentes Promotores de una Afore se encuentra atrasada por dos días contra la información que se encuentra en la Empresa Procesadora

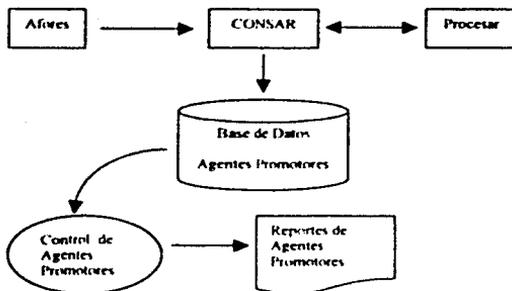


Figura II 2 1

Diagrama de contexto de Agentes Promotores

### II.2.1.2 DIRECCIONES ESTRATÉGICAS PARA EL PROYECTO

A continuación se listan algunos beneficios que recibirá la Comisión después de desarrollar el proyecto Data Warehouse:

- Permite tener un conocimiento amplio de la ejecución y producción de los procesos
- No sufren problemas debido a constantes consultas y reportes.
- El Mantenimiento de datos se limita a los datos de la producción.
- El acceso a los datos se realiza de una manera oportuna ya que los usuarios requieren la información crucial fácilmente.
- Los datos son estructurados.
- Con la exploración y descubrimiento de los datos se abren nuevas posibilidades para valorar tendencias, modelos, o relaciones que puedan conducir a mejoras en los procesos existentes o decisiones en beneficio directo de la Comisión

Definición y cuantificación de los problemas actuales en la Comisión.

- Desconocimiento parcial de la información.
- El tiempo de acceso a datos es muy elevado
- Explotación de la información es complicada y lenta
- Mantenimiento costoso del sistema
- No es fácil realizar reingeniería al sistema ya que los procesos actuales no están bien documentados



## II.2.2 CONDUCIR LA VALORACIÓN

### II.2.2.1 OBJETIVOS DEL PROYECTO Y DE LA CONSAR

#### Misión de la CONSAR

Su misión es proteger el interés de los trabajadores y de sus beneficiarios, asegurando una administración eficiente de su ahorro para el retiro, desempeñando con responsabilidad y compromiso las funciones que le son asignadas

#### Necesidades de la CONSAR

Verificar que se lleve a cabo con fluidez la Vigilancia y Mantenimiento a Agentes Promotores, cuidando que los participantes cumplan con las leyes relativas, así como con las disposiciones de la CONSAR o medidas de control practicadas y así corregir desviaciones en la operación.

#### Objetivos del proyecto

La CONSAR a través de la Dirección General de Inspección requiere Vigilar y Consultar los movimientos efectuados de los Agentes Promotores en las Afores y Empresas Procesadoras en fechas determinadas

Así como reportar los diferentes estatus de los Agentes Promotores por cada Afore, Región, Estado y Periodo ofreciendo así a las personas que toman decisiones en la CONSAR un acceso integrado, consistente, fiable y rápido a los datos, que les permite tomar decisiones basadas en una mejor información ayudando de esta forma al entendimiento del pasado y contar con los elementos para la planeación del futuro de corto, mediano y largo plazo sin necesidad de que el usuario final solicite que se elaboren y ejecuten procesos especiales. Por lo tanto, es importante asegurar que los datos que contenga este Data Warehouse sean adecuados, suficientes y seleccionados de acuerdo a las necesidades de decisión que requiera la CONSAR

### II.2.2.2 NECESIDADES DE ALTO NIVEL DE LA CONSAR

#### Necesidades estratégicas para el Data Warehouse

- Diseñado para reflejar los procesos de control de los Agentes Promotores
- Apoyo de decisiones
- Acceso fácil a datos
- Mantenimiento menos costoso

### II.2.2.3 REQUERIMIENTOS EMPRESARIALES DE ALTO NIVEL

- Subjects requeridos

A continuación presentamos una tabla que contiene los subjects requeridos para la construcción del Data Warehouse



Nombre del Subject	Dimensiones	Descripción
AGENTES PROMOTORES CON SAR	PAIS	Es la variable del DW que categoriza los datos por país.
	REGIÓN	Es la variable del DW que categoriza los datos por región de nuestro país.
	ESTADO	Es la variable del DW que categoriza los datos por estado de nuestro país.
	AFORE	Es la variable del DW que categoriza los datos Afores que se encuentran dadas de alta en el SAR.
	ESTATUS	Es la variable del DW que categoriza los datos por estatus de los Agentes Promotores.
	PERÍODO	Es la variable del DW que categoriza los datos por periodo comprendido entre los años 1997 y 2002.
	AÑO	Es la variable del DW que categoriza los datos por año, de 1997 a 2002.

- Áreas de la CON SAR para la cual el Data Warehouse esta siendo desarrollado.

Departamento	Descripción
Dirección General de Inspección	Dinje y coordina la supervisión e investigación, emitiendo informes respectivos y aplicando medidas preventivas y correctivas a que haya lugar en la operación de las instituciones y entidades participantes en los Sistemas de Ahorro para el Retiro.
Dirección General de Vigilancia	Determinar, establecer y supervisar la instrumentación de Programas de Vigilancia así como evaluar y realizar las medidas correctivas necesarias para su implementación y seguimiento.

#### II 2 2 4 NECESIDADES DE INFORMACION DE ALTO NIVEL

Describimos la información requerida para apoyar a las direcciones y las decisiones necesarias basándose en esa información.



Información Requerida	Decisiones Necesarias
Cantidad de Agentes Promotores	<ul style="list-style-type: none"><li>• Estimular la participación de los Agentes Promotores a nivel nacional de las diversas Afores registradas en el SAR.</li><li>• Implementar programas de capacitación a los Agentes Promotores.</li><li>• Apoyar a las Afores participantes en el Sistema de Ahorro para el Retiro</li><li>• Ayudar a establecer nuevos mecanismos que permitan controlar el estatus de los Agentes Promotores a nivel nacional</li><li>• Implementar una relación directa entre los registros de los Agentes Promotores con el tiempo</li></ul>

### II.2.2.5 FLUJO DEL PROCESO DE LA CONSAR

Para cada proceso realizado en las siguientes áreas definimos el flujo de información y fuentes de datos.

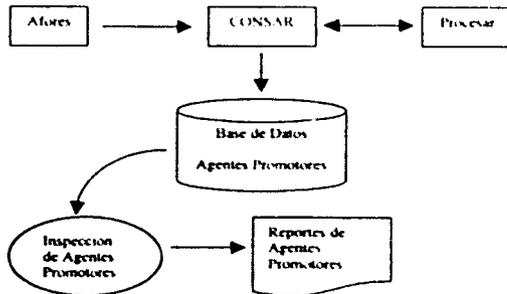


Figura II 2 2

Dirección General de Inspección



Nombre de la Fuente de Datos	Fuente de Datos	Descripción
AFORE	Afore_nombre	Nombre de la Afore
	Estado_id	Identificador del Estado
	Afore_criterio_valuacion	Criterio de valuación de la Afore
ESTADO	Estado_id	Identificador del Estado
	Region_id	Identificador de la Región a la cual pertenece dentro del País
	Estado_nombre	Nombre del Estado
AP_AGENTE_PROMOTOR	Afore_id	Identificador de la Afore participante
	Estado_id	Identificador del Estado
	Agente_numero	Número único del Agente Promotor
	Agente_f_Autorizacion	Fecha de autorización del Agente Promotor
AP_SEGUIMIENTO	Id_seguimiento	Identificador del seguimiento del Agente Promotor
	Afore_id	Identificador de la Afore participante
	Id_estatus	Identificador del estatus del Agente Promotor
	Agente_numero	Número único del Agente Promotor
AP_ESTATUS	Id_estatus	Identificador del estatus del Agente Promotor
	Estatus_descripcion	Descripción del estatus del Agente Promotor

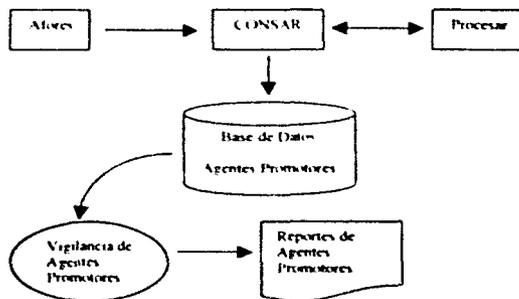


Figura II 2 3  
Dirección General de Vigilancia



Nombre de la Fuente de Datos	Fuente de Datos	Descripción
AFORE	Afore_nombre	Nombre de la Afore
	Estado_id	Identificador del Estado
	Afore_criterio_valuacion	Criterio de valuación de la Afore
ESTADO	Estado_id	Identificador del Estado
	Region_id	Identificador de la Región a la cual pertenece dentro del País
	Estado_nombre	Nombre del Estado
AP_AGENTE_PROMOTOR	Afore_id	Identificador de la Afore participante
	Estado_id	Identificador del Estado
	Agente_numero	Número único del Agente Promotor
	Agente_f_Autorizacion	Fecha de autorización del Agente Promotor
AP_SEGUIMIENTO	Id_seguimiento	Identificador del seguimiento del Agente Promotor
	Afore_id	Identificador de la Afore participante
	Id_estatus	Identificador del estatus del Agente Promotor
	Agente_numero	Número único del Agente Promotor
AP_ESTATUS	Id_estatus	Identificador del estatus del Agente Promotor
	Estatus_descripcion	Descripción del estatus del Agente Promotor

### II.2.2.6 PERFILES DE USUARIO

A continuación describimos el tipo de usuarios que utilizará el Data Warehouse resultante en términos de

Frecuencia de Uso	Tipo de Uso	Limitaciones de Seguridad	Número aproximado de cada nivel	Categoría de usuario
Diana	Total	Password	16	Administración
Mensual	Read Only	Password	14	Usuario



### II.2.3 REQUERIMIENTOS IT DE ALTO NIVEL

#### II.2.3.1 ARQUITECTURA OBJETIVO DE IT

Definimos la siguiente plataforma, infraestructura de red, sistemas operativos y recursos de la organización para apoyar el Data Warehouse.

Nombre y tipo de los sistemas operativos	Plataforma	Conectividad de red y protocolos	Espacio disponible de almacenamiento
Unix 8.0	Sun Solaris Enterprise 10000	Red LAN con topología de anillo Protocolo TCP	1.5 TB en 56 discos de 36 GB

#### II.2.3.2 DISPOSICIÓN IT DE LA CONSAR

Identificamos la plataforma, infraestructura de red, sistemas operativos y recursos de la organización actuales e incluyen lo siguiente:

Nombre y tipo de los sistemas operativos	Plataforma	Conectividad de red y protocolos	Espacio disponible de almacenamiento
Unix 2.5.1	Sun Solaris Enterprise 3000	Red LAN con topología de anillo Protocolo TCP	6 discos duros de 4.2 GB y 5 de 9 GB

#### II.2.3.3 FUENTES DE DATOS DE ALTO NIVEL

Definimos los siguientes elementos:

Depósito	Data Warehouse
Plataformas	Sun Solaris Enterprise 10000
Volumen de datos estimado	5 GB de información

Fuentes de datos que serán usados para actualizar el Data Warehouse:

Nombre de la Fuente de Datos	Fuente de Datos	Descripción
AFORE	Afore_nombre	Nombre de la Afore
	Estado_id	Identificador del Estado
	Afore_criterio_valuacion	Criterio de valuación de la Afore
ESTADO	Estado_id	Identificador del Estado
	Region_id	Identificador de la Region a la cual pertenece dentro del País
	Estado_nombre	Nombre del Estado



AP_AGENTE_ PROMOTOR	Afore_id	Identificador de la Afore participante
	Estado_id	Identificador del Estado
	Agente_numero	Número único del Agente Promotor
	Agente_f_Autorizacion	Fecha de autorización del Agente Promotor
AP_ SEGUIMIENTO	Id_seguimiento	Identificador del seguimiento del Agente Promotor
	Afore_id	Identificador de la Afore participante
	Id_estatus	Identificador del estatus del Agente Promotor
	Agente_numero	Número único del Agente Promotor
AP_ESTATUS	Id_estatus	Identificador del estatus del Agente Promotor
	Estatus_descripcion	Descripción del estatus del Agente Promotor

Métodos usados para actualizar las fuentes de datos:

- Procesos de actualización del Data Warehouse

Frecuencia de actualización:

- Mensualmente



### II.2.3.4 MODELO DEL SUBJECT

Creemos el Modelo del Subject el cual cubre las necesidades de información de la CONSAM.

