



Universidad de las Ciencias Informáticas

Facultad 6

*Trabajo de Diploma para optar por el título de
Ingeniero en Ciencias Informáticas*

*"Sistema de Información de Gobierno. Mercado de datos Balance
económico financiero: capa de visualización"*

***Autores:** Cecilia Elisa Abreu Sánchez*

Leidiana Torres Roca

***Tutores:** Lic. Elena Leonila Fernández García*

Ing. Yaneisy López Marrero

"Ciudad de La Habana, 2011 "

"Año 53 de la Revolución"

“Piensa como piensan los sabios, mas habla como habla la gente sencilla.”

Aristóteles

DECLARACIÓN DE AUTORÍA

Declaramos que somos las únicas autoras del trabajo titulado “Sistema de Información de Gobierno. Mercado de datos Balance económico financiero: capa de visualización” y autorizamos a la Universidad de las Ciencias Informáticas los derechos patrimoniales del mismo, con carácter exclusivo.

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

Firma del Autor

Cecilia Elisa Abreu Sánchez

Firma del Autor

Leidiana Torres Roca

Firma del Tutor

Elena Leonila Fernández García

Firma del Tutor

Yaneisy López Marrero

Tutores:

Tutor: Lic. Elena Leonila Fernández García

Especialidad de graduación: Matemática

Categoría docente: Instructor

Categoría Científica: Licenciada

Años de experiencia en el tema: 32

Años de graduado: 32

Correo Electrónico: efernandezg@uci.cu

Tutor: Ing. Yaneisy López Marrero

Especialidad de graduación: Ingeniero en Ciencias Informáticas

Categoría Científica: Ingeniera

Años de experiencia en el tema: 2

Años de graduado: 2

Correo Electrónico: yaneisylm@uci.cu

-A la Revolución, a nuestro comandante Fidel y a Raúl Castro por darnos la oportunidad de estudiar en esta universidad y formarnos como profesionales.

- A nuestros padres, por enseñarnos el camino a seguir, por apoyarnos en los momentos difíciles y haber sonreído con nosotras en los momentos felices, por depositar tanta confianza y enseñarnos que lograr un sueño es posible, estimulándonos a alcanzar siempre metas más altas.

- A nuestras familias, por ser la cuna donde hemos nacido y crecido, por ser las manos que nos han forjado sin esperar nada a cambio. Les queremos dar este hermoso regalo, que aunque pequeño ha de ser, es nuestro sueño añorado.

- A todos los profesores que contribuyeron a nuestra formación como ingenieros, compartiendo sus experiencias durante estos años. Siempre les estaremos agradecidos.

- A nuestras tutoras Elena y Yanesiy por su apoyo y dedicación. Por guiarnos y aconsejarnos en todo momento.

- A nuestro tribunal en especial a nuestro oponente Yunier, por corregirnos cada detalle y

estar dispuesto a ayudarnos cada vez que hizo falta.

- A nuestros compañeros de los eternos grupos 6105, 6107 y 6512, que durante todos estos años nos han acompañado, ayudado, que formarán parte de nuestras vidas para siempre, aunque quizás nos sea difícil volver a encontrarnos. Gracias por tantos recuerdos inolvidables.

-A todos nuestros amigos, gracias por estar a nuestro lado cuando los hemos necesitado, por hacernos sonreír y por preocuparse tanto por nosotras: Vania, Taimí, Yndira, Aylenis, Bety, Mayelín, Lili, Yesenia, Katy, Edelis, Maray, Edel, Michael, Flavio, Nani, Jandro, Raiko.

-A todos los que de una forma u otra nos han ayudado, sepan que este logro también es gracias a ustedes.

- Y más que nada, agradecemos a Dios, por darnos la vida y el amor de nuestros padres.

Leidiana Torres y Cecilia Elisa Abreu

A mi mamá por ser mi mejor amiga, mi ejemplo de ser humano, por darme siempre tanto amor, tanto apoyo y tanta alegría. Gracias por soportar tanto tiempo separada de mí sin decir ni una palabra, siempre exhortándome a esforzarme para lograr mi sueño.

A mi abuelita, por estar conmigo en todos los momentos de mi vida, por ayudarme a dar los primeros pasos, por complacerme en todo, por ser la mejor abuela y madre del mundo.

A mi papá, mi tío querido Pipo, tía Kenia y abuela Mercy por que sé que están allí para mí, por siempre, por el amor y el cariño con que me han criado, por permitirme vivir en una familia feliz.

A mi novio Jorge Alberto por todo su amor, cariño y apoyo siempre que lo necesito. Por enseñarme tantas cosas y sobre todo por ser parte importante de mi vida.

A mi suegritos Gina, Pablo y Albertico por acogerme como una hija y brindarme apoyo y la familia que tengo lejos.

A Yinet por ser como una hermana y una amiga a lo largo de este tiempo, a Edgar por su ayuda en todo lo que he necesitado.

A mi dúo de tesis Elisa, por ser tan perseverante, comprensiva, por saber ser amiga más que nada, por darme fuerzas en los momentos en que más lo necesité. Ceci hoy logramos nuestra meta.

Gracias a todos los que pusieron un granito de arena en esta investigación y en mi carrera, sin ustedes nada de esto sería posible.

Leidiana Torres Roca

A mis padres por su eterno amor, dedicación y toda la confianza que han depositado en mí. Por guiarme en la vida y encontrar siempre en ellos todo el apoyo y el cariño que necesito para seguir adelante.

A mi abuelita Chilito y mi tío Pachi que aunque no se encuentren hoy con nosotros, se que estarían muy orgullosos de mí, por ver este sueño hecho realidad. Gracias por dedicar parte de sus vidas a ayudarme a crecer, complacerme en mis caprichos y entregarme todo su amor.

A mi hermanito Félix José por todo su cariño y por ser la luz de mi vida.

A toda mi familia, en especial a mi abuela Bertha, mi tía Mayra y mis primos Liset, Yamilet, Albertico e Isachy por hacerme sentir una persona especial en sus vidas e inspirarme fuerzas para seguir adelante.

A mi novio Alexeis por todo su amor, cariño y apoyo incondicional cada vez que lo he necesitado, su compañía y dedicación para que yo lograra mi sueño.

A mis suegros Nery y Alexis por apoyarme en todas las decisiones, que durante este largo andar, me ha impuesto el destino.

A mi compañera de tesis y amiga Leidiana, por todo su apoyo, por darme fuerzas para seguir adelante, por luchar junto a mí, y alegrarme los momentos difíciles. Gracias por compartir este momento que quedará grabado en mí por siempre.

A todos los que de una forma u otra contribuyeron con el éxito de la investigación, les agradezco de todo corazón.

Cecilia Elisa Abreu Sánchez

RESUMEN

La Inteligencia de negocios permite aprovechar toda la información operacional y convertirla en información estratégica, útil para tomar decisiones. Aprovechando las múltiples oportunidades que brinda la Inteligencia de negocios, el presente trabajo de diploma propone una capa de visualización, para el mercado de datos Balance económico financiero de la Oficina Nacional de Estadísticas que facilitará los procesos de consulta de información, referente a los estados financieros de todas las unidades presupuestadas del país. Para el desarrollo de la investigación se toma como base una adaptación de la metodología Kimball y la tesis de doctorado de Leopoldo Zenaido Zepeda Sánchez, que incorpora los casos de uso para guiar el proceso de desarrollo de software, ajustado a las normas de la Universidad de las Ciencias Informáticas. El mercado de datos fue modelado con la herramienta CASE Visual Paradigm 6.4 y el gestor de base de datos utilizado PostgreSQL 8.4.

El presente trabajo es el resultado de una investigación sobre los temas relacionados con almacenes de datos, tanto en el marco nacional como internacional. Culmina con la puesta en práctica del mercado de datos Balance económico financiero perteneciente al almacén de datos Sistema de Información de Gobierno.

Palabras claves: Inteligencia de negocios, Mercado de datos.

TABLA DE CONTENIDO

INTRODUCCIÓN.....1

CAPÍTULO 1: FUNDAMENTOS TEÓRICOS SOBRE EL DESARROLLO DE UN MERCADO DE DATOS5

1.1 Introducción5

1.2 Almacenes de datos.....5

1.3 Mercado de datos8

1.4 Diferencias entre almacén de datos y mercado de datos8

1.5 Actualidad de los almacenes de datos9

1.6 Bases de datos multidimensionales11

1.7 Modos de almacenamiento de la información12

 1.7.1 Tipos de procesamiento analítico en línea13

1.8 Inteligencia de negocios.....14

1.9 Metodologías y herramientas seleccionadas.....15

1.10 Conclusiones20

CAPÍTULO 2: ANÁLISIS Y DISEÑO DEL MERCADO DE DATOS BALANCE ECONÓMICO FINANCIERO.....21

2.1 Introducción21

2.2 Definición del negocio21

2.3 Necesidades de los usuarios22

2.4 Reglas del negocio.....23

2.5 Especificación de requerimientos.....24

 2.5.1 Requisitos de información24

 2.5.2 Requisitos funcionales25

 2.5.3 Requisitos no funcionales26

2.6 Modelo de casos de uso del sistema28

 2.6.1 Roles y permisos28

 2.6.2 Diagrama de casos de uso del sistema.....28

2.7 Especificación del modelo dimensional31

 2.7.1 Matriz BUS.....32

 2.7.2 Modelo de datos32

2.8 Esquema de seguridad38

2.9 Política de respaldo y recuperación39

2.10 Conclusiones39

CAPÍTULO 3: IMPLEMENTACIÓN Y PRUEBAS DEL MERCADO DE DATOS BALANCE ECONÓMICO FINANCIERO.....41

3.1 Introducción	41
3.2 Implementación del modelo de datos físico.....	41
3.3 Mapa de navegación.....	43
3.4 Implementación de los cubos OLAP	44
3.5 Arquitectura de información	45
3.6 Implementación de los reportes candidatos	46
3.7 Pruebas	48
3.7.1 Diseño de casos de pruebas.....	48
3.7.2 Lista de chequeo.....	51
3.7.3 Resultados y discusión de las pruebas	54
3.8 Conclusiones	54
CONCLUSIONES GENERALES.....	56
RECOMENDACIONES	57
REFERENCIAS BIBLIOGRÁFICAS	58
BIBLIOGRAFÍA.....	61
GLOSARIO DE TÉRMINOS.....	65