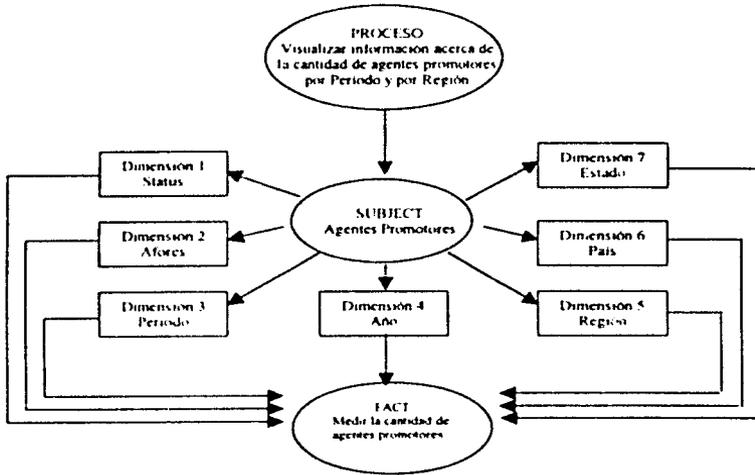


Figura II.2.4  
Modelo del Subject



# **CAPÍTULO III**

# **REQUERIMIENTOS**

**CONTENIDO:**

- III.1 REQUERIMIENTOS DE LA EMPRESA**
- III.2 REQUERIMIENTOS DE LA ARQUITECTURA DE INFORMACIÓN**



### CAPÍTULO III REQUERIMIENTOS

El propósito de la fase de requerimientos es unir los requerimientos IT de la CONSAR y definir los subjects del Data Warehouse.

#### III.1 REQUERIMIENTOS DE LA EMPRESA III.1.1 REQUERIMIENTOS DE INFORMACIÓN DETALLADOS

Determinar los requerimientos detallados de la empresa que deben ser apoyados por el Data Warehouse. Incluyen:

- Fuentes de datos específicos

Actualmente se lleva un registro de los Agentes Promotores a través de una base de datos relacional llamada supervisión.

Esta base de datos contiene las siguientes tablas, en las cuales residen los datos sobre los que nos apoyaremos para poblar nuestra aplicación Data Warehouse:

Nombre de la Fuente de Datos	Descripción
AFORE	Catálogo de Afores participantes en el SAR
ESTADO	Catálogo de regiones y estados en el país
AP_AGENTE_PROMOTOR	Tabla de los movimientos de los Agentes Promotores a nivel nacional registrados en las Afores
AP_SEGUIMIENTO	Tabla del seguimiento de los movimientos de los Agentes Promotores
AP_ESTATUS	Tabla del estatus que guardan los Agentes Promotores

- Requerimientos de información

La CONSAR necesita consultar información de manera sumariada acerca de la cantidad de Agentes Promotores a nivel nacional

Esta información deberá contener diversos campos, tales como región, periodo, año, status, afore, país, y estado

Dado que es a través de estos campos de donde vamos a obtener la información a ser analizada, necesitamos que se pueda consultar cualquier cruce de información entre ellos, es decir, si tenemos los siguientes campos:



Renglón		Columna	Columna	Renglón
Periodo, Año	Pais, Región, Estado, Afore,	VS	Periodo, Año	Pais, Región, Estado, Afore,
Estatus			Estatus	
Periodo			Periodo	
Año			Año	
Pais			Pais	
Región			Región	
Estado			Estado	
Afore			Afore	
Estatus			Estatus	

En la siguiente tabla se muestra cada cruce de información que es posible, así como una explicación de los datos que se generan al realizar dicho cruce.

Cruce de Información	Explicación
PAÍS VS. PERIODO	Este cruce de información nos muestra los datos existentes acerca de los Agentes Promotores a nivel nacional a través de un periodo de tiempo comprendido entre 1997 y 2002
PAÍS VS. AÑO	A diferencia de la anterior intersección de información, en esta nos muestra la información referente a los Agentes Promotores existentes en el país, pero a través de cada uno de los años (1997, 1998, 1999, 2000, 2001 y 2002)
REGIÓN VS. PERIODO	En este cruce encontramos la información referente a las diferentes regiones del país comprendida en un periodo de tiempo (1997-2002)
REGIÓN VS. AÑO	En este otro cruce encontramos información más detallada que en el anterior ya que aquí se nos muestran los datos referentes a cada una de las regiones del país referentes a cada año entre 1997 y 2002
ESTADO VS PERIODO	En esta intersección encontramos los datos de los Agentes Promotores agrupados por cada estado de la República Mexicana en un periodo de tiempo (1997-2002)
ESTADO VS AÑO	Este cruce nos muestra la información agrupada por estado de la República Mexicana y desglosada por cada año comprendido de 1997 a 2002
AFORE VS PERIODO	En este cruce se nos proporciona información referente a las Afores dadas de alta en la CONRSAR en un periodo de tiempo (1997-2002)



AFORE VS. AÑO	En este otro cruce se nos proporciona la información referente a las Afores dadas de alta pero por cada año comprendido entre 1997 y 2002
STATUS VS. PERÍODO	En esta intersección encontramos la información referente al status (desempeño de los Agentes Promotores) a través de un periodo de tiempo (1997-2002).
STATUS VS AÑO	En este cruce encontramos la información referente al status (desempeño de los Agentes Promotores) a través de cada año entre 1997 y 2002

Se requiere que para cada consulta se elabore su respectivo reporte. Los reportes deben incluir una gráfica ilustrativa del reporte. El sistema debe ser capaz de exportar el reporte a una hoja de cálculo de Excel. Además de contar con este cruce de información, la consulta deberá permitir aplicar al resultado de la cantidad de Agentes Promotores al menos los siguientes estadísticos:

- Promedio
  - Desviación Estándar
  - Máximo
  - Mínimo
- 
- usuarios del negocio

Departamento	Usuarios
Dirección General de Inspección	Javier Martínez
	Ma. Luisa Mejía
	José Luis Flores
	Julio Viramontes
	Lizeth Jiménez
Dirección General de Vigilancia	Martha Martínez
	Jaime Sosa
	Clara Rodríguez
	Edmundo Salazar
	Martin Gasca

- ciclo de frecuencia

La frecuencia de actualización de los datos en el Data Warehouse es de un mes, llevándose a cabo el primer día hábil de cada mes, esto es debido a que mensualmente se registran los movimientos efectuados de los Agentes Promotores en las Afores y Empresas Procesadoras



### III.1.2 PERFILES DE ACCESO A LOS USUARIOS

Documentamos los tipos de usuarios que usarán el Data Warehouse.

Para cada perfil de usuario identificar el tipo de uso, restricciones de seguridad, ubicación, ambiente de cómputo o cualquier otro tipo de restricción.

Tipo	Nivel de Acceso
Usuario	Read Only
Administración	Total

### III.1.3 REQUERIMIENTOS DE SEGURIDAD.

Las siguientes consideraciones de seguridad se realizan con base en la implementación de un sistema Data Warehouse que soporta transacciones en línea, servidores remotos, así como una sólida robustez en la información que nos va a proporcionar.

#### Seguridad Física.

Para considerar la seguridad física se pueden implementar diversos mecanismos tales como:

#### A nivel preventivo

- Acceso controlado al equipo en base a diversas políticas tales como
- Uso del equipo por personal autorizado
- Sólo podrá tener acceso al equipo aquella persona que cuente con autorización para operar sobre el Data Warehouse
- El personal que carezca de autorización y que requiera hacer uso de la información contenida en el Data Warehouse podrá solicitar la ayuda de los analistas de sistemas para hacer uso del sistema mediante una autorización escrita por el responsable del área que solicite la información
- Introducción de una clave personal de usuario para aquellas personas que cuenten con la autorización para operar sobre el Data Warehouse, la cual se aloja en una base de datos junto con su ID de usuario. Esto con el objetivo de tener un control sobre quién entra al sistema
- Respaldo continuo de la información, para asegurar la integridad total de la información



### A nivel correctivo

- El aseguramiento del equipo en alguna agencia de seguros para que en caso de posible siniestro no exista una pérdida total por parte de los elementos físicos que controlan y dan soporte al sistema.

### Seguridad Lógica

La seguridad lógica se refiere al control lógico dentro del software, y esto se implementa mediante la construcción de passwords para los usuarios que estén autorizados para utilizar el Data Warehouse, así como también se restringirá el acceso por medio de permisos, en base al sistema operativo que use como plataforma el sistema del Data Warehouse.

### Conducta del usuario

Para asegurar el éxito del sistema es necesario establecer campañas constantes sobre la funcionalidad y logros que se alcanzarán con el sistema, los cuales crearán una conciencia en el usuario y logren formar en él un interés en el uso del sistema.

También es importante el análisis del uso y trayecto del sistema por parte de los usuarios para poder detectar fugas de información y posibles problemas con los datos accedados del sistema

Usuario	Tipo	Nivel de Acceso	Ubicación	Ambiente de Cómputo
Javier Martínez	Usuario	Total	DGI	PC
Ma. Luisa Mejía	Usuario	Total	DGI	PC
José Luis Flores	Usuario	Total	DGI	PC
Julio Viramontes	Usuario	Total	DGI	PC
Lizeth Jiménez	Usuario	Total	DGI	PC
Martha Martínez	Usuario	Total	DGAR	PC
Jaime Sosa	Usuario	Total	DGAR	PC
Clara Rodríguez	Administración	Read Only	DGAR	PC
Edmundo Salazar	Administración	Read Only	DGAR	PC
Martín Gasca	Administración	Read Only	DGAR	PC



### III.1.4 MODELO DIMENSIONAL PRELIMINAR

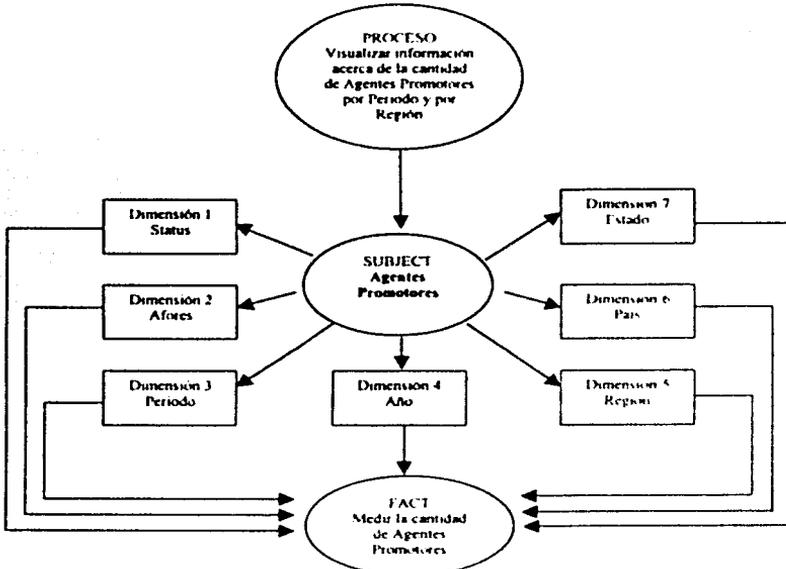


Figura III 1 1  
Modelo Dimensional Preliminar

## III.2 REQUERIMIENTOS DE LA ARQUITECTURA DE INFORMACIÓN

### II.2.1 REQUERIMIENTOS DE LA ARQUITECTURA TÉCNICA DEL DATA WAREHOUSE

Para la implementación del proyecto Data Warehouse en la CONSAR, se planea utilizar equipo de vanguardia para soportar toda la funcionalidad que traera este proyecto así como para responder a cualquier tipo de proceso al que sea sometido.

Para tal efecto se utilizará el dominio del servidor Sun Enterprise 10000 con las siguientes características técnicas.



Datos \ Dominio	DW CONSAR (Dominio a utilizar)	Supervisión	Riesgos
Nombre del Sistema Operativo	Solaris 8.0	Solaris 8.0	Solaris 8.0
Plataforma	Unix	Unix	Unix
Procesadores	8 @ 466 MHz	4 @ 466 MHz	4 @ 466 MHz
Memoria RAM	8 Gb	8 Gb	8 Gb
Almacenamiento	Ymetrix 1.5 Tb	Ymetrix 1.5 Tb	Ymetrix 1.5 Tb
RDBMS	Sybase 12.0	Sybase 12.0	Sybase 12.0
Conectividad de red	LAN	LAN	LAN
Protocolo	TCP/IP	TCP/IP	TCP/IP

### III.2.2 REQUERIMIENTOS DE PLATAFORMA IT.

La plataforma IT, es la infraestructura tecnológica con la que se cuenta actualmente y con la que vamos a proveer de datos a nuestro Data Warehouse. Podríamos decir que se refiere a los sistemas fuente o aquellos que nos van a servir para alimentar de información al Data Warehouse.

Estos recursos son en esencia, la red con que se cuenta, el tipo de servidores en donde descansan los sistemas de bases de datos relacionales, y en general cualquier medio tecnológico con el que se cuente

En particular hay una base de datos que nos interesa, es la base de datos llamada Supervisión, esta base de datos cuenta con todas las tablas que requerimos para el proyecto Data Warehouse y es en si la que nos va a proveer de toda nuestra información en este proyecto

Esta base de datos se encuentra en el servidor SunConsar con la IP 8.39.3.151.

Este servidor es del tipo Sun Enterprise 3000 con las siguientes características técnicas:

Datos \ Servidor	SunConsar
Nombre del Sistema Operativo	Solaris 2.5.1
Plataforma	Unix
Procesadores	4 @ 400 MHz
Memoria RAM	4 Gb
Almacenamiento	6 DD de 4.2 Gb y 5 DD de 9 Gb
Otros medios de almacenamiento	Cinta tipo DDS
RDBMS	Sybase 11.5
Conectividad de red	LAN
Protocolo	TCP/IP
Dirección IP	8 39 3 151



### III.2.3 FUENTES DE DATOS DETALLADOS

La siguiente tabla contiene información de las fuentes de datos.

NOMBRE DE LA FUENTE DE DATOS	FUENTE DE DATOS	DESCRIPCIÓN
AFORE	Afore_nombre	Nombre de la Afore
	Estado_id	Identificador del Estado
	Afore_criterio_valuacion	Criterio de valuación de la Afore
ESTADO	Estado_id	Identificador del Estado
	Region_id	Identificador de la Región a la cual pertenece dentro del País
	Estado_nombre	Nombre del Estado
AP_AGENTE_PROMOTOR	Afore_id	Identificador de la Afore participante
	Estado_id	Identificador del Estado
	Agente_numero	Número único del Agente Promotor
	Agente_f_Autorizacion	Fecha de autorización del Agente Promotor
AP_SEGUIMIENTO	Id_seguimiento	Identificador del seguimiento del Agente Promotor
	Afore_id	Identificador de la Afore participante
	Id_estatus	Identificador del estatus del Agente Promotor
	Agente_numero	Número único del Agente Promotor
AP_ESTATUS	Id_estatus	Identificador del estatus del Agente Promotor
	Estatus_descripcion	Descripción del estatus del Agente Promotor



# CAPÍTULO IV

## DISEÑO

### CONTENIDO:

- IV.1 DISEÑO DEL MODELO FÍSICO DEL DATA WAREHOUSE
- IV.2 DISEÑO DEL MODELO LÓGICO PARA EL DATA WAREHOUSE
- IV.3 DISEÑO DE LOS PROCESOS DEL DATA WAREHOUSE
- IV.4 DISEÑO DE LA SEGURIDAD
- IV.5 DISEÑO DE LA EXPLOTACIÓN DEL DATA WAREHOUSE

**CAPÍTULO IV****DISEÑO****IV.1 MODELO FÍSICO DEL DATA WAREHOUSE****IV.1.1 DISEÑO DEL MODELO FÍSICO DEL DATA WAREHOUSE**

Basándonos en la fuente de datos, y en las transformaciones que se requieren, determinamos el ambiente de hardware y software del Data Warehouse. Para esta determinación tomamos en cuenta lo siguiente:

- o Componentes de software que serán usados para extraer los datos

<b>Nombre del Software</b>	<b>Versión</b>
The SAS System for Windows 98	8 2
SAS Enterprise Guide	1.0
Sybase	12 0

- o Como serán transformados los datos

Al realizar una validación para proyectar los datos para identificar valores potencialmente inválidos, fuera de rango, perdidos o duplicados encontramos los siguientes datos:

<b>AP_AGENTE_PROMOTOR</b>	<b>Tipo de Dato</b>	<b>Long</b>	<b>Nulos</b>	<b>Descripción</b>
Agente_nombre	Varchar	40	Not null	Nombre del Agente Promotor
Agente_RFC	Varchar	13	Not null	Registro Federal de Causantes
Agente_NSS	Varchar	11	Not null	Número de Seguridad Social
Agente_CURP	Varchar	18	Not null	Clave Unica de Registro de Población

<b>AFORE</b>	<b>Tipo de Dato</b>	<b>Long</b>	<b>Nulos</b>	<b>Descripción</b>
Afore_razon_social	varchar	100	Not null	Razón social de la Afore
Afore_fecha_autonización	smalldat etime		Not null	Fecha de Autorización de la Afore



AP_SEGUIMIENTO	Tipo de Dato	Long	Nulos	Descripción
Seguimiento_f_evento	Datetime		Not null	Fecha del evento
Seguimiento_f_examen	Datetime		Not null	Fecha de examen
Seguimiento_calificacion	Datetime		Not null	Fecha de calificación
Seguimiento_tipo_registro	int	2	Not null	Tipo de registro
Seguimiento_f_renuncia	Datetime		Not null	Fecha de renuncia

Depuramos para limpiar los datos grabando o removiendo valores inválidos, duplicados, fuera de rango o perdidos.

AP_SEGUIMIENTO	Tipo de Dato	Valores Inválidos
Seguimiento_f_evento	Datetime	Mes Dia

La Integración de los datos la realizamos conciliando diversos atributos de los datos, así como sus valores, en una forma estándar y consistente.

Nombre de la Fuente de Datos	Nombre de los Datos	Descripción	Tipo de Dato	Long
Estado	Estado_nombre	Estado al que pertenece el Agente Promotor	Varchar	50
Estado	Estado_id	Identificador del estado	Smallint	2
Estado	Region_id	Región del país a la cual pertenece el Agente Promotor	Smallint	2
Afore	Afore_id	Identificador de la Afore	Smallint	2
Afore	Afore_nombre_corto	Nombre de la Afore participante en el SAR	Varchar	50
Afore	Afore_criterio_valuación	Criterio de valuación que aplica la CONSAR a las Afores participantes	Char	1
AP_AGENTE_PROMOTOR	Num_agente	Es el número de identificación del Agente Promotor	Char	10
AP_AGENTE_PROMOTOR	Estado_id	Identificador del Estado	Smallint	2
AP_AGENTE_PROMOTOR	Afore_id	Identificador de la Afore	Smallint	2
AP_AGENTE_PROMOTOR	Afore_f_autorizacion	Fecha de autorización del Agente Promotor	Datetime	



AP_SEGUIMIENTO	Id_estatus	Identificador del status del Agente Promotor	Smallint	2
AP_SEGUIMIENTO	Num_agente	Es el número de identificación del Agente Promotor	Char	10
AP_ESTATUS	Id_estatus	Identificador del status del Agente Promotor	Smallint	2
AP_ESTATUS	Estatus_descripción	Descripción del status	Varchar	100

Realizamos la Estructuración de los datos creando nuevos (derivados) valores de los datos y modificando los valores ya existentes.

Datos nuevos:

Nombre de los Datos	Descripción	Tipo	Long
Pais	Agrupar a los Agentes Promotores a nivel nacional	Varchar	10
Periodo	Agrupar a los Agentes Promotores desde 1997 a 2002	Datetime	
Cantidad_Agente	Cuenta la cantidad de Agentes Promotores	int	1

Datos modificados:

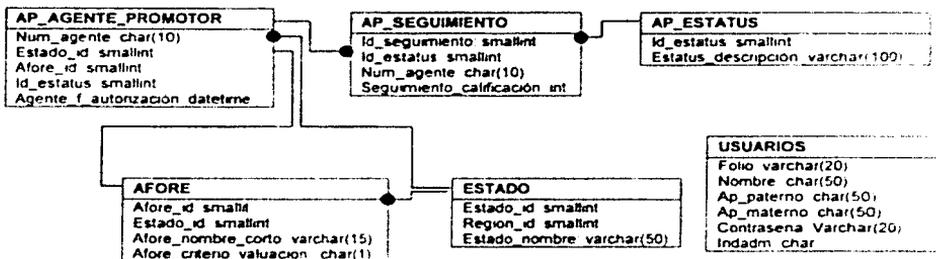
AP_SEGUIMIENTO	Tipo de Dato	Valores Válidos
Seguimiento_f_evento	Datetime	Año

Realizamos el proceso de Denormalización el cual consiste en liberar los datos de su estructura OLTP. Este proceso busca optimizar la ejecución del Data Warehouse, y sin embargo frecuentemente resulta en poca redundancia de datos.

TESIS CON  
FALLA DE ORIGEN



## Modelo Entidad – Relación



## Modelo Denormalizado

Nombre de los Datos	Tipo de Dato	Long
Num_agente	char	19
Agente_f_autorizacion	datetime	
Seguimiento_calificación	int	
Estatus_descripción	varchar	100
Afore_nombre_corto	varchar	15
Afore_criterio_valuacion	char	1
Region_id	smallint	
Estado_nombre	Varchar	50

La sumarización consiste en agregar datos mediante el cálculo de estadísticos sumarizados sobre datos continuos o discretos. A continuación mostramos una tabla que contiene la información del dato al que se le aplicará dicho proceso.