ÍNDICE DE FIGURAS

Figura 1 Estructura general de un almacén de datos7

Figura 2 Diagrama de casos de uso del sistema.....29

Figura 3 Modelo de datos33

Figura 4 Mapa de navegación.....44

Figura 5 Diseño de los cubos utilizando Schema Workbench45

Figura 6 Elementos que componen los cubos.....45

Figura 7 Arquitectura de información46

Figura 8 Resultado de la lista de chequeo54

ÍNDICE DE TABLAS

Tabla 1 Comparación entre almacén de datos y mercado de datos	9
Tabla 2 Comparación entre OLTP y OLAP	14
Tabla 3 Roles y permisos.....	28
Tabla 4 Caso de uso: Analizar información del estado de situación	31
Tabla 5 Matriz BUS.....	32
Tabla 6 Hecho indicadores generales	34
Tabla 7 Hecho indicadores seleccionados de la contabilidad.....	34
Tabla 8 Hecho cumplimiento del plan económico	35
Tabla 9 Hecho estado de situación	36
Tabla 10 Hecho disponibilidad de fondo	36
Tabla 11 Hecho estado de gastos.....	37
Tabla 12 Hecho inversiones y donaciones.....	38
Tabla 13 Hecho estado de resultado.....	38
Tabla 14 Esquemas y tablas.....	42
Tabla 15 Tabla de índices.....	43
Tabla 16 Caso de prueba Analizar información del estado de situación.....	50
Tabla 17 Descripción de las variables.....	50
Tabla 18 Lista de chequeo	54

INTRODUCCIÓN

En el mundo actual el desarrollo económico de los países, está dado por la capacidad de crear riquezas a fin de mantener el bienestar económico y social de sus habitantes. Es necesario recopilar toda la información financiera actualizada, para tomar las decisiones correspondientes sobre sus futuras operaciones, aportando información sobre el pasado y el presente para realizar previsiones sobre el futuro, con el propósito de ayudar a quienes toman decisiones en la evaluación de la situación financiera y las perspectivas del país.

Cuba como nación en vías de desarrollo, se ha dado a la tarea de informatizar los procesos que se llevan a cabo en las diferentes esferas del país, permitiéndoles lograr más eficiencia. Un proceso de vital importancia para la sociedad cubana, es la manera de recoger los datos que se generan en las instituciones de todo tipo del país. En la actualidad no se pueden analizar fenómenos, ni proyectarse soluciones, si no se cuentan con datos procesados, que es lo que se consideran estadísticas. Estas contribuirán a realizar estudios, para entender cómo se puede lograr el desarrollo económico y social del país, de forma más rentable posible.

El país posee una larga historia en materia de estadística. La entidad rectora de este tema es la Oficina Nacional de Estadísticas (ONE), que tiene como objetivo fundamental garantizar la calidad estadística a través del Sistema Estadístico Nacional. Dicha oficina se encuentra distribuida territorialmente por todas las provincias y municipios del país. Está estructurada por departamentos, direcciones, áreas, donde cada una de ellas se encarga de recoger información por las diferentes ramas y eslabones económicos. La información es captada a través de los centros informantes (CI) existentes a lo largo de la isla, y se procesa a través de una red de más de 2500 computadoras conectadas en tiempo real.

Una de las limitaciones que a diario presenta la máxima dirección del gobierno, es que la información que tiene sobre el uso del presupuesto, no le permite determinar realmente como en la base se utilizan los mismos, por la ausencia de un sistema de consulta que facilite el proceso de obtención de la información.

Este trabajo de diploma refiere su investigación a la dirección de Cuentas Nacionales, quien trabaja en conjunto con el Ministerio de Finanzas y Precios, determinando este último como se hará la captación de los estados financieros.

La dirección de Cuentas Nacionales trabaja con las transformaciones de los estados financieros que el ministerio define, y algunos de los indicadores que captan de los diferentes modelos con los que trabajan, se encuentran repetidos debido a que la información es recogida en varios espacios de tiempos y diferentes CI, creando confusión a los especialistas a la hora de digitalizar esta información, por el cúmulo de datos que estos contienen, trayendo como resultado que el proceso de recuperación y elaboración de informes sea muy costoso en esfuerzo y tiempo, provocando más demora en recopilar toda la información, que analizarla que es el propósito.

La ONE cuenta con un sistema llamado MSET_NT, el mismo es capaz de captar la información pero no permite consultarla, no deja relacionar las tablas puesto que no es una aplicación de base de datos (BD), tiene una salida independiente de ficheros, es capaz de emitir reportes de estos ficheros pero separados uno de otro.

Todos estos factores ocasionan un retraso significativo en la entrega de la información, a las áreas vitales de la economía nacional y esto afecta negativamente a las estrategias, políticas sociales y económicas del país.

Teniendo en cuenta lo expuesto previamente se plantea como **problema científico** de la investigación: ¿Cómo convertir en información los datos almacenados en el área del Balance económico financiero del Sistema de Información de Gobierno?

Por lo que el **objeto de estudio** es: inteligencia de negocios teniendo como **campo de acción**: técnicas de inteligencia de negocios para el área Balance económico financiero del Sistema de Información de Gobierno.

El presente trabajo de diploma tiene como **objetivo general**: desarrollar una capa de visualización para el área Balance económico financiero del Sistema de Información de Gobierno que convierta los datos almacenados en información.

Con el propósito de cumplir el objetivo general, se trazaron los siguientes **objetivos específicos**:

- Realizar el análisis y diseño del mercado de datos del área Balance económico financiero.
- Implementar la capa de visualización del área Balance económico financiero.
- Validar la capa de visualización del área Balance económico financiero.

Para dar cumplimiento a los objetivos se plantean las siguientes **tareas de la investigación**:

- Caracterización de las metodologías, herramientas y tecnologías a utilizar en el desarrollo de almacenes de datos.
- Levantamiento de requisitos.
- Descripción de los casos de uso del mercado de datos.
- Definición de los hechos, las medidas y las dimensiones del mercado de datos.
- Diseño del modelo de datos.
- Definición de la arquitectura del mercado de datos.
- Diseño del subsistema de visualización.
- Diseño de los casos de pruebas.
- Implementación del subsistema de visualización.
- Aplicación de la lista de chequeo.
- Aplicación de los casos de pruebas.

El siguiente trabajo estará compuesto por **3 capítulos**:

Capítulo 1: Fundamentos teóricos sobre el desarrollo de un mercado de datos.

Aborda algunos elementos teóricos acerca de la tecnología de almacenes de datos (AD) y los mercados de datos (MD). Además, se hace referencia a las principales características de las herramientas que se van a utilizar para el trabajo con estos sistemas. Se exponen las razones por las cuales se escogió una adaptación a la metodología Kimball y la tesis de doctorado de Leopoldo Zenaido Zepeda Sánchez para desarrollar todo el ciclo de vida del software.

Capítulo 2: Análisis y diseño del mercado de datos Balance económico financiero.

Contiene la descripción de los pasos a seguir durante el análisis y el diseño de la solución. Se identifican los requisitos de información, funcionales y no funcionales, se describen los casos de uso (CU), así como la realización del modelo de datos del MD. Se definen las políticas de seguridad y se realiza el diseño de lo que constituirá el sistema.

Capítulo 3: Implementación y pruebas del mercado de datos Balance económico financiero.

Contiene todos los elementos referentes a la implementación del sistema, las dimensiones, medidas y hechos que formarán parte de la posible solución, y la realización de las vistas de análisis del MD. Se incluirán los resultados de los casos de prueba y la lista de chequeo aplicada al sistema.



CAPÍTULO 1: FUNDAMENTOS TEÓRICOS SOBRE EL DESARROLLO DE UN MERCADO DE DATOS

1.1 Introducción

Durante el desarrollo del capítulo se exponen algunos elementos teóricos acerca de la tecnología de AD y los MD. Se brinda información de cada una de las herramientas que se seleccionan, así como sus características, ventajas y desventajas. Se toma como base para la realización del MD, la adaptación a la metodología Kimball y la tesis de doctorado de Leopoldo Zenaido Zepeda Sánchez, que incorpora los casos de uso para desarrollar todo el ciclo de vida del software.

1.2 Almacenes de datos

En la actualidad, y debido al desarrollo tecnológico de campos como la informática y las comunicaciones, la mayoría de las BD están en formato digital, eliminando así su dependencia del trabajo manual ofreciendo un amplio rango de soluciones al problema de almacenar datos. Cada día la información es fundamental e indispensable en la toma de dediciones por parte de la gerencia de cualquier empresa.