Nombre de los Datos Sumarizados	Descripción	Tipo de Dato	Long
Cantidad_Agente	Cuenta la cantidad de Agentes Promotores	int	1

- o Infraestructura para realizar dichas transformaciones

Depósito	Agentes_Promotores
Data Warehouse	
Ubicación	//opt/SAS/sasuser
Plataforma de Hardware	Sun Solaris Enterprise 10000
Volumen de Datos Estimado	5 GB de información



### NIVELES DE DETALLE EN EL DATA WAREHOUSE

DESCRIPCIÓN

INFORMACIÓN

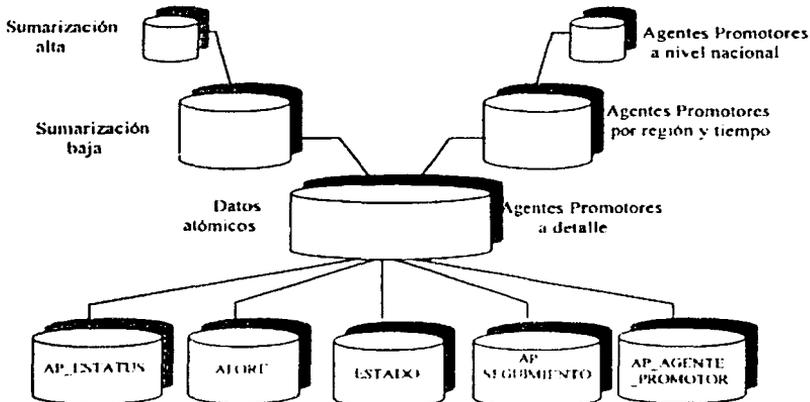


Figura IV.1.1  
Nivel de detalle del Data Warehouse

#### IV.1.2 DISEÑO DE PROCESOS DE METADATOS IV.1.2.1 DISEÑO DEL AMBIENTE DE METADATOS

A continuación mostramos la información de la herramienta que se utilizará para dar soporte a los metadatos:

<b>Herramienta de soporte de los metadatos</b>
EIS / OLAP Application Builder

#### IV.1.2.2 DISEÑO DEL MODELO DE DATOS PARA EL DEPÓSITO DE METADATOS

La siguiente tabla muestra información necesaria para el mantenimiento de los metadatos:

<b>Nombre del repositorio</b>	Nombre del repositorio de metadatos
<b>Nombre de la tabla</b>	Nombre de la tabla de metadatos
<b>Host</b>	Host en donde se almacenarán los metadatos



<b>Atributos</b>	Atributos de la tabla de metadatos
------------------	------------------------------------

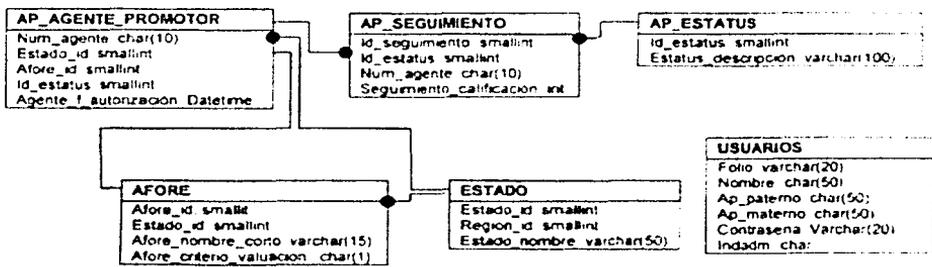
### IV.1.2.3 DISEÑO DE LOS PROCESOS DE EXPLOTACIÓN DE METADATOS

Los procesos por medio de los cuales se proveerán los metadatos a los usuarios será por medio de la interfaz que proporciona la herramienta EIS/OLAP Application Builder.

## IV.2 DISEÑO DEL MODELO LÓGICO PARA EL DATA WAREHOUSE

### IV.2.1 DISEÑO DEL MODELO LÓGICO PARA EL DATA WAREHOUSE

El modelo lógico define las entidades de datos y las relaciones entre estas entidades requeridas para resolver el problema en cuestión



### Funcionalidad deseada del Data Warehouse

<b>Subject para los datos</b>		Agentes Promotores
<b>Variación en el tiempo</b>		Mensualmente
<b>Granularidad del tema</b>		Cantidad de Agentes Promotores
<b>Dimensiones</b>		Periodo
		Año
		País
		Región
		Estado
		Afore
		Estatus Descripción
<b>Jerarquías</b>	<b>Tiempo</b>	Periodo Año
	<b>Región</b>	País Región Estado Afore Estatus
<b>Hechos (facts)</b>		Cantidad de Agentes Promotores
<b>Variable de Análisis</b>		Cantidad de Agentes Promotores
<b>Granularidad</b>		Cantidad de Agentes Promotores



## IV.2.2 DISEÑO DEL MODELO DE DATOS

### IV.2.2.1 CATEGORIZAR LOS REQUERIMIENTOS EMPRESARIALES

Determinamos cuáles requerimientos empresariales se pueden satisfacer con el Data Warehouse.

- Estimular la participación de los Agentes Promotores a nivel nacional de las diversas Afores registradas en el SAR (Sistema de Ahorro para el Retiro).
- Implementar programas de capacitación para los Agentes Promotores.
- Apoyar a las Afores participantes para fortalecer el SAR.
- Establecer nuevos mecanismos que permitan controlar el status de los Agentes Promotores a nivel nacional.
- Implementar una relación directa entre los registros de los Agentes Promotores con el tiempo.

### IV.2.2.2 DISEÑO DE LOS REQUERIMIENTOS DE EXPLOTACIÓN DE INFORMACIÓN

Documentamos los requerimientos de información que serán satisfechos en el Data Warehouse.

El siguiente cuadro muestra todas las combinaciones de cruces de información que se requieren.

Renglón o Columna		Columna o Renglón
Tiempo (Jerarquía)	VS	Tiempo (Jerarquía)
Región (Jerarquía)		Región (Jerarquía)
Periodo		Periodo
Año		Año
País		País
Región		Región
Estado		Estado
Afore		Afore
Estatus		Estatus

Es importante mencionar que el renglón y la columna no deben ser iguales.



### IV.3 DISEÑO DE LOS PROCESOS DEL DATA WAREHOUSE

#### IV.3.1 DISEÑO DEL PROCESO DE EXTRACCIÓN

Para realizar los procesos de extracción de los datos se requieren de librerías de acceso. A continuación se muestra una tabla con la información para el diseño de las librerías necesarias.

Librerías de acceso:

Nombre de la librería	Engine	Ubicación	Habilitar al inicio de sesión	User	Password	Database	Server
Sybase	Sybase		si	sa	supervmc	supervision	desarrollo
Agentes	Default	//opt/SAS/ sasuser/dataset	si				
sashelp	Default	//opt/SAS/ sasuser/dataset	si				

Diseñamos el proceso de extracción para cada una de las fuentes de datos para cargar el Data Warehouse.

<b>Procedimiento</b>	Proc sql
<b>Ubicación</b>	//opt/SAS/sasuser/Data_Sets
<b>Librería de acceso</b>	Sybase
<b>Librería de almacén</b>	Agentes
<b>Data Set</b>	Afore Estado AP_Seguimiento AP_Agentes_Promotores AP_Estatus

A continuación mostramos un ejemplo para un proceso de extracción.

Data Set	Afore
<pre>proc sql;   connect to Sybase     (user=sa;      password=supervmc      server=desarrollo      database=supervision);   create table Agentes.Afore as   select *   from connection to sybase   (     select * from Afore   );   disconnect from sybase; run;</pre>	



### IV.3.2 DISEÑO DEL PROCESO DE TRANSFORMACIÓN

Diseñamos los procesos y métodos usados en el proceso de carga del Data Warehouse por medio de los cuales las fuentes de datos son transformadas al formato apropiado definido en el modelo físico de datos

<b>Procedimiento</b>	Proc data
<b>Data Set</b>	AP_SEGUIMIENTO1 AP_AGENTE_PROMOTOR ESTADO
<b>Ubicación</b>	//opt/SAS/sasuser/Data_Sets
<b>Librería de almacén</b>	Agentes

A continuación presentamos un ejemplo del proceso de transformación de la fecha de seguimiento al formato adecuado para el Data Warehouse.

<b>Data Set</b>	AP_SEGUIMIENTO1
<b>data</b>	<pre> Agentes.AP_SEGUIMIENTO1; set Agentes.AP_SEGUIMIENTO; length Id_seguimiento \$6.; length Id_estatus \$2.; length Num_agente \$10.; length Anio \$4.;  fecha_evento=datpart(seguimiento_f_evento); format fecha_evento date9.; anio=year(fecha_evento);  run; </pre>

### IV.3.3 DISEÑO DEL PROCESO DE CARGA

Mostramos el proceso de carga del Data Warehouse, el cual consiste en vaciar los datos al Data Warehouse.

<b>Procedimiento</b>	Proc mddb
<b>Data Set</b>	Data_Set10
<b>Ubicación</b>	//opt/SAS/sasuser/Data_Sets
<b>Librería de almacén</b>	Agentes
<b>proc mddb</b>	<pre> data=sasuser.Data_Set10 out=SasHELP.agentes_consar label="Información del Estatus de los Agentes";  Class Periodo; Class Anio; Class Pais; Class Region; Class Estado; Class Afore; Class Estatus_Descripcion; Var Cantidad_Agente / sum; </pre>



```
HIERARCHY Período Anio
/name='Período'
Display=yes;
HIERARCHY País Region Estado Afore
Estatus_Descripcion /name='Region'
Display=yes;
run;
```

#### IV.3.4 REGISTRO DE METADATOS

Registramos los metadatos acerca del proceso de transformación en la herramienta de administración de metadatos.

<b>Repositorio de metadatos</b>	SASHELP repositorio
<b>Tabla</b>	Información de los Agentes Promotores
<b>Tabla Fuente</b>	Data_Set10
<b>Jerarquias</b>	Período Región
<b>Host</b>	Sun Enterprise 10000
<b>Label</b>	Información de los Agentes Promotores
<b>Lib</b>	Sashelp
<b>MDDB</b>	Agentes_Consar
<b>MTYPE</b>	MDDB
<b>Path</b>	//opt/SAS/sasuser/cubos

#### IV.4 DISEÑO DE LA SEGURIDAD

##### IV.4.1 DISEÑO DEL ESQUEMA DE SEGURIDAD DE DATOS

Para el modelo de seguridad del Data Warehouse utilizamos las medidas existentes en la COSAR que son las siguientes

##### Componentes reutilizables

Acceso a la red, restringido por:

- Nombre de usuario
- Contraseña

Acceso a las fuentes de datos, restringido por:

- Nombre de usuario
- Contraseña
- Tipo de usuario
- Tipo de acceso

Acceso a los recursos Data Warehouse, restringido por:

- Nombre de usuario



• Contraseña

Modelo de seguridad del Data Warehouse

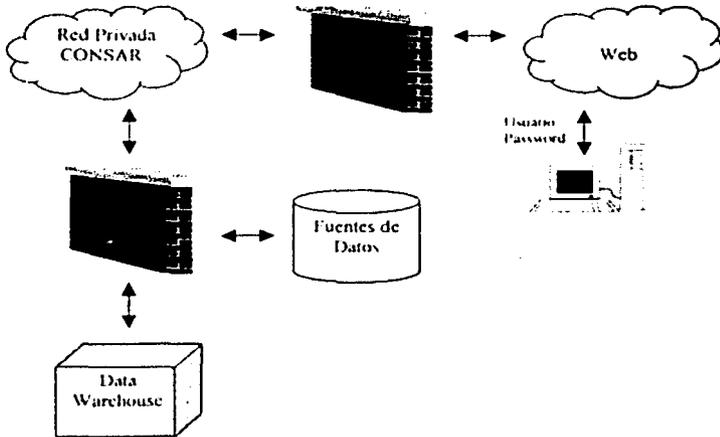


Figura IV.4.1  
Modelo de seguridad del Data Warehouse

IV.4.2 DISEÑO DE LA INTERFAZ DE SEGURIDAD

Para la interfaz de seguridad del Data Warehouse diseñamos las medidas siguientes:  
Componentes necesarios para la seguridad.

<b>Nombre del programa para la seguridad</b>	Crearsesion sas	
<b>Ubicación</b>	//opt/SAS/sasuser/progs	
<b>Procedimientos utilizados</b>	Proc sql	Extrae los datos del usuario
	Data _null	Ejecuta la macro %salida
	Macro %salida	Valida al usuario, crea código html
	Acceso html	Ejecuta el programa crear sesion sas
	Indicadores html	Ejecuta la interfaz de explotación
<b>Librería</b>	Sashelp	
<b>Dataset</b>	Users	



<b>Campos</b>	Folio Contraseña Indadm
---------------	-------------------------------

A continuación presentamos la pantalla de acceso y seguridad

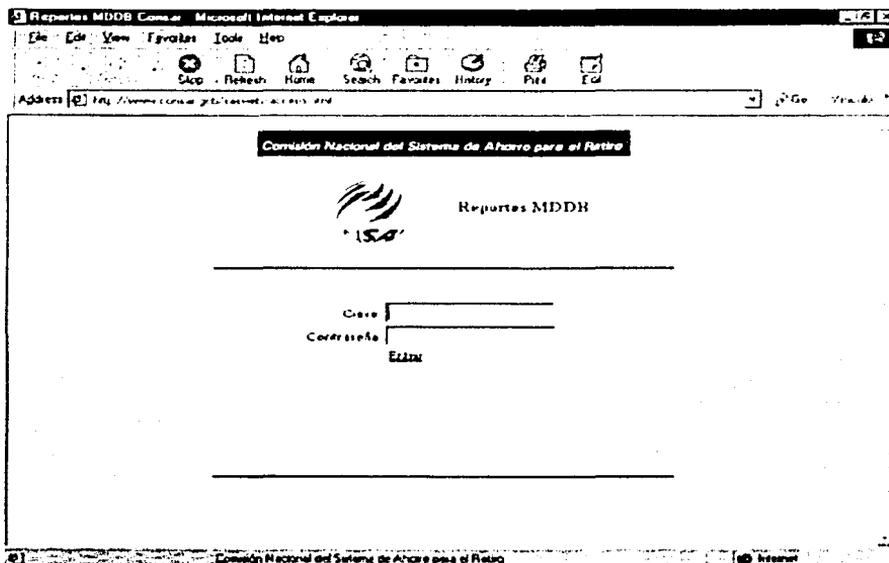


Figura IV.4.2  
Interfaz de seguridad

## IV.5 DISEÑO DE LA EXPLOTACIÓN DEL DATA WAREHOUSE

### IV.5.1 DISEÑO DE LA INTERFAZ DE EXPLOTACIÓN

Para diseñar la interfaz de explotación del Data Warehouse utilizamos el Reporteador de SAS llamado Report Viewer 8.2.

Para esto configuramos los siguientes parámetros.

<b>Servidor de Web</b>	Servidor de Red Apache
<b>Servidor de Aplicaciones</b>	Appstart
<b>allocate library</b>	//opt/SAS/sasuser
<b>Programa</b>	//opt/SAS/sasuser/programs
<b>allocate file</b>	//opt/SAS/sasuser/programs
<b>proglibs</b>	//opt/SAS/sasuser/programs



## IV 5.2 DISEÑO DE LA INTERFAZ DE APLICACIÓN

Para diseñar la interfaz de aplicación creamos la siguiente página y configuramos la siguiente información:

Nombre de la Página	Sasweb.html
Librería	Trnplib
Ruta	//opt/SAS/sasuser
Dataset	Agentes_Consar
Metabase	SASHELP.MBEIS

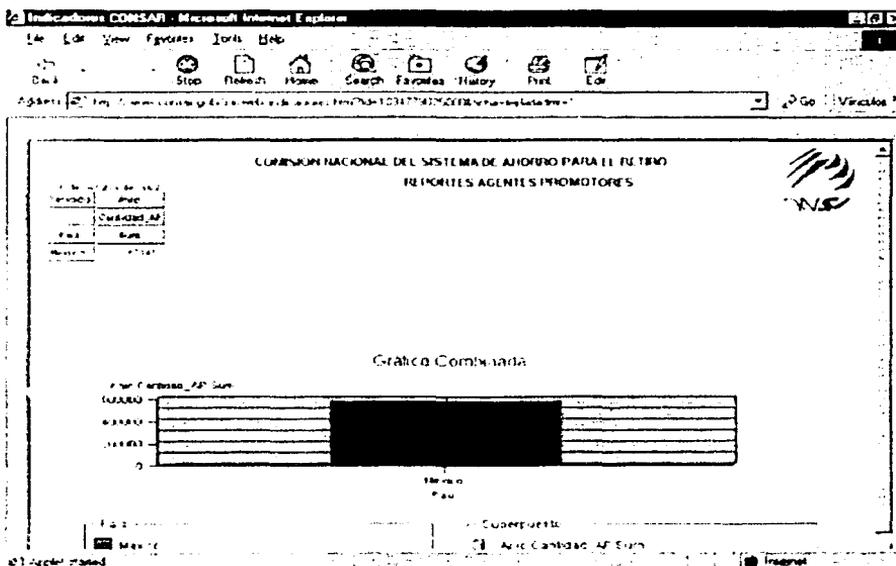


Figura IV 4 3  
Interfaz de aplicación

TESIS CON  
FALLA DE ORIGEN



# **CAPÍTULO V**

## **PRUEBAS FINALES**

**CONTENIDO:**

**V.1 PRUEBAS AL SISTEMA**

**V.1.1 EJECUTAR PRUEBAS AL SISTEMA**

**V.1.2 VERIFICAR LA EJECUCIÓN DEL DATA  
WAREHOUSE**



## CAPÍTULO V

### PRUEBAS FINALES

#### V.1 PRUEBAS AL SISTEMA

Asegurar la calidad del Data Warehouse para garantizar que cumpla con los requerimientos funcionales de acuerdo al documento de especificaciones de diseño.

##### V.1.1 EJECUTAR PRUEBAS AL SISTEMA

Nombre y función de los componentes

Describimos los componentes probados, así como su funcionalidad

Componente	Funcionalidad
Extracción del Data Warehouse	Extracción de las fuentes de datos de la Base de Datos
Transformación del Data Warehouse	Proceso mediante el cual los datos de las Fuentes de Datos son transformados al formato definido en el Modelo Físico de Datos
Carga del Data Warehouse	Proceso que consiste en vaciar los datos de la fuente de datos al Data Warehouse
Registro de Metadatos	Almacenamiento de los Metadatos
Seguridad del Data Warehouse	Seguridad del Data Warehouse
Interfaz de Explotación del Data Warehouse	Funcionalidad del Report Viewer
Interfaz de Aplicación del Data Warehouse	Interfaz de aplicación en la web

Propósito y alcance de las pruebas.

Propósito	Asegurar la calidad de la funcionalidad de los componentes del Sistema del Data Warehouse
Alcance	Identificar las fallas que permitan corregir la funcionalidad de los componentes del Sistema del Data Warehouse

Lista de Pruebas

Componentes	Característica probada	Satisfactoria SI/No
Extracción del Data Warehouse	Acceso y extracción de las Fuentes de Datos	SI
Transformación del Data Warehouse	Transformación de los Datos al formato apropiado	SI
Carga del Data Warehouse	Que los Datos sean almacenados en el Data Warehouse	SI
Registro de Matadatos	Registro de los datos en la Metabase del Sistema	SI



Componentes	Característica probada	Satisfactoria Si/No
Seguridad del Data Warehouse	Que los datos estén resguardados de manera segura y confiable	SI
Explotación del Data Warehouse	Que la configuración del Reporteador Report Viewer v 8.2 opere de manera rápida, segura y confiable	SI
Interfaz de Aplicación del Data Warehouse	Que la interfaz de aplicación opere de manera segura y confiable	SI

Excepciones en los Resultados de las Pruebas

Como resultado de las pruebas encontramos que para cada componente no se presentó ninguna falla en las actividades realizadas.

Característica probada	Componente	Comentario	Acción
No se presentaron fallas			

Conclusiones

Describimos el status de las pruebas para cada característica de los componentes

Característica probada	Status (completo, parcial, o no aceptado)
Conexión a la Fuente de Datos	COMPLETO
Extracción de los Datos Fuente	COMPLETO
Transformación de los Datos Fuente	COMPLETO
Carga de los Datos Transformados	COMPLETO
Configuración de la Metabase	COMPLETO
Seguridad del Data Warehouse	COMPLETO
Interfaz de Explotación	COMPLETO
Interfaz de Aplicación	COMPLETO

### V.1.2 VERIFICAR LA EJECUCIÓN DEL DATA WAREHOUSE

Verificar que el Data Warehouse cumple con los requerimientos de ejecución como se especificó en el documento de especificaciones de diseño

Nombre y función de los componentes

Describimos los componentes a ser probados, así como su funcionalidad



Componente	Funcionalidad
Cruce de Información	Selecciona el tipo de cruce de información
Estadísticos Aplicables	Estadísticos aplicables a la variables de análisis
Opciones de Gráficos	Opciones de gráficos para ilustrar el tipo de reporte seleccionado
Exportar los Datos a Excel	Exporta los datos a Excel

## Propósito y alcance

Propósito	Verificar que el Data Warehouse cumple con los requerimientos de ejecución como se estableció en el documento de especificaciones de diseño
Alcance	Valorar la información arrojada por el Data Warehouse

## Lista de Pruebas

Componentes	Característica probada	Satisfactoria SI/No
Cruce de Información	Probamos los diferentes tipos de reportes para los diferentes tipos de cruce de información establecidos en el documento de requerimientos	SI
Estadísticos Aplicables	Probamos los estadísticos establecidos en los documentos de requerimientos configurados en el Reporteador de SAS Report Viewer v 8 2	SI
Opciones de Gráficos	Probamos las Opciones de Gráficos establecidas en el documento de requerimientos configurados en el Reporteador de SAS Report Viewer v 8 2	SI
Exportar los Datos a Excel	Probamos la opción para exportar los Datos del Data Warehouse a Excel configurados en el Reporteador de SAS Report Viewer v 8 2	SI

## Excepciones en los Resultados de las Pruebas

Como resultado de las pruebas encontramos que para cada componente no se presentó ninguna falla en las actividades realizadas



Característica probada	Componente	Comentario	Acción
No se presentaron fallas			

### Conclusiones

Describimos el status de las pruebas para cada característica de los componentes.

Característica probada	Status (completo, parcial, o aceptado)
Cruce de Información	ACEPTADO
Estadísticos Aplicables	ACEPTADO
Opciones de Gráficos	ACEPTADO
Exportar los Datos a Excel	ACEPTADO



# CAPÍTULO VI

# REVISIÓN

## CONTENIDO:

### VI.1 REVISIÓN DEL IMPACTO EN LA ORGANIZACIÓN



---

## CAPÍTULO VI REVISIÓN

### VI.1 REVISIÓN DEL IMPACTO EN LA ORGANIZACIÓN

El proyecto propuesto en la presente tesis consiste en una solución integral con base en tecnologías de vanguardia como es **The SAS System™** que permitirá apoyar a los directivos de la CONSAR (Comisión Nacional del Ahorro para el Retiro) en la toma de decisiones sobre el control y monitoreo de los Agentes Promotores a nivel nacional que participan en el Sistema de Ahorro para el Retiro.