Los AD se han convertido en piezas clave en el entorno empresarial, cuyo objetivo fundamental es el de ayudar a los usuarios a entender mejor el pasado, así como planear el futuro. Esta tecnología no es más que la integración de las técnicas de BD, con las de análisis de datos.

Según Ralph Kimball, un AD es “una copia de las transacciones de datos específicamente estructurada para la consulta y el análisis”. [1]

Características del almacén de datos

Según William H. Inmon un AD es “(...) una colección de datos orientados al dominio, integrados, no volátiles y variables en el tiempo, organizados para dar apoyo al proceso de toma de decisiones”. [2]

- **Orientados al dominio:** Los datos están almacenados por diferentes materias o temáticas. Los mismos se organizan desde la perspectiva del usuario final con vista a lograr una mayor eficiencia en el acceso a la información.
- **Integrado:** Los datos proceden de distintas fuentes, en general heterogéneas, por lo que deben estar integrados en un mismo formato y una misma estructura consistente que lo soporte.
- **No volátil:** El Almacén de Datos sólo permite cargar nuevos datos y acceder a los ya almacenados, pero no permite ni borrar ni modificar los datos. [3]
- **Variables en el tiempo:** El tiempo debe estar presente en todos los registros contenidos en un AD. Esto posibilita realizar análisis de tendencias y comparaciones.

Etapas del almacén de datos

Los principales procesos que conforman un AD son: análisis y diseño, extracción, transformación y carga generalmente conocido por sus siglas en inglés ETL (Extraction, Transformation and Load) y la generación de la capa de visualización (Business Intelligence (BI)).

Análisis y diseño: Durante la etapa de análisis es donde se determinan los requisitos iniciales, se definen las reglas del negocio e identifican las fuentes de datos operacionales y externas, es aquí donde se define el alcance mediante entrevistas con los clientes finales.

La etapa de diseño comprende el análisis del negocio en cuestión, para así poder identificar las necesidades o requerimientos del negocio. De esta forma definir un modelo conceptual que debe traducirse más tarde en el modelo lógico del DWH. [4]

Es en esta etapa donde se define la estructura del AD.

Extracción: El proceso de extracción consiste en estudiar y entender los datos desde diferentes fuentes, tomando aquellos que son de utilidad para el AD.

Transformación: Una vez que los datos son extraídos desde de las diferentes fuentes pueden ser incoherentes, tener errores o estar incompletos, estos se transforman y homogeneizan. Este proceso incluye corrección de errores, integración de los datos con el propósito de eliminar problemas de redundancia e identificar las fuentes de datos más fiables, borrado de campos que no son de interés y usar conversiones y combinaciones para generar nuevos campos.

Carga: Organización y actualización de los datos en la BD.

Inteligencia de negocios: Microsoft, define la Inteligencia de Negocios como la manera de organizar y acceder rápidamente a información relevante, precisa y actual, con la que pueden adoptar mejores decisiones y potenciar la capacidad de su negocio.

La etapa ETL antes mencionada constituye aproximadamente un 70% del trabajo de la construcción de un AD. Por lo que es fundamental realizar correctamente los procesos que la conforman, debido a que se pueden obtener datos incorrectos, afectando todo el proceso de toma de decisiones. En la Figura 1 se muestra de forma general cada una de estas etapas.

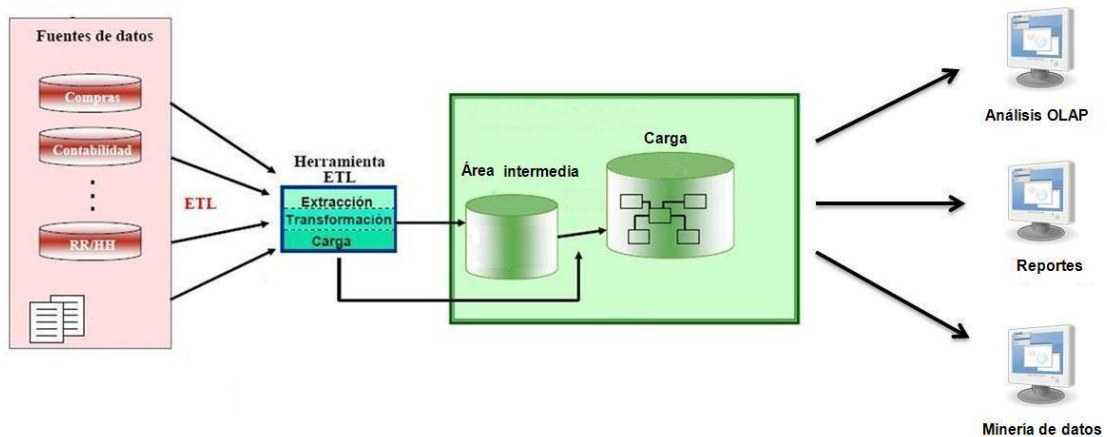


Figura 1 Estructura general de un almacén de datos

Ventajas de los almacenes de datos

- La inversión que realiza una organización para una correcta implantación de un sistema de Almacén de Datos conlleva un coste muy elevado, sin embargo, el retorno de la inversión es garantizado en gran medida.
- Integración de información procedente de múltiples sistemas.
- Los Almacenes de Datos hacen más fácil el acceso a una gran variedad de datos.
- Se obtiene una base de datos clasificada por temas e histórica. [3]
- Permiten disponer de la información necesaria en poco tiempo.
- Los directivos pueden dedicar más tiempo al análisis de los datos en vez de a su búsqueda.
- Cuando se aplica al marketing permite profundizar en el comportamiento del cliente.

- Debido a la acumulación de datos y experiencias se pueden detectar tendencias y hacer previsiones.

Desventajas de los almacenes de datos

- La subestimación del tiempo requerido para extraer, limpiar y cargar los datos en el Almacén.
- Problemas con los sistemas de origen de los datos.
- Pueden suponer altos gastos, además de los gastos de mantenimiento que son muy elevados.
- La construcción de un Almacén de Datos puede requerir de mucho tiempo. [3]

1.3 Mercado de datos

Es un subconjunto de datos de un Almacén relativos a los requisitos de un departamento o área de negocio concreto. Este subconjunto de datos puede funcionar de forma autónoma, o bien enlazado al Almacén de Datos. El motivo por el cual se crean mercados de datos es el crecimiento que tiene el Almacén y así facilitar su construcción y utilización. [3]

Características:

- Se centran en los requisitos de los usuarios asociados a un departamento o área de negocio concreto. [3]
- Son más sencillos a la hora de utilizarlos y comprender sus datos, debido a que la cantidad de información que contienen es mucho menor que en los Almacenes de Datos. [3]

Ventajas de los MD:

- Convertir los datos operacionales en información relacionada y estructurada, llegando a generar el conocimiento necesario para la toma de decisiones.
- Centralizar y homogeneizar la información.
- Permitir la visión global de la información en base a los conceptos de negocio que tratan los usuarios.
- Reducir costes evitando costosas extracciones manuales, permitiendo dedicar recursos a otras tareas.

1.4 Diferencias entre almacén de datos y mercado de datos

Almacén de Datos	Mercado de Datos
-------------------------	-------------------------

Corporativo o red empresarial.	Departamental
Es la unión de todos los MD.	Un proceso de negocios.
Tecnología óptima para el acceso a los datos y el análisis de gran cúmulo de información.	Tecnología óptima para el acceso a los datos y el análisis, dirigido a un departamento específico.
Estructura para vista corporativa de los datos.	Estructura para adaptarse a la vista de los datos departamentales.

Tabla 1 Comparación entre almacén de datos y mercado de datos

Los MD como estructura, son un Almacén de Datos pero reducido a un departamento específico sirviendo como fuente de análisis del tema que concierne a dicho departamento. La unión de todos los MD de la organización es lo que conformaría la vista global de los datos, es decir, el Almacén de Datos Corporativo. [5]

1.5 Actualidad de los almacenes de datos

En el mundo actual existe una gran necesidad por parte de las grandes empresas y compañías de responder con rapidez y eficiencia a las solicitudes de los usuarios y a las suyas propias, esto provoca que se haya elevado la necesidad de incorporar tecnologías de última generación, para crear estrategias de comercialización y mercado, basadas principalmente en los clientes. Las empresas basan su poder en el uso de la información, esta es su principal fuente de búsqueda como cualquier otro aspecto clave y estratégico de la organización. Y es aquí donde entran los AD, que permiten desarrollar y mantener los sistemas de datos donde el factor más determinante es el tiempo.