Este proyecto sigue la metodología **SAS® Rapid Warehousing**, para la creación de un Data Warehouse, diseñado específicamente en apoyo de decisiones, proporcionando integración y acceso fácil a los datos de los Agentes Promotores de la CONSAR desde cualquier explorador Web, además de que esta información la podremos visualizar en un reporte o bien descargarla a una hoja de cálculo (Excel).

Hoy en día la CONSAR necesita depositar mucha confianza en la toma de decisiones acerca de todos sus asuntos referentes al SAR, y para tomar dichas decisiones se requiere no sólo de hechos y cifras, ya que sabemos que las exigencias crecen en todo momento, y entonces las decisiones que debemos tomar en la CONSAR deben ser más aceleradas; pero que pasa si tenemos una montaña de información la cual debe ser analizada, lógicamente lo primero que pensaríamos es que vamos a requerir de mucho tiempo.

Por ese motivo se requieren herramientas que nos ayuden a minimizar el tiempo para analizar mucha información con mayor velocidad y precisión; y utilizando las herramientas de **The SAS System™** logramos mantener una ventaja competitiva, ya que los procesos empresariales deben reaccionar ante los cambios del mercado. De otro modo el mercado cada vez más globalizado, la presión inmensa de los cambios y exigencias sociales, así como los arranques tecnológicos, se convierten en factores que debilitarían a esta empresa. El adquirir este tipo de herramientas nos muestra que la CONSAR invierte en tecnología y soluciones con las cuales se mantiene a la vanguardia en este mundo cada vez más competitivo, ya que ahora las empresas no dependen tan sólo de factores tales como ubicación, productos, etc., sino también del conocimiento. Tal conocimiento es la metodología **SAS® Rapid Warehousing**, ya que está basada en técnicas comprensibles, detalladas y relevantes que son cruciales para lograr y sostener una ventaja competitiva. El poseer este tipo de conocimientos significa tener respuestas correctas y realizar decisiones estratégicas para la ejecución de la empresa.



Con la creación de este proyecto se consiguieron una serie de beneficios que permiten a los directivos de la CONSAR apoyarse en él para la toma de decisiones estratégicas con respecto a los Agentes Promotores a nivel nacional.

Sin lugar a dudas, uno de los beneficios más palpables es el de poder visualizar la información de dichos Agentes de manera sumariada o resumida a través de varios campos, lo cual ahorra tiempo a la hora de tomar alguna decisión respecto a los Agentes Promotores.

Otra de las grandes ventajas es, como se mencionó anteriormente, el poder ver esta información a través de cualquier navegador de Internet, cualidad que hace que este sistema esté a la vanguardia en cuanto a las necesidades de información de la actualidad, brindando accesibilidad desde cualquier ubicación, y claro está que, esta accesibilidad va de la mano con un sistema de seguridad, para evitar el acceso no autorizado a consultar la información contenida en el sistema.

Sin lugar a dudas con la realización de este proyecto, y al utilizar tecnología de punta como es el caso de las herramientas de **The SAS System™** se brinda a la CONSAR un importante avance en cuanto a las nuevas tecnologías de información como lo es el Data Warehouse, además de brindar un avance tecnológico más, que debe ser en un principio imitado por varias empresas que requieran la rapidez y confiabilidad de sistemas de este tipo, y después tratar de que éste sea superado para continuar con el gran avance en cuanto a las tecnologías de información.

## IMPACTOS DEL DATA WAREHOUSE

El éxito del Data Warehouse implementado en la CONSAR no está sólo en su construcción, sino en el uso que se le dé para mejorar todos aquellos procesos, operaciones y decisiones de la empresa. Posicionar un Data Warehouse para que sea usado efectivamente, requiere entender los impactos de implementación en los siguientes ámbitos:

### 1.- Impactos Humanos.

#### - Efectos sobre el personal de la CONSAR

##### ➤ Construcción del Data Warehouse

Construir un Data Warehouse requiere la participación activa de quienes lo usarán. A diferencia del desarrollo de aplicaciones tradicionales, donde los requerimientos de la empresa logran ser relativamente bien definidos, producto de la estabilidad de los procesos empresariales a través del tiempo, el construir este Data Warehouse para los Agentes Promotores de la CONSAR dependió en gran medida de la realidad de la empresa, así como de las condiciones existentes, las



cuales determinaron qué es lo que debe contener el Data Warehouse. El personal de la CONSAR participó activamente durante el desarrollo del Data Warehouse, desde una perspectiva de construcción y creación.

➤ **Accesando al Data Warehouse**

Este Data Warehouse intenta proveer los datos que dan la posibilidad a los usuarios de acceder a la información cuando ellos la necesiten. Esta aproximación al entregar información tiene varias implicaciones:

- a) Que el personal de la CONSAR puede necesitar aprender las nuevas funciones que trae consigo la implementación del Data Warehouse.
- b) Los análisis extensos y las demoras en la programación para obtener la información deberán ser eliminadas. Como la información estará lista para ser accesada desde cualquier ubicación via un navegador de Internet las expectativas de que este fin se lleve a cabo probablemente aumentarán.
- c) Las nuevas oportunidades que pueden existir en la CONSAR para los especialistas existentes en esta área de información, y para los que no lo sean.
- d) La gran cantidad de reportes en papel serán reducidas o eliminadas.
- e) La madurez del Data Warehouse dependerá del uso activo y la retroalimentación de sus usuarios.

**2.- Impactos Empresariales**

➤ **Procesos y Decisiones Empresariales**

Se ponen a consideración los beneficios empresariales potenciales en la CONSAR de los siguientes impactos:

- a) Los Procesos de Toma de Decisiones pueden ser mejorados mediante la disponibilidad de la información. Decisiones empresariales se hacen más rápidas por personal más informado, que puede ser realizado mediante cursos de capacitación.
- b) Los procesos de la CONSAR relacionados con los Agentes Promotores pueden ser optimizados. El tiempo perdido esperando por información que finalmente es incorrecta o no encontrada, debe ser eliminado.



c) Las conexiones y dependencias entre estos procesos se vuelven más claros y entendibles. Las secuencias de estos mismos pueden ser optimizados para ganar eficiencia y reducir costos.

d) Procesos y datos de los sistemas operacionales, así como los datos en el Data Warehouse son usados y examinados. Cuando los datos son organizados y estructurados para tener un significado que se relacione de manera más directa con la actividad de cada área, hace que el personal aprenda mucho de los sistemas de información. Pueden quedar expuestos posibles defectos en las aplicaciones actuales, siendo posible entonces mejorar la calidad de nuevas aplicaciones.

➤ **Comunicación e Impactos Organizacionales**

Apenas el Data Warehouse comience a ser fuente primaria de la información empresarial consistente, los siguientes impactos pueden comenzar a presentarse:

- a) El personal tiene mayor confianza en las decisiones empresariales que se toman. Ambos, quienes toman las decisiones como los afectados saben que está basada en información confiable.
- b) La CONSAR y el personal del cual se compone quedan determinados por el acceso a la información. De esta manera, el personal queda mejor habilitado para entender su propio rol y responsabilidades, así como también los efectos de sus contribuciones; y a la vez, desarrollan un mejor entendimiento con las contribuciones de otros.
- c) La información compartida conduce a un lenguaje común, conocimiento común y mejoramiento de la comunicación en la CONSAR. Se mejora la confianza y cooperación entre distintos sectores de la CONSAR, viéndose reducida la sectorización de funciones.

Visibilidad, accesibilidad, y conocimiento de los datos producen mayor confianza en los sistemas operacionales.

**3.- Impactos Técnicos del Data Warehouse.**

Considerando las etapas de construcción, soporte del Data Warehouse y soporte de sistemas operacionales, se tienen los siguientes impactos técnicos:

➤ **Nuevas destrezas de desarrollo**

Cuando se construyó el Data Warehouse, el impacto más grande sobre el personal técnico está dado por la curva de aprendizaje ya que muchas habilidades nuevas se deben aprender, incluyendo:



a) Conceptos y estructura del Data Warehouse.

b) El Data Warehouse introduce muchas tecnologías nuevas (ETT, Carga, Acceso de Datos, Catálogo de Metadatos, Implementación de DSS/EIS), y cambia la manera en que se utiliza la tecnología existente. Nuevas responsabilidades de soporte, nuevas demandas de recursos y nuevas expectativas, son los efectos de estos cambios.

c) Habilidades de diseño y análisis donde los requerimientos empresariales no son posibles de definir de una forma estable a través del tiempo.

d) Técnicas de desarrollo incremental y evolutivo

e) Trabajo en equipo con personal de diversas áreas de la CONSAR como participantes activos en el desarrollo del proyecto.

➤ Nuevas responsabilidades de operación

Cambios sobre los sistemas y datos operacionales deben ser examinados más cuidadosamente para determinar el impacto que estos cambios tienen sobre ellos, y sobre el Data Warehouse.



# CONCLUSIONES



## CONCLUSIONES

El proyecto desarrollado en la presente tesis consistió en una solución integral basada en tecnologías de punta que permitirá apoyar a los directivos de la CONSAR en la toma de decisiones sobre el control y monitoreo de los Agentes Promotores a nivel Nacional que participan en el Sistema de Ahorro para el Retiro.

Siguiendo la metodología desarrollada en esta tesis, observamos que los sistemas tradicionales y las aplicaciones de Data Warehousing son polos opuestos en cuanto a sus requerimientos de diseño y sus características de operación; y no debemos perder de vista este detalle, ya que es de suma importancia comprender perfectamente esta diferencia para evitar caer en el diseño de un Data Warehouse como si fuera una aplicación tradicional.

Las aplicaciones tradicionales están organizadas para ejecutar las transacciones para los cuales fueron hechas, como por ejemplo: mover dinero entre cuentas, un cargo o abono, una devolución de inventario, etc. Por otro lado, un Data Warehouse está organizado en base a conceptos, como por ejemplo: clientes, facturas, productos, etc

Otra diferencia radica en el número de usuarios. Normalmente el número de usuarios de un Data Warehouse es menor al de un sistema tradicional. Es común encontrar que estos sistemas son accedidos por cientos de usuarios simultáneamente, mientras que los Data Warehouse sólo por decenas. Otro factor es que frecuentemente los sistemas tradicionales son menores en tamaño que los Data Warehouse, esto es debido a que un Data Warehouse puede estar formado por información de varios de estos sistemas

Existe también una gran diferencia en el diseño, ya que mientras que un sistema tradicional es extremadamente normalizado, un Data Warehouse tiende a ser denormalizado, además de que estos sistemas normalmente están formados por un número mayor de tablas, cada una con pocas columnas, mientras que en un Data Warehouse el número de tablas es menor, pero cada una de éstas tiende a ser mayor en el número de columnas.

Otra gran diferencia la encontramos en la estructura, ya que mientras las estructuras de los sistemas tradicionales son muy estables, y rara vez cambian, la estructura de un Data Warehouse sufre cambios constantes derivados de su evolución. Esto se debe a que los tipos de consultas a los cuales están sujetos, son muy variados y es imposible preverlos todos de antemano

En esta tabla podemos visualizar mejor estas diferencias

ESTA TESIS NO SALE  
DE LA BIBLIOTECA



<b>Sistemas Tradicionales</b>	<b>Data Warehouse</b>
Orientada a transacciones	Orientada a conceptos
Detallada	Sumarizada
Actualizada en línea	Representa valores a un tiempo
Usuarios a nivel operativo	Usuarios a nivel gerencial
Corre en base a repeticiones	Corre heurísticamente
Muy sensitivo al desempeño	Poco sensitivo al desempeño
Accesa unidades a la vez	Accesa conjuntos de unidades a la vez
Orientado a una operación	Orientado al análisis
Estructura estática	Estructura flexible
Sin redundancia	Con mucha redundancia
Alta probabilidad de acceso	Modesta probabilidad de acceso
Administrada como un todo	Administrada por partes
Información bruta (datos)	Información procesada (información)
Actualizada en línea	Actualizada en batch
Muchas tablas con pocas columnas	Pocas tablas con muchas columnas

Hay que resaltar de manera muy importante que los objetivos de los sistemas tradicionales y el Data Warehouse para la toma de decisiones son muy diferentes: tratar de diseñar un Data Warehouse pensando en un sistema operacional es un fracaso. El éxito en la implementación de una arquitectura de Data Warehouse en las empresas radica en parte en el éxito del diseño de los sistemas operacionales, ya que éstos son los proveedores de los datos y los que se deben de adaptar de manera rápida y flexible a los cambios empresariales. Se deben tener ideas muy claras de lo siguiente: ¿qué datos utilizar?, ¿cómo se deben transformar?, ¿cómo se deben transferir, almacenar y organizar? y finalmente, ¿cómo se deben acceder y analizar?

En la CON SAR era innegable que los sistemas de información tradicionales construidos utilizando manejadores de bases de datos relacionados, rendían los resultados requeridos, ya que para cumplir con demandas de información actualizada, de proyecciones en el tiempo, de análisis comparativos entre regiones en diferentes periodos de tiempo, los desarrolladores utilizaban herramientas y manejadores de bases de datos que tenían el alcance, y esto es debido a que es una tecnología madura que provee las facilidades necesarias, sin salirse del concepto seguido por tanto tiempo en cuanto al desarrollo y diseño de estos sistemas.

Sin embargo el introducir en la CON SAR la metodología de Data Warehouse implicó el concepto de valor de la información. Datos sobre los mismos procesos realizados en la CON SAR son útiles cuando el tiempo es un factor en contra. Las personas que toman las decisiones claves necesitan un acceso rápido a datos actuales aportados por diferentes sistemas existentes y bases de datos que



operan tanto de forma interna como externa para tomar decisiones claves sobre el rumbo de las operaciones de la CONSAR.

Y para que estas decisiones se lleven a buen fin son factor clave los datos históricos de esta empresa, ya que debido a que los Data Warehouse tienen el potencial de convertir esta información en conocimientos, y la CONSAR al contar con información profunda de sus operaciones tiene el potencial de convertirlos en beneficios. Este conocimiento, extraído de los datos históricos de sus operaciones, es único para esta empresa y, por lo tanto, se puede emplear para crear elementos que marquen una diferencia, tales como servicios adecuados para cada segmento del mercado

Tan cierto es esto que quizá la razón más importante de construir un Data Warehouse para la CONSAR sea lo inadecuado de los sistemas actuales para las exigencias del mundo moderno. Muchos sistemas con los que se cuentan no satisfacen este tipo de exigencias, ya que por lo regular, los datos son inaccesibles e inconsistentes, tanto en forma como en significado.

Otro de los objetivos principales del Data Warehouse implementado en la CONSAR es servir como base de información para la toma de decisiones. Los beneficios obtenidos por la utilización de este sistema se basan en el acceso interactivo e inmediato a información estratégica de un área de la CONSAR (en este caso los Agentes Promotores). Este acercamiento de la información al usuario final permite una toma de decisiones rápida y basada en datos objetivos obtenidos a partir de las bases de datos de la empresa. Estos beneficios aumentan cuanto más importantes son las decisiones a tomar y cuanto más crítico es el factor tiempo

Esta necesidad de obtener información para una amplia variedad de usuarios es la principal razón que conduce al concepto de Data Warehouse. El énfasis no está sólo en llevar la información hacia lo alto sino que a través de la institución, para que todos los empleados que la necesiten la tengan a su disposición.

Con el Data Warehouse implementado podemos realizar las siguientes tareas:

Consultas más fáciles
Facilita la toma de decisiones
Práctica de negocios inteligentes
Obtener información valiosa y oportuna
Impacto positivo sobre los procesos de la empresa
Usuarios mejor informados

Adicionalmente nos va a proporcionar los siguientes beneficios en las siguientes áreas:



#### a) Usuarios

- **Confianza en el uso de sistemas:** El incremento en la rapidez de las consultas afianzará el uso de los sistemas operativos de la CONSAR.
- **Reducción en tiempos de consultas:** Es una solución que reduce tiempos de espera por parte de los usuarios.
- **Apoyo a la toma de decisiones:** Los ejecutivos de la CONSAR tendrán a la mano los resúmenes de información para armar reportes comparativos, utilizando los indicadores más comunes de la empresa, para apoyarse en la toma de decisiones.

#### b) Sistemas

- **Facilidad de uso:** La facilidad de uso de esta herramienta permitirá tener un DW para otra área específica en muy poco tiempo.
- **Acceso a cualquier base de datos:** el soporte para acceder múltiples bases de datos amplía las áreas de aplicación de esta herramienta, para generar resúmenes operativos de su información.
- **Mayor rendimiento:** el rendimiento de los servidores de la CONSAR se verá altamente beneficiado, debido a que el tiempo de consulta se verá disminuido.

Además de obtener los beneficios mencionados anteriormente, encontramos estos otros de carácter más general:

- Permite información en línea (online) sin ocupar un documento en papel, lo que reduce además de costo, el espacio que ocupa un documento físico.
- La información está disponible de forma rápida y precisa, únicamente al abrir un explorador de Web.
- Acceso directo, fácil y económico a todos los datos de la CONSAR relacionados con los Agentes Promotores.
- Poseer una base de datos consistente en toda la empresa.
- Visualizar en varias formas los resultados de sus consultas (gráficas, reportes, etc.)
- Identificar y analizar las tendencias en el tiempo más breve.
- Capacidad para utilizar una variedad de herramientas para procesar y resumir los datos, y así finalmente reagrupar toda la información.
- Capacidad para explorar e identificar relaciones de datos a través de múltiples bases de datos.
- Los datos históricos aumentan su valor a lo largo del tiempo.
- Crecimiento escalable.



- Los logros son evidentes: mayor productividad por decisiones correctas en un tiempo más corto.

Por último se puede decir que el proyecto Data Warehousing implementado para la CONSAR se puede considerar exitoso, cuando su objetivo final comience a concretarse, es decir que el personal de la CONSAR use el Data Warehouse para satisfacer sus necesidades de información más efectiva, ya que como hemos comentado, son variados los cambios que comenzarán a producirse al implementar un Data Warehouse, y es importante entonces anticiparse a estos cambios, considerar sus implicaciones y planificarlas en la empresa. Las siguientes situaciones, pueden advertirnos el inicio de estos cambios:

- El personal de la CONSAR depende del Data Warehouse como un recurso primario de información
  - El personal de la CONSAR se vuelve menos dependiente de los sistemas operacionales y de sus bases de datos para sus necesidades de información.
  - Se ve reducida o eliminada la demanda por programación especializada para encontrar la información necesaria
  - Los usuarios y uso del Data Warehouse crecen, con un correspondiente incremento en la demanda de soporte
  - La complejidad de cambios en los sistemas operacionales se incrementa, y su efecto sobre el Data Warehouse debe ser considerado.
  - La finalidad del Data Warehouse consiste en asistir al personal que toma las decisiones a comprender el pasado y planear el futuro. Aunque el Data Warehouse contiene información estratégica crucial, extraer un valor empresarial del Data Warehouse es un esfuerzo complejo. Obtener este valor requiere de una mezcla de aptitudes, habilidad técnica, intuición y experiencia
  - El Data Warehouse puede producir la información correcta para los ejecutivos, gerentes y analistas de la CONSAR. Conectarse al potencial del Data Warehouse sólo está limitado por la capacidad, habilidad y creatividad de los usuarios
- Las experiencias anteriores apuntan hacia un amplio rango de usos: análisis de rentabilidad y crecimiento, administración estratégica, etc



Esta serie de beneficios que mencionamos pueden servir como una recomendación a quienes decidan iniciar un modelo Data Warehouse en sus empresas, ya que el Data Warehouse se puede ver como una colección de datos relacionados al oficio o a la actividad profesional a la que se dedique la empresa, además de ser integrada y relacionada a la fecha en que se produjeron estos datos para poder analizarlos periódicamente. El enfoque está en la disposición eficiente de estos datos, así como la ejecución de evaluaciones y análisis; podemos mencionar además que el Data Warehouse apoya a todos los ejecutivos que tienen que tomar decisiones con información que se obtiene de manera rápida y confiable.

Sin embargo, las aplicaciones del Data Warehouse son muy variadas y en diferentes campos de la industria, como por ejemplo en el área de Marketing donde un Data Warehouse nos puede ayudar en:

- Investigación comercial
- Segmentación de mercados
- Identificación de necesidades no cubiertas y generación de nuevos productos, o modificación de productos existentes
- Fijación de precios y descuentos
- Definición de la estrategia de canales de comercialización y distribución
- Definición de la estrategia de promoción y atención al cliente
- Relación con el cliente
- Lanzamiento de nuevos productos
- Campañas de ventas cruzadas, vinculación, etc
- Apoyo al canal de venta con información cualificada

Otra área en donde el uso del Data Warehouse se ha utilizado es en el Análisis de Riesgo Financiero, donde ofrece capacidades avanzadas de desarrollo de aplicaciones para dar soporte a las diversas actividades en la administración de los riesgos. Es posible desarrollar cualquier herramienta utilizando las funciones que incorpora la plataforma, gracias a la potencialidad estadística aplicada al riesgo de crédito

Así se puede usar para llevar a cabo las siguientes funcionalidades

- Para la gestión de la posición: determinación de la posición, cálculo de sensibilidades, simulaciones, monitorización, riesgos contra límites, etc
- Para la medición de riesgos: simulación de escenarios históricos, simulación de Montecarlo, modelos de valoración, análisis de rentabilidad, establecimiento y seguimiento de límites, etc



Y los logros al utilizar la tecnología Data Warehouse en cualquier área en donde se aplique son más que evidentes:

**"MAYOR PRODUCTIVIDAD POR DECISIONES CORRECTAS EN UN TIEMPO MÁS CORTO".**



# APÉNDICE A

## GLOSARIO



## GLOSARIO

### A

#### ACCESO A DATOS

Proceso mediante el cual se ingresa a las fuentes primarias de datos para posteriormente realizar la carga de estos mismos al Data Warehouse.

#### AFORE

Son instituciones financieras creadas exclusivamente para administrar los fondos de retiro y dar servicio a los trabajadores afiliados. Para ofrecer mejores beneficios, y mayores ganancias a los fondos de retiro de los trabajadores.

#### AGENTE PROMOTOR

Aquella persona física que realiza actividades de registro de cuentas individuales, así como de comercialización, promoción y atención de solicitudes de traspasos para los fondos de ahorro para el retiro.

#### ARCHIVO

Colección de registros de un tipo único o de tipos relacionados.

#### ARCHIVO PLANO

Matriz de columnas (campos) de datos, en el que cada fila representa un registro. Equivalente al término 'Tabla' o 'Relación' en una base de datos relacional.