Mundialmente la utilización de los AD ha tenido un auge considerable:

➤ Santo Domingo, 11 de febrero de 2010. La Oficina Nacional de Estadística (ONE) puso en línea el Almacén Central de Datos (Data Warehouse) del Sistema Estadístico Nacional. El Data Warehouse es una herramienta electrónica que tiene como objetivo principal disponer de estadísticas precisas y actualizadas para mejorar los servicios al alcance de los tomadores de decisiones, los investigadores y la sociedad en general. Actualmente, las informaciones disponibles para los usuarios a través del Data Warehouse son el Censo Nacional de Población y Vivienda 2002, el Directorio de Establecimientos Económicos 2009, los nacimientos, defunciones, matrimonios y divorcios registrados

entre el 2001-2008, las encuestas de Ingresos y Gastos de los Hogares 2007 y las de Hogares de Propósitos Múltiples 2005, 2006 y 2007, entre otras. [6]

➤ DATATUR Almacén de Datos para el Análisis y Difusión de la Información Estadística del Turismo en España. Los principales objetivos de DATATUR son: Conservar y mantener al día los datos estadísticos, complementados con las correspondientes informaciones documentales para su correcta utilización. [7]

➤ En México “surge como proyecto institucional el Almacén de Datos Estadísticos cuyo propósito es integrar la información estadística que se genera e integra en el INEGI en un repositorio para consulta y análisis que permita a los usuarios tomar mejores decisiones en el contexto del Sistema Nacional de Estadística e Información Geográfica “. [8]

En Cuba la aplicación de estas nuevas tecnologías aún se halla reducida y faltan muchos aspectos por mejorar, se han dedicado esfuerzos en este sentido, pero en general se puede decir que estos constituyen los primeros pasos para el futuro desarrollo de las organizaciones en este campo. Entre las organizaciones que han dado pasos firmes en este sentido se encuentran:

➤ El AD comercial de la corporación CIMEX, que se dedica fundamentalmente a la exportación e importación de mercancías. Forman parte de ella un conjunto de empresas que se encuentran enfocadas en diversos negocios, aquí se puede citar la red de comercio minorista y la dirección de logística, esta última dedicada al comercio mayorista. El mismo centra su atención en la gestión de inventario, permitiendo una gestión de compra-venta eficiente, con la finalidad de disminuir los costos, sin afectar al cliente, permitiendo prestaciones eficientes y con la calidad requerida, aumentando las utilidades de las mismas.

➤ En el XIII Concurso Nacional de Computación y en la Feria de Informática del 2002, se presentó un AD para CUBACEL desarrollado sobre la plataforma Oracle, con grandes resultados obtenidos a partir de su implantación.

➤ El sistema SanoMillenium (...) contribuye a lograr el bienestar general del hombre a través de la Sanología y la Promoción de Salud. Este sistema constituye la fuente que nutrirá al Data Warehouse enfocado a la Promoción de Salud, respondiendo a las necesidades informacionales y apoyando la toma de decisiones en este campo y lograr una comunidad saludable. [9]

➤ Diseño e implementación de una herramienta que permita gestionar y organizar homogéneamente la información relevante actual e histórica sobre los indicadores de eficiencia de la gestión contable de

la Empresa de Proyectos de Arquitectura e Ingeniería (EMPAI) de Matanzas, utilizando las posibilidades que brindan los Almacenes de Datos (AD). [10]

➤ Almacenes de datos aplicados a la seguridad ciudadana: Es necesario establecer un modelo iterativo e incremental para constituir las bases del desarrollo del Almacén de Seguridad Ciudadana, enfocado en la asimilación paulatina de la información por parte de los analistas, así como en el incremento del dominio de análisis del propio almacén. Esta información, al ser presentada de forma dinámica al usuario final, le permite a este, poder combinarla con distintos parámetros y alcanzar mejores interpretaciones útiles a la institución en apoyo a la toma de decisiones. [11]

➤ En la actualidad, la Universidad de las Ciencias Informáticas, se encuentra desarrollando un AD para la ONE, aún en proceso de construcción.

1.6 Bases de datos multidimensionales

Los AD son soportados por un modelo multidimensional, a diferencia de los sistemas de BD relacionales que se sustentan en el modelo Entidad-Relación. El modelo multidimensional almacena la información de forma tal que sea de fácil entendimiento por parte del usuario y que, a su vez, garantice la rápida y eficiente ejecución de las consultas. Esta estructura de almacenamiento tiene forma de cubo multidimensional, donde las aristas representan las dimensiones, y las intersecciones entre estas aristas son los hechos.

Las **tablas de hechos** son los elementos primarios del modelo dimensional y contienen los valores del negocio. Una fila de la tabla de hechos corresponde a una medida, que no es más que la intersección entre todas las dimensiones

Las **tablas de dimensiones** las tablas de dimensiones almacenan los valores que se utilizan en las tablas de hechos. Desempeñan un papel vital dentro de los AD, puesto que son la fuente de casi todas las consultas que se le pueden realizar a este. Debido a esto se les considera la llave para hacer un AD usable y comprensible.

El modelo multidimensional implica tres variantes de modelamiento de acuerdo a la complejidad del sistema:

➤ **Esquema estrella:** Es la técnica más común, en este esquema existe un único elemento central (tabla de hechos) conectado radialmente con las tablas de dimensiones.

- **Esquema copo de nieve:** Es un esquema derivado del de estrella, donde las tablas de dimensiones se ramifican en más puntas.
- **Esquema constelación de hechos:** Este esquema es más complejo que las otras arquitecturas debido a que contiene múltiples tablas de hechos. Con esta solución las tablas de dimensiones pueden estar compartidas entre más de una tabla de hechos. [12]

Niveles

Un nivel representa un nivel particular de agregación dentro de una dimensión; cada nivel sobre el nivel base representa la sumarización total de los datos desde el nivel inferior. [13]

Jerarquías

Una jerarquía es un orden determinado dentro de las variables o atributos de una dimensión. Estas jerarquías son utilizadas a la hora de agregar (navegación ascendente) o desagregar (navegación descendente) la información. [14]

Granularidad

El grano o la granularidad de la tabla queda determinada por el nivel de detalle que se almacenará. [13]

Es el nivel más atómico por el cual se definen los datos. La granularidad afecta a la cardinalidad, tanto de las dimensiones como de las tablas de hechos, a mayor granularidad mayor será el número de registros final de la tabla de hechos.

1.7 Modos de almacenamiento de la información

Procesamiento transaccional en línea

Online Transactional Processing (OLTP), tiene como objetivo mantener la integridad de la información (relaciones entre los datos) necesaria para operar un negocio de la manera más eficiente. Sin embargo, no son eficientes cuando se trata de grandes cantidades de información, ni de consultas muy complejas. Los datos están normalizados, es decir, se han eliminado las redundancias y cualquier problema que pueda crear dificultades al actualizar cierto dato. Para analizar los datos los usuarios deben examinarlos registro a registro, tomando así grandes cantidades de tiempo cuando se trata de mucha información, admiten el acceso simultáneo de muchos usuarios que agregan y modifican datos. Representan el estado, en cambio constante de una organización, pero no guardan su historial. Por lo tanto, no es un sistema óptimo para ser utilizado en la toma de decisiones.

Procesamiento analítico en línea

“OLAP es un proceso analítico de datos online que permite al usuario seleccionar y extraer la información desde diferentes puntos de vista, y que se traduce en un conjunto de técnicas (...) OLAP permite hacer un análisis correcto y rápido de la información desde los puntos de vista o "dimensiones" que el usuario desee.” [15]

1.7.1 Tipos de procesamiento analítico en línea

➤ ROLAP (Relational On-line Analytical Process). Son herramientas OLAP que crean vistas multidimensionales extrayendo los datos de bases de datos SQL ordinarias, es decir, relacionales. Estas herramientas simulan los datos multidimensionales usando sofisticadas técnicas de indexación, caches, metadato, etc. [15]

➤ Herramientas MOLAP (Multidimensional On-line Analytical Process). Son herramientas que acceden a datos que no están almacenados en registros de tablas, sino que almacenan los datos en arrays de varias dimensiones, llamados cubos. Estos cubos utilizan índices para optimizar el acceso a los datos. [15]

➤ Herramientas HOLAP (Hybrid On-line Analytical Process). “Permiten un análisis híbrido de la información, es decir, que une lo mejor de los dos tipos anteriores. El análisis HOLAP nos ayudará a reducir costes de hardware (...) Además, la respuesta de las consultas sobre las bases de datos multidimensionales son más rápidas que sobre las relacionales. Como aspecto negativo, los datos multidimensionales deben ser cargados antes de ser consultados y refrescados cuando se actualizan los datos de la organización“. [15]

Comparación entre OLTP y OLAP

En la tabla siguiente se exponen las diferencias principales entre los dos tipos de procesos, OLTP y OLAP.

Comparación	OLTP	OLAP
Acceso	Acceso de miles de usuarios.	Acceso de un número de usuarios moderado.
Finalidad de los datos.	Para controlar y ejecutar las tareas fundamentales del negocio.	Para ayudar con la planificación, resolución de problemas y apoyo a

		las decisiones.
Redundancia	Se eliminan las redundancias (más eficiencia en las transacciones, más dificultades para la navegación del usuario).	Datos redundantes (acceso fácil para el usuario y buenos tiempos de respuesta).
Normalización	Altamente normalizados con muchas tablas.	Posiblemente sin normalizar parcialmente por razones de rendimiento.
Modificación de los datos.	Permite la modificación de los datos.	No se modifican los datos.
Representan	El estado de una organización.	Guardan el historial de una organización.
Estructura	Estructuras de BD complejas.	Estructuras de BD simples.

Tabla 2 Comparación entre OLTP y OLAP

1.8 Inteligencia de negocios

La Inteligencia de negocios “permite la explotación de la información de una organización para apoyar la toma de decisiones de sus ejecutivos”. [16]

La capa de visualización tiene la función de convertir la información de una enorme cantidad de datos que solo son útiles para el trabajo diario, en datos que pueden proporcionar conocimiento y ayudar a tomar decisiones. Con herramientas de BI se pueden generar vistas de análisis, informes, gráficos y todo lo necesario para que los interesados puedan analizar y comprender los datos fácilmente, permitiendo llegar a conclusiones o realizar decisiones en su área de trabajo.

Ventajas:

- Generar reportes globales o por secciones.
- Crear una BD de clientes.
- Crear escenarios con respecto a una decisión.

- Análisis multidimensionales.
- Generar y procesar datos.
- Cambiar la estructura de toma de decisiones.
- Mejorar el servicio al cliente.

1.9 Metodologías y herramientas seleccionadas

Metodología

Una metodología es aquella guía que se sigue a fin de realizar las acciones propias de una investigación, es el conjunto de métodos que rigen una investigación científica o en una exposición doctrinal, indicando qué hacer y cómo actuar cuando se quiere obtener algún tipo de investigación.

En esta investigación se toma como base la metodología Kimball, la cual es muy eficaz y conduce a una solución completa en una cantidad muy pequeña de tiempo, lo que indudablemente es muy efectivo en el proceso de la toma de decisiones. Además, propone el desarrollo iterativo incremental, donde se construye una pieza a la vez. Permite construir un MD como primer elemento del sistema de análisis, y luego ir añadiendo otros que compartan las dimensiones ya definidas o incluyen otras nuevas. Es una metodología madura y reconocida por el resto de la comunidad dedicada al tema, posee una mayor velocidad de respuesta al cliente, comprensible para el usuario, resistente y preparada para cambios, cuenta con respaldo y soporte. Tiene bien definidas las etapas, actividades, artefactos y roles. Una característica importante para la selección es que existe gran cantidad de documentación y se pueden encontrar respuestas a casi todas las preguntas sobre el tema.

En unión a la metodología de Kimball se utiliza también la tesis de doctorado de Leopoldo Zenaido Zepeda Sánchez, propone un nuevo estándar para el desarrollo de sistemas dirigido por modelos, se enmarca en el área del diseño de AD que utiliza la incorporación de los CU para guiar el proceso de desarrollo de software. Este método es empleado con una metodología compuesta y consta de tres fases.

- La primera fase está dedicada a examinar el esquema entidad relación de la BD operacional, generando los esquemas multidimensionales candidatos para el AD.
- En la segunda fase, los requisitos del usuario son recogidos por medio de entrevistas. El propósito de las entrevistas es obtener información acerca de las necesidades de análisis de los usuarios de datos.

➤ La tercera fase contrasta la información obtenida en la segunda fase, con los esquemas multidimensional candidatos formados en la primera fase generando así, la mejor solución (soportada por las BD operacionales) que mejor reflejan los requisitos de usuario.

Tomando como base la metodología de Kimball y la tesis de doctorado de Leopoldo Zenaido Zepeda Sánchez, se logra un mayor alineamiento con las tendencias y normas de la universidad.

Gestor de bases de datos

Un sistema gestor de bases de datos (SGBD) es un software informático especializado y disponible en el mercado que se utiliza para la creación, acceso, control, y gestión de las bases de datos. [17]

Los componentes principales de un SGBD son el lenguaje de definición de datos (DDL), y el lenguaje de manipulación de datos (DML).

PostgreSQL

“PostgreSQL se diseñó como una base de datos orientada a objetos (...) Esto significa, que las tablas no son tablas, sino objetos, y las tuplas son instancias de ese objeto (...) PostgreSQL tiene transacciones, integridad referencial, vistas, y multitud de funcionalidades”. [18]

Cuenta con más de 15 años de desarrollo activo, es un sistema fiable y tiene una arquitectura probada, que se ha ganado una sólida reputación, tiene características avanzadas tales como:

- Consistencia es la propiedad que asegura que sólo se empieza aquello que se puede acabar. Por lo tanto, se ejecutan aquellas operaciones que no van a romper las reglas y directrices de integridad de la base de datos.
- Aislamiento es la propiedad que asegura que una operación no puede afectar a otras. Esto asegura que dos transacciones sobre la misma información nunca generará ningún tipo de error.
- Durabilidad es la propiedad que asegura que una vez realizada la operación, ésta persistirá y no se podrá deshacer aunque falle el sistema.
- Corre en casi todos los principales sistemas operativos: Linux, Unix, BSDs, Mac OS, Beos, Windows, etc.
- Documentación muy bien organizada, pública y libre, con comentarios de los propios usuarios.
- Altamente adaptable a las necesidades del cliente.
- Soporte nativo para los lenguajes más populares del medio: PHP, C, C++, Perl, Python, etc.
- Soporte de protocolo de comunicación encriptado por SSL.

- Utilidades para limpieza de la base de datos (Vacuum). [19]

Mejoras de la versión 8.4

Para la realización del trabajo de diploma se utilizará la versión 8.4. Entre las mejoras más populares de esta versión están:

- Restauración de bases de datos en procesos paralelos, que acelera recuperación de un respaldo hasta 8 veces.
- Privilegios por columna, que permiten un control más granular de datos confidenciales.
- Configuración de ordenamiento configurable por base de datos, lo cual hace a PostgreSQL más útil en entornos con múltiples idiomas.
- Nuevas herramientas de monitoreo de consultas que le otorgan a los administradores mayor información sobre la actividad del sistema. [20]

PGAdmin III 1.10

PgAdmin III es una aplicación gráfica para gestionar el gestor de bases de datos PostgreSQL, siendo la más completa y popular con licencia Open Source. Está escrita en C++. PgAdmin III está diseñado para responder a las necesidades de todos los usuarios, desde escribir consultas SQL simples hasta desarrollar bases de datos complejas. La interfaz gráfica soporta todas las características de PostgreSQL y facilita enormemente la administración. [21]

Características:

- Soporte en español.
- Modificación de datos activos o en uso.
- Manipulación de propiedades activas o en uso.
- Completo y muy utilizado.
- Es una interfaz comprensible para el diseño y administración de una BD PostgreSQL, diseñada para ejecutarse en la mayoría de los sistemas operativos.

Herramienta de modelado

Las herramientas CASE (Computer Aided Software Engineering), no son más que un conjunto de programas y ayudas que dan asistencia a los analistas, ingenieros de software y desarrolladores.

Visual Paradigm 6.4

Se selecciona Visual Paradigm en su versión 6.4 para darle solución al problema por las ventajas que posee para el análisis y diseño del MD.

Visual Paradigm es una herramienta de Lenguaje Unificado de Modelado (UML) profesional que soporta el ciclo de vida completo del desarrollo de software: análisis y diseño orientados a objetos, construcción, pruebas y despliegue.

Ventajas:

- Diseño centrado en casos de uso y enfocado al negocio que genera un software de mayor calidad.
- Uso de un lenguaje estándar común a todo el equipo de desarrollo que facilita la comunicación.
- Capacidades de ingeniería directa (versión profesional) e inversa.
- Modelo y código que permanece sincronizado en todo el ciclo de desarrollo.
- Disponibilidad de múltiples versiones, para cada necesidad.
- Disponibilidad en múltiples plataformas. [22]

Desventajas:

- Una desventaja de esta herramienta es que posee licencia privada de Visual Paradigm International Limited, pero la UCI posee dicha licencia. [23]

Herramientas para el proceso de BI

Las aplicaciones de BI han evolucionado en muchas direcciones, debido al crecimiento exponencial de la información. Las aplicaciones de BI son herramientas de soporte de decisiones, que permiten en tiempo real, acceso interactivo, análisis y manipulación de información crítica para las empresas.

- **Pentaho BI Server 3.6:** Es una aplicación 100% Java 2 Platform Enterprise Edition (J2EE) que permite gestionar todos los recursos de BI. Cuenta con una interfaz de usuario donde se encuentran disponibles vistas de análisis, reportes, entre otros. Además, posee una consola de administración que permite gestionar y supervisar tanto la aplicación como los usuarios.

Características:

- Aplicación J2EE 100% extensible, adaptable y configurable.
- Provee la gestión de la configuración tanto en la instalación inicial como en el mantenimiento.

- Se integra con la mayoría de entornos y se puede comunicar con otras aplicaciones vía web services.
- Integra todos los recursos informacionales en una única plataforma de explotación
- Proporciona mucha libertad al usuario y los desarrolladores para crear contenidos nuevos.
- Explotación de sus recursos como SOAP web services. [24]
- **Mondrian OLAP server 3.0.4:** Es un motor OLAP escrito en Java. Ejecuta las preguntas escritas en el lenguaje Multi Dimensional Expressions (MDX), la lectura de los datos de una base de datos relacional, y presenta los resultados en un formato multidimensional a través de un API de Java. [25]

Entre sus principales características se encuentra la facilidad para el análisis de grandes volúmenes de información, que se encuentren almacenados en BD. Permite realizar consultas al AD y que los resultados sean presentados mediante un navegador, donde usando JPivot el usuario puede realizar las actividades típicas de navegación sobre cubos OLAP.

- **Pentaho Schema Workbench 3.2.0:** Esta es una herramienta para el desarrollo del esquema del modelo multidimensional, el cual se guarda en formato XML, desarrollado en Java. Este programa fue publicado en el 2007, entrega todas las facilidades para poder realizar el modelo lógico del cubo OLAP al cual se le realizarán las consultas. Este programa se conecta directamente con la BD, para así poder diseñar los cubos OLAP que se requieren, para que el usuario final pueda visualizar los indicadores en el MD.

- **Servidor Apache Tomcat 5.5**

El Servidor Apache Tomcat fue escrito en Java, por lo que funciona en cualquier sistema operativo que tenga instalado la máquina virtual de Java. Tomcat es el servidor web más utilizado a la hora de trabajar con Java en entornos web. Es una implementación completamente funcional de los estándares de JSP y Servlets. También puede especificarse, como el manejador de las peticiones de JSP y Servlets recibidas por servidores webs populares. Tomcat está integrado en la implementación de referencia J2EE de Sun Microsystems. Puede funcionar como servidor Hypertext Transfer Protocol (HTTP), o conectado a otro servidor HTTP como Apache HTTP Server, o Internet Information Services (IIS).