### B

#### BASE DE DATOS MULTIDIMENSIONAL (MDDB)

Es un diseño que almacena registros en ejes de n-dimensiones, formando una matriz o cubo de dimensiones.

---



## **BASE DE DATOS RELACIONAL**

Es un conjunto de datos entre los que existe una correlación y que se almacenan con criterios independientes con respecto a los programas que los utilizan. Esta agrupación de datos se realiza teniendo en mente su fácil acceso, recuperación y actualización.

## **BUSSINESS INTELLIGENCE (INTELIGENCIA DEL NEGOCIO)**

La inteligencia de negocios esta expresada en reportes y graficas disponibles en workstations locales.

# **C**

## **CAMBIO DE ADQUISICIÓN DE DATOS**

Cambios en sumas, eliminaciones, o actualizaciones para la fuente de datos desde la última carga.

## **CARGA DE DATOS**

Proceso en el cual se alimenta al Data Warehouse de los datos extraídos de las fuentes.

## **CICLO DE REFRESCO**

Es cuando en cada intervalo fijo el Dataware físico es remplazado en su integridad.

## **CONSTRAINT**

Es una restricción que afecta a una columna o a una combinación de columnas de la tabla que estamos definiendo.



---

## **CONSULTA**

Herramienta capaz de sintetizar datos procedentes de varias tablas de bases de datos relacionales, mostrando estos datos de una forma concreta.

## **D**

### **DATA MART (MERCADO DE DATOS)**

Deposito que contiene únicamente datos.

### **DATA WAREHOUSE**

Es un almacén de información de asuntos orientados diseñado específicamente en apoyo de decisiones (inteligencia de negocios), proporcionando integración y acceso fácil a los datos de una organización.

### **DATA WAREHOUSE DISTRIBUIDO**

Es un Data Warehouse implementado en un ambiente multiplataforma.

### **DATOS**

Cualquier información que pueda ser usada para cálculo, comparación u otro procesamiento o que requiera ser recordada para un uso futuro. Algunas veces, se usa para referirse a registros u otra información involucrada en un programa a diferencia del programa mismo

### **DATOS SUMARIZADOS**

Son elementos de datos requeridos para las consultas más frecuentes que se guardan en forma pre-resumida.

### **DENORMALIZACIÓN**

Es un proceso mediante el cual se liberan a los datos de su estructura OLTP. Este proceso busca optimizar la ejecución del Data Warehouse, y frecuentemente resulta en algo de redundancia en los datos.



## **DEPURACIÓN**

Limpieza de datos grabando o removiendo valores inválidos, duplicados, fuera de rango, perdidos o duplicados.

## **DIMENSIÓN**

Las dimensiones son valores que categorizan a otros valores.

## **DOMINIO**

Sistema de denominación de Hosts en Internet. Los dominios van separados por un punto y jerárquicamente están organizados de derecha a izquierda.

# **E**

## **EIS**

Sistemas de información empresarial.

## **ENTIDAD-RELACIÓN (ER)**

Tradicionalmente, los modelos entidad relación han sido usados para diseñar la mayoría de los sistemas OLTP; es altamente normalizado, es decir tiene redundancia de información pequeña o nula, además de ser optimizado para cambios rápidos de actualización y transacción.

## **ESQUEMA COPO DE NIEVE**

El esquema copo de nieve es una versión mas normalizada del esquema de estrella, con la adición de tablas de dimensiones normalizadas

## **ESQUEMA DE ESTRELLA**

El esquema de estrella esta basado en una tabla central de facts rodeada por varias tablas de dimensiones.



### **EXPLORADOR WEB**

Un explorador Web contiene el software básico que necesita para encontrar, recuperar, ver y enviar información en Internet.

### **EXPLOTACIÓN**

Proceso mediante el cual se visualiza la información contenida en el Data Warehouse.

### **EXTRACCIÓN DE DATOS**

Proceso mediante el cual se extraen datos de las fuentes respectivas.

### **ESTRUCTURACIÓN**

Etapas donde se crean nuevos (derivados) valores de los datos y se modifican los valores ya existentes

## **F**

### **FACT**

Los facts son valores que contabilizan personas, lugares y cosas.

### **FOREIGN KEY (FK)**

Clave alterna que identifica en forma única un archivo o registro.

### **FOUNDATION LAYER**

Capa del Data Warehouse que tiene el máximo nivel de detalle.

### **FUENTE DE DATOS**

Sistemas que contienen una variedad de estructuras de archivos que van a servir para alimentar al Data Warehouse



## **G**

### **GRANULARIDAD**

La granularidad (o grano) describe el nivel de detalle almacenado en el Data Warehouse físicamente.

## **H**

### **HEURISTICO**

Realizar alguna acción de forma intuitiva o empírica, es decir sin una metodología.

### **HOST**

Anfitrión. Computadora conectada a Internet.

## **I**

### **IMPACTO EMPRESARIAL**

Resultados de la implementación del Data Warehouse a nivel corporativo.

### **IMPACTO HUMANO**

Resultados de la implementación del Data Warehouse a nivel de usuario.

### **IMPACTO TECNICO**

Resultados de la implementación del Data Warehouse a nivel infraestructura.

---



### **INFO MART**

Deposito que contiene reportes, graficas o programas

### **INFORMATION TECHNOLOGY (IT)**

Técnica utilizada para tratar la información.

### **INTEGRACION**

Conciliar diversos atributos de los datos así como sus valores, en una forma estándar y consistente.

### **INTELIGENCIA DEL NEGOCIO**

Bussiness Intelligence.

### **INTERFAZ**

Se denomina así a todo aquel medio físico que conecta un dispositivo periférico con la computadora; también se le conoce así a todo el software que comunica al usuario con la misma.

### **INTERNET PROTOCOL (IP)**

Protocolo Internet. Es un protocolo de bajo nivel para redes que describe la manera cómo el usuario puede comunicarse en Internet.

## **K**

### **KEY**

Clave principal que identifica en forma única un archivo o registro.



## L

### LAN

Local Area Network. Red de Área Local. Una red de área local es un sistema de comunicación de alta velocidad de transmisión. Estos sistemas están diseñados para permitir la comunicación y transmisión de datos entre estaciones de trabajo inteligentes. Todas las PCs, conectadas a una red local, pueden enviar y recibir información. Como su mismo nombre lo indica, una red local es un sistema que cubre distancias cortas. Una red local se limita a una planta o un edificio

## M

### MAINFRAME

Término inglés empleado para designar computadoras de grandes dimensiones.

### MDDB

Base de datos multidimensional.

### MERCADO DE DATOS

Data mart.

### MERCADO DE INFORMACION

Info mart.

### METADATA REPOSITORY

Un Metadata Repository documenta e implementa las reglas y terminología del negocio, búsqueda, carga y mantenimiento del almacén de datos.



## **METADATO**

Los metadatos contienen la explicación y estructura de los datos mismos.

## **METODOLOGÍA**

Serie de etapas sistemáticas para el desarrollo de un proyecto.

## **METODOLOGÍA SAS DATA WAREHOUSE**

Metodología creada por The SAS Institute.

## **MODELO FÍSICO**

El modelo físico define la arquitectura actual de almacenamiento del Data Warehouse, además de que se definen las tablas del Data Warehouse.

## **MODELO LÓGICO**

El modelo lógico define las entidades de datos y las relaciones entre estas entidades requeridas para resolver el problema en cuestión.

## **MODELADO DIMENSIONAL**

El modelado dimensional es el proceso de establecer apropiadamente facts y dimensiones y ordenarlos en uno o más esquemas de estrella.

# **N**

## **NORMALIZACIÓN**

Minimizar la redundancia en los datos.



## O

### **OLAP (On-Line Analytical Processing)**

#### **Procesos Analíticos En Línea**

Procesos que accedan a datos basados en reglas de negocios, además de comparar puntos similares de datos a cualquier nivel de sumariación.

### **OLTP (On-Line Transaction Process)**

#### **Procesos de Transacción en Línea**

Es la puerta principal a la información en una organización. Administra las actividades diarias, además de que coleccionan y editan los datos almacenados

### **ON LINE**

Término inglés que significa en línea o conectado. Se aplica a dispositivos conectados directamente a una computadora que están preparados para su funcionamiento inmediato y que están bajo el control directo de un programa

## P

### **PRIMARY KEY**

Llave primaria.

## Q

### **QUERY**

Consulta.

---



## R

### **REDUNDANCIA**

Repetición en la información.

### **REPORTE**

Forma en que se visualiza la información.

## S

### **SEGURIDAD FÍSICA**

Mecanismos para el control de acceso a los equipos.

### **SEGURIDAD LÓGICA**

La seguridad lógica se refiere al control lógico dentro del software.

### **SQL**

Lenguaje normalizado que permite comunicarse con la mayor parte de los sistemas manejadores de bases de datos relacionales.

### **SUBJECT**

Son tareas enfocadas a los procesos de negocios no a transacciones individuales.

### **SUMARIZACIÓN**

Es agregar datos de los resúmenes estadísticos por computadora sobre datos continuos o crear cuentas sobre datos discretos.



## T

### **TABLA SUMARIZADA**

Las tablas resumizadas almacenan facts en categorías por dimensiones.

### **TECNOLOGIA DE INFORMACIÓN**

Information technology.

### **TRADEOFF**

Es una opción para elegir el tamaño y el rendimiento del modelo lógico de datos, por ejemplo el tamaño del Data Warehouse y el tiempo de respuesta a las consultas.

### **TRANSACCIÓN**

Operaciones específicas en un sistema OLTP.

### **TRANSFORMACIÓN DE DATOS**

La transformación es el término de cobertura para todas las áreas que preparan a los datos para ser cargados dentro del Data Warehouse.

## V

### **VALIDACIÓN**

Proyección de los datos para identificar valores potencialmente inválidos, fuera de rango, perdidos o duplicados.



## W

### WEB

Site. Sitio en el World Wide Web. Conjunto de páginas Web que forman una unidad de presentación, como una revista o libro. Un sitio está formado por una colección de páginas Web.



## REFERENCIAS



Data Warehouse Overview  
Theory And Business Concepts Course Notes  
SAS Institute Inc.  
Cary, NC USA 1999  
194 p.



Rapid Warehousing Methodology  
SAS Institute Inc.  
Cary, NC USA 2000  
263 p.



Building A Data Warehouse  
Using SAS/Warehouse/Administrator Software  
SAS Institute Inc.  
Cary, NC USA 2000  
440 p.



Introduction To SAS/Eis And SAS/MDDB  
Server Software Course Notes  
SAS Institute Inc.  
Cary, NC USA 2000  
388 p.



SAS Web Tools  
Static And Dynamic Solutions  
Using SAS/Intranet  
Software Course Notes  
SAS Institute Inc.  
Cary, NC USA 2000  
430 p.



SAS Web Tools  
Accessing MDDB Data Using Webeis  
Software course notes  
SAS Institute Inc.  
Cary, NC USA 1999  
212 p.



SAS Programming I  
Essentials Course Notes  
SAS Institute Inc.  
Cary, NC USA 1999  
490 p.



SAS Programming II  
Manipulating Data with Data Step Course Notes  
SAS Institute Inc.  
Cary, NC USA 1999  
437 p.



SAS Programming III  
Advanced Techniques Course Notes  
SAS Institute Inc.  
Cary, NC USA 1999  
516 p.



Getting Started with Enterprise Guide  
Software Course Notes  
SAS Institute Inc.  
Cary, NC USA 2001  
240 p.