Herramienta para el control de versiones

Subversion (SVN 1.5.4) Subversion es un sistema de control de versiones libre y de código fuente abierto. Es decir, Subversion maneja ficheros y directorios a través del tiempo. Hay un árbol de ficheros en un repositorio central. El repositorio es como un servidor de ficheros ordinario, excepto porque recuerda todos los cambios hechos a sus ficheros y directorios. Esto le permite recuperar versiones antiguas de sus datos, o examinar el historial de cambios de los mismos (...) Subversion puede acceder al repositorio a través de redes, lo que le permite ser usado por personas que se encuentran en distintos ordenadores. [26]

1.10 Conclusiones

Tomando como punto de partida la fundamentación teórica y la importancia que reviste la misma para la realización de una correcta investigación sobre los AD y MD en el capítulo anterior se arriban a las siguientes conclusiones:

- Se exponen las ventajas de las herramientas seleccionadas para confeccionar el MD, evidenciando que dicha selección ayudará a desarrollar el mercado de manera más eficiente y robusta.
- Se explican las características que componen la adaptación a la metodología seleccionada, la cual guiará todo el proceso de desarrollo del MD.
- Se muestran ejemplos del desarrollo que han alcanzado los AD, tanto a nivel nacional como internacional.



CAPÍTULO 2: ANÁLISIS Y DISEÑO DEL MERCADO DE DATOS BALANCE ECONÓMICO FINANCIERO

2.1 Introducción

En este capítulo se definen las reglas del negocio, se diseñan los artefactos que ayudan a construir la propuesta, se identifican las necesidades del usuario, requisitos de información, requisitos funcionales y no funcionales, además se modelan y validan los requerimientos. Se definen las dimensiones del MD y los hechos asociados.

Análisis

El análisis consiste en lograr una visión general de la propuesta, para conseguir una comprensión más precisa de los requisitos, y una descripción de los mismos que ayude a estructurar todo el sistema.

2.2 Definición del negocio

El Ministerio de Finanzas y Precios pone en vigor las normas cubanas sobre la información financiera, como base para el registro de los hechos económicos en las entidades del país. El mismo propone la legislación y los sistemas, que aseguren la integridad y el control financiero de los intereses del estado cubano, en entidades públicas, privadas y asociaciones con capital extranjero, incluyendo los principios, normas y procedimientos de contabilidad y costos. Establece las proformas y las indicaciones metodológicas de los estados financieros, con el propósito de homogeneizar y posibilitar los procesos de agregación de la información a diferentes niveles.

Hasta este momento lo que estaba definido por la ONE era la captación de indicadores contables en tres modelos estadísticos (0005 Indicadores Generales, 5903 Cumplimiento del Plan Económico, 5901 Indicadores Seleccionados de la Contabilidad). Pero se ha podido comprobar que en muchos casos, estos indicadores no se correspondían con los registrados en los balances financieros de las unidades presupuestadas (UP). Por esta razón se decide captar a partir del 2011 los registros primarios (5910

Estado de Situación, 5911 Estado de Resultado, 5912 Inversiones y Donaciones, 5913 Estado de Disponibilidad de Fondo, 5914 Estado de Gastos) definidos por Finanzas como soporte para la captación de los estados financieros, y aprovechando las posibilidades que nos brindan los AD, poder obtener los propios registros primarios además de los indicadores estadísticos. Esta información será captada y procesada a través de la red interna del Sistema de Información de Gobierno, y estará al servicio tanto de la ONE como del gobierno para una mejor toma de decisiones.

Tema de análisis

Para la construcción de la propuesta se definió como tema de análisis: Balance económico financiero.

2.3 Necesidades de los usuarios

Son las publicaciones de información, desde diferentes fuentes, donde el usuario además de consultarla puede realizar su análisis. A continuación se detallan las necesidades de usuarios identificadas:

- El usuario necesita analizar la información del formulario 5910 “Estado de Situación”, por alguna de las siguientes dimensiones: división política administrativa, organismo, entidad, subordinación, nomenclador de actividades económicas de todos los indicadores que se solicitan en el formulario en el tiempo por: apertura y en esta fecha.
- El usuario necesita analizar la información del formulario 5911 “Estado de Resultado”, por alguna de las siguientes dimensiones: división política administrativa, organismo, entidad, subordinación, nomenclador de actividades económicas de todos los indicadores que se solicitan en el formulario en el tiempo por: hasta esta fecha año anterior, plan anual y hasta la fecha.
- El usuario necesita analizar la información del formulario 5912 “Inversiones y Donaciones”, por alguna de las siguientes dimensiones: división política administrativa, organismo, entidad, subordinación, nomenclador de actividades económicas de todos los indicadores que se solicitan en el formulario en el tiempo por: plan, apertura y en esta la fecha.
- El usuario necesita analizar la información del formulario 5913 “Disponibilidad de Fondo”, por alguna de las siguientes dimensiones: división política administrativa, organismo, entidad, subordinación, nomenclador de actividades económicas de todos los indicadores que se solicitan en el formulario en el tiempo por: importe.
- El usuario necesita analizar la información del formulario 5914 “Estado de Gastos”, por alguna de las siguientes dimensiones: división política administrativa, organismo, entidad, subordinación,

nomenclador de actividades económicas, forma de gastos y nomenclador de actividades económicas por gastos de todos los indicadores que se solicitan en el formulario en el tiempo por: plan y real.

- El usuario necesita analizar la información del modelo 0005-11 “Indicadores Generales”, por alguna de las siguientes dimensiones: división política administrativa, organismo, entidad, subordinación, nomenclador de actividades económicas de todos los indicadores que se solicitan en el formulario en el tiempo por: real del año actual.
- El usuario necesita analizar la información del modelo 5901-08 “Indicadores Seleccionados de la Contabilidad”, por alguna de las siguientes dimensiones: división política administrativa, organismo, entidad, subordinación, nomenclador de actividades económicas de todos los indicadores que se solicitan en el formulario en el tiempo por: saldo inicial y saldo final.
- El usuario necesita analizar la información del modelo 5903-04 “Cumplimiento del Plan Económico”, por alguna de las siguientes dimensiones: división política administrativa, organismo, entidad, subordinación, nomenclador de actividades económicas de todos los indicadores que se solicitan en el formulario en el tiempo por: real acumulado del período del año actual.

2.4 Reglas del negocio

En las reglas del negocio se definen y controlan la estructura, el funcionamiento y la estrategia que la propuesta debe contemplar para cumplir con las necesidades del cliente. Durante las entrevistas con el cliente, se obtuvieron un total de 97 reglas del negocio, a continuación se especifican algunas de ellas, para ver las restantes reglas del negocio remitirse al Anexo 1.

1. Para obtener el activo circulante es necesario realizar la adición de los siguientes indicadores: Efectivo en Caja, Efectivo en Banco, Efectos por Cobrar, Cuentas por Cobrar a Corto Plazo, Cuentas por Cobrar Compra de Moneda, Pagos Anticipados a Suministradores, Pagos Anticipados del Proceso Inversionista, Anticipos a Justificar, Adeudos del Presupuesto del Estado, Adeudos del Organismo Superior, Reparaciones Generales en Proceso, Inventarios y Créditos Documentarios.

2. Para obtener el inventario es necesario realizar la adición de los siguientes indicadores: Materias Primas y Materiales, Combustibles y Lubricantes, Partes y Piezas de Repuesto, Envases y Embalajes, Útiles y Herramientas y Producción Terminada menos la adición de: Descuentos Comerciales e Impuesto, Mercancía para la Venta, Medicamentos, Base Material de Estudio, Vestuario y Lencería, Alimentos, Otros Inventarios, Producciones en Proceso, Producción Propia para Insumo, Inventarios Ociosos e Inventarios de Lento Movimiento.

3. Los Activos a Largo Plazo se obtienen a través de las Inversiones a Largo Plazo o Permanente.
4. El Activo Fijo se obtiene de la adición de los siguientes indicadores: Activos Fijos Tangibles Netos, Activos Fijos Intangibles Netos, Fondos Bibliotecarios, Inversiones Materiales, Equipos por Instalar y Materiales del Proceso Inversionista.
5. Para obtener los Activos Fijos Tangibles Netos es necesario realizar la adición de los siguientes indicadores: Depreciación de Activos Fijos Tangibles, Compra de Activos Fijos Tangibles y Activos Fijos Tangibles.

2.5 Especificación de requerimientos

2.5.1 Requisitos de información

Los requisitos de información, describen qué información debe almacenar el sistema para satisfacer las necesidades de clientes y usuarios. Identifican los conceptos relevantes sobre los que se debe almacenar información.

En el negocio se identificaron un total de 118 requisitos de información a continuación se mencionan algunos, para ver los restantes requisitos de información remitirse al Anexo 2.

RI-1 Obtener los indicadores seleccionados del estado de situación por apertura de las unidades presupuestadas por organismo en el tiempo.

RI-2 Obtener los indicadores seleccionados del estado de situación por apertura de las unidades presupuestadas por subordinación en el tiempo.

RI-3 Obtener los indicadores seleccionados del estado de situación por apertura de las unidades presupuestadas por división política administrativa en el tiempo.

RI-4 Obtener los indicadores seleccionados del estado de situación por apertura de las unidades presupuestadas por nomenclador de actividades económicas en el tiempo.

RI-5 Obtener los indicadores seleccionados del estado de situación por apertura de las unidades presupuestadas por indicadores del estado de situación en el tiempo.

RI-6 Obtener los indicadores seleccionados del estado de situación por apertura de las unidades presupuestadas por entidad en el tiempo.

RI-7 Obtener los indicadores seleccionados del estado de situación en esta fecha de las unidades presupuestadas por organismo en el tiempo.

RI-8 Obtener los indicadores seleccionados del estado de situación en esta fecha de las unidades presupuestadas por subordinación en el tiempo.

RI-9 Obtener los indicadores seleccionados del estado de situación en esta fecha de las unidades presupuestadas por división política administrativa en el tiempo.

RI-10 Obtener los indicadores seleccionados del estado de situación en esta fecha de las unidades presupuestadas por nomenclador de actividades económicas en el tiempo.

2.5.2 Requisitos funcionales

El proceso de obtención de los requisitos es importante en el desarrollo del MD, pues de la calidad de estos depende el éxito del producto. Los requisitos funcionales describen la función que el sistema debe hacer. Para el desarrollo del MD se identificaron los siguientes requisitos funcionales:

RF1- Autenticar usuario

RF2- Adicionar rol

RF3- Eliminar rol

RF4- Adicionar usuario

RF5- Eliminar usuario

RF6- Insertar reporte

RF7- Eliminar reporte

RF8- Modificar reporte

RF9- Abrir navegador OLAP.

RF10- Mostrar editor MDX.

RF11- Configurar tabla OLAP.

RF12- Mostrar Padres.

RF13- Ocultar repeticiones.

RF14- Mostrar propiedades.

RF15- Suprimir filas/columnas vacías.

RF16- Intercambiar ejes.

RF17- Detallar miembros.

RF18- Abrir detalles.

RF19- Entrar en detalles.

RF20- Mostrar datos origen.

RF21- Mostrar gráfico.

RF22- Configurar gráficos.

RF23- Configurar impresión.

RF24- Exportar a PDF.

RF25- Exportar a Excel.

2.5.3 Requisitos no funcionales

Se define como las propiedades o cualidades que el producto debe tener. Debe pensarse como las características que hacen al producto atractivo, usable, rápido y confiable. A continuación se especifican algunos de los requisitos no funcionales propuestos para el MD, para ver los restantes requisitos no funcionales remitirse al Anexo 3.

Fiabilidad

RNF 1. Asegurar la disponibilidad del sistema.

El sistema debe permanecer disponible las 24 horas del día. En caso de fallo, restablecer el sistema lo más pronto posible.

Restricciones de diseño

RNF 2. Utilizar las herramientas y lenguaje de programación definido durante la investigación.

Para la programación en el MD se utilizará PL/pgSQL como lenguaje dentro del SGBD. Además, se utilizará el lenguaje MDX para realizar las consultas. El SGBD PostgreSQL en su versión 8.4, como interfaz de administración del SGBD se usará el PgAdmin en su versión 1.10. Como herramienta de modelado el Visual Paradigm en su versión 6.4. Para el control de versiones se utiliza el Subversion que permite manejar los archivos, carpetas y sus modificaciones en el transcurso del tiempo, en su versión 1.5.4.

De la suite de Pentaho, Pentaho BI Suite Community Edition en su versión 3.5, se usarán los siguientes componentes.

- Schema Workbench 3.2.1: Es la herramienta gráfica que se utiliza para construir el esquema multidimensional.
- Pentaho BI Server 3.5: Es el servidor de visualización de la suite, que se encarga de mostrar las vistas de análisis, controlar el acceso a la información y unificar en una solución de inteligencia de negocios, el uso de las demás herramientas que componen la suite.

- Pentaho Administrator Console 3.5: Es la herramienta que se utiliza para administrar el Pentaho BI Server, permite la configuración de las conexiones a las BD, tareas programadas, así como los roles y usuarios.
- Pentaho Desing Studio 3.5: Es la herramienta que proporciona el ambiente gráfico para construir y probar documentos de Secuencia de Acción a través de su colección de editores, visores y módulos de administración integrados.

Se requiere para el uso de las herramientas anteriores la máquina virtual de Java.

Requisitos de software

RNF 3. Instalar el software necesario en las máquinas clientes para el correcto funcionamiento del sistema.

Las configuraciones de software de las máquinas clientes deben contar al menos con los siguientes elementos:

- Navegador web.
- Máquina Virtual de Java.
- La aplicación debe soportar los sistemas operativos Windows o Linux.

RNF 4. Instalar el software necesario en los servidores para el correcto funcionamiento del sistema.

Las configuraciones de software de los servidores deben contar al menos con los siguientes elementos:

Los mismos elementos de las máquinas clientes, además de:

- Pentaho BI Server en caso de que un usuario capacitado requiera la construcción de esquemas multidimensionales para el diseño de nuevas vistas de análisis.
- PostgreSQL como gestor de BD.

Requisitos de hardware

RNF 5. Proporcionar características mínimas de hardware a las estaciones de trabajo.

- Características de un cliente ligero: procesador Modelo Intel Celeron, velocidad de 1.33 GHz, 256 de memoria RAM.
- Ordenador 256 MB RAM, Pentium III.

RNF 6. Proporcionar características mínimas de hardware a los servidores.

Los servidores deben contar con los siguientes requerimientos de hardware para lograr una explotación aceptable del sistema:

- 1 GB RAM permite hasta 9 conexiones.
- 1 Microprocesador Dual Core / 2.2 GHz.
- Disco duro de 160 GB.

2.6 Modelo de casos de uso del sistema

Los CU ayudan a describir qué es lo que se debe hacer, están enfocados al punto de vista del usuario. Es decir, describen cómo este interactúa con el usuario.

2.6.1 Roles y permisos

Se definen los roles y permisos en dependencia de la actividad y del tipo de información que maneje en el MD.

Roles	Permisos		Descripción
	Lectura	Escritura	
Administrador	X	X	Se encargará de administrar roles, usuarios y reportes.
Analista	X		Se encargará de analizar la información correspondiente a las UP.

Tabla 3 Roles y permisos

2.6.2 Diagrama de casos de uso del sistema

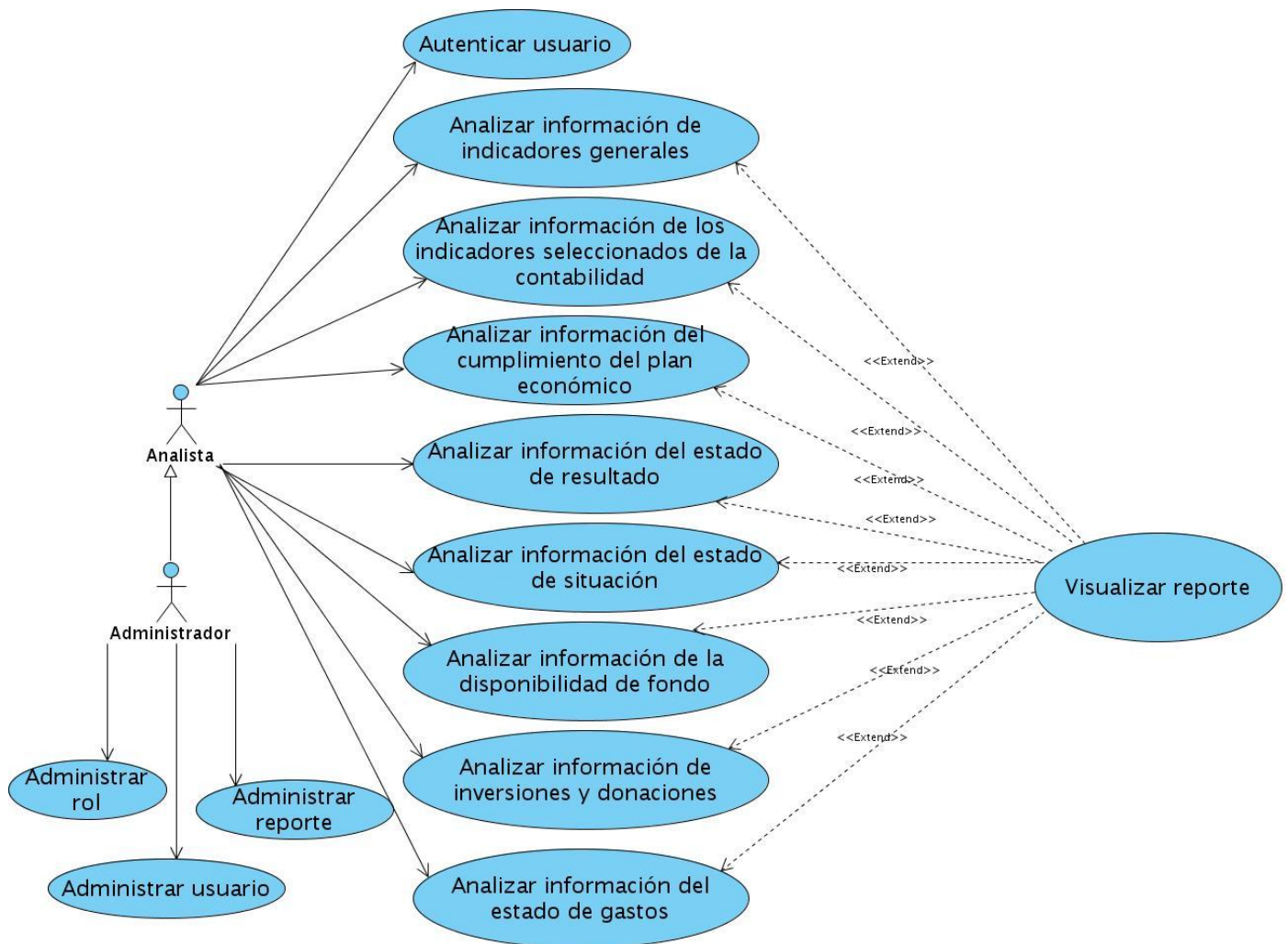


Figura 2 Diagrama de casos de uso del sistema

A continuación se realiza la descripción textual del caso de uso Analizar información del estado de situación, para ver las descripciones de los restantes CU, remitirse al Anexo 4.

Caso de Uso:	Analizar información del estado de situación.
Tipo:	Información
Actores:	Analista, Administrador.
Resumen:	El caso de uso se inicia cuando el usuario desea consultar la información relacionada con los indicadores del estado de situación, desde diferentes perspectivas de análisis. Luego de seleccionar el reporte deseado, el sistema muestra la información contenida en él y las opciones de los posibles cambios

	que le puede hacer al reporte. El caso de uso finaliza cuando el usuario termina el análisis de la información relacionada con el estado de situación.		
Precondiciones:	<ul style="list-style-type: none"> • Que el usuario esté autenticado en el sistema. • Completitud del mercado de datos. • Carga de los datos. • Disponibilidad de los reportes. 		
Referencias:	RI1 hasta RI12, CU Visualizar reporte.		
Prioridad:	Crítico		
Flujo Normal de Eventos			
Acción del Actor	Respuesta del Sistema		
1. El usuario se autentica en el sistema.	2. El sistema muestra la interfaz principal con las áreas de análisis existentes.		
3. El usuario selecciona el área de análisis A.A.G SIGOB.	4. El sistema muestra las áreas de análisis que están contenidas dentro del A.A.G SIGOB.		
5. El usuario selecciona el área de análisis A.A Balance económico financiero.	6. El sistema muestra los diferentes libros de trabajo que están contenidos dentro del A.A Balance económico financiero.		
7. El usuario selecciona el libro de trabajo L.T Estado de situación.	8. El sistema muestra los reportes contenidos dentro del L.T Estado de situación.		
9. El usuario selecciona el reporte deseado.	10. El sistema muestra la información contenida en el reporte seleccionado, y brinda la posibilidad al usuario de hacerle cambios al reporte para su análisis. Ir al caso de uso Visualizar reporte . Finaliza el caso de uso.		
Opciones de reportes del estado de situación			
Entradas	Posibles resultados		
	Salidas	Periodicidad	
Variables de entrada relacionadas con el caso	Variables de salida	Rango de tiempo en	

de uso Analizar información del estado de situación. ✓ DPA ✓ Nomenclador de Actividades Económicas ✓ Organismo ✓ Entidad ✓ Indicadores ✓ Subordinación ✓ Tiempo	disponibles en el caso de uso Analizar información del estado de situación. ✓ Apertura ✓ En esta Fecha	que se solicitan las variables de salidas. ✓ Mensual ✓ Anual
--	--	--

Prototipo de Interfaz

			Tiempo	
			Febrero	
			Medidas	
DPA	Entidad	Indicadores	● Apertura	● En esta Fecha
[-] Cuba	[+] Total	ACTIVO		
		Activo Circulante	6.434.619	8.156.659
		Activo Fijo	45.664.264	47.410.857
		Activos a Largo Plazo		
		Activos Fijos Intangibles	10.290	13.320
		Activos Fijos Intangibles Netos	9.624	9.624
		Activos Fijos Tangibles	40.457.041	43.111.538
		Activos Fijos Tangibles Netos	35.289.004	36.834.489
		Adeudos del Organismo Superior		
		Adeudos del Presupuesto del Estado	15.554	1.104.067

Poscondiciones:	Los reportes correspondientes al libro de trabajo L.T Estado de situación han sido consultados por el usuario.
------------------------	--

Tabla 4 Caso de uso: Analizar información del estado de situación

2.7 Especificación del modelo dimensional

El diseño es un refinamiento del análisis partiendo de los requerimientos definidos, traduce los requisitos en una representación del software. Debe ser suficiente para que la propuesta pueda implementarse sin ambigüedades, debido a que proporciona el cimiento bajo los cuales se va a desarrollar el software.

2.7.1 Matriz BUS

La matriz BUS especifica la relación entre los hechos y las dimensiones, donde las columnas contienen todas las dimensiones y las filas las tablas de hechos, la intersección de una fila con una columna, especifica si hay relación entre una tabla de hechos y una dimensión.

HECHOS/DIMENSIONES	Organismo	Subordinación	División Política Administrativa	Nomenclador de Actividades Económicas	Temporal mes	Indicador general	Nomenclador de Actividades Económicas por gastos	Forma de gastos	Entidad	Indicador balance
Estado de situación	X	X	X	X	X	-	-	-	X	X
Estado de resultado	X	X	X	X	X	-	-	-	X	X
Disponibilidad de fondo	X	X	X	X	X	-	-	-	X	X
Inversiones y donaciones	X	X	X	X	X	-	-	-	X	X
Estado de gastos	X	X	X	X	X	-	X	X	X	X
Indicadores generales	X	X	X	X	X	X	-	-	X	-
Indicadores seleccionados de la contabilidad	X	X	X	X	X	X	-	-	X	-
Cumplimiento del plan económico	X	X	X	X	X	X	-	-	X	-

Tabla 5 Matriz BUS

2.7.2 Modelo de datos

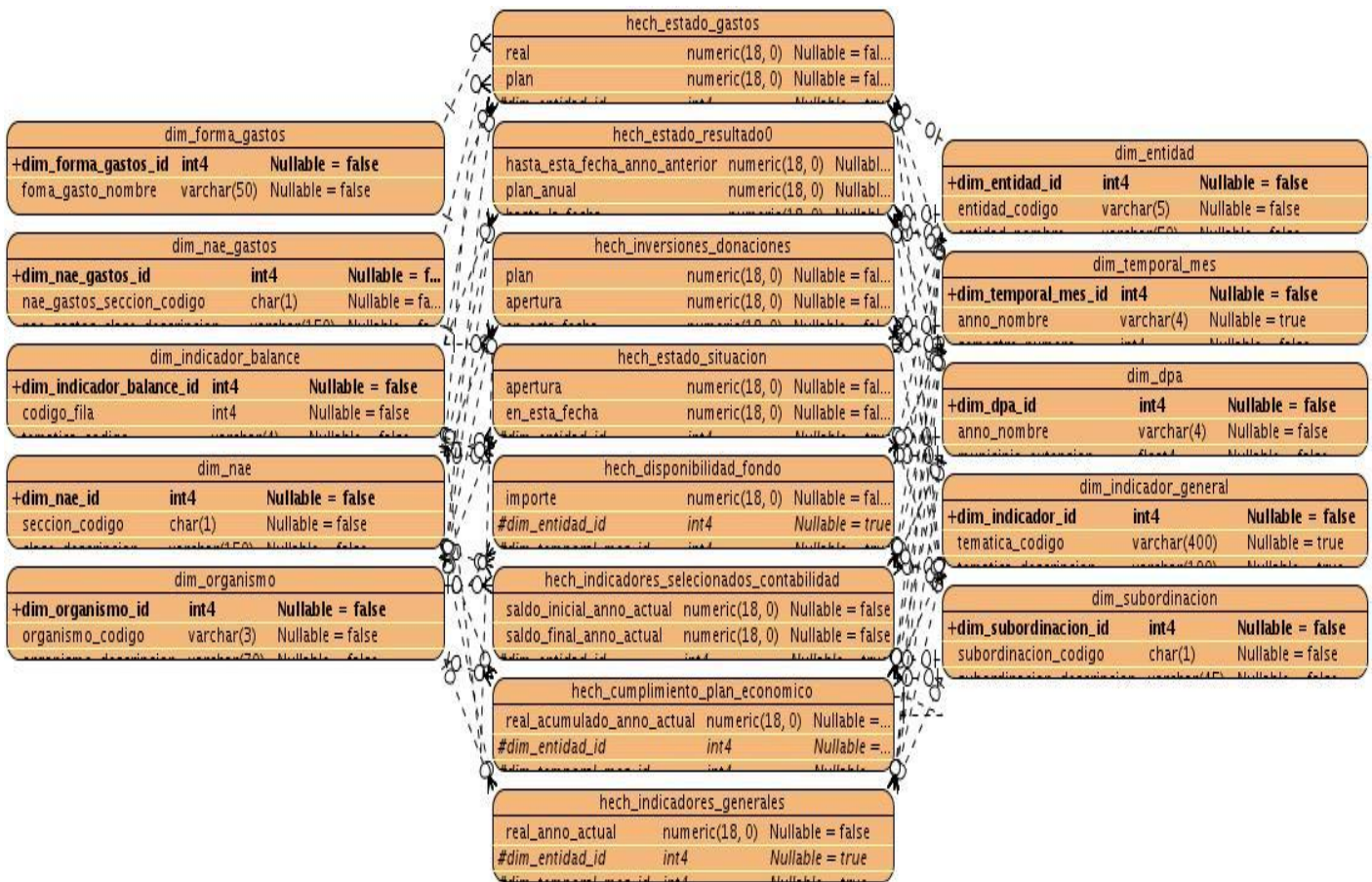


Figura 3 Modelo de datos

Tablas de hechos y medidas asociadas

➤ Indicadores generales

Atributos	ID	Tipo de dato
dim_subordinacion	dim_subordinacion_id	integer
dim_organismo	dim_organismo_id	integer
dim_dpa	dim_dpa_id	integer
dim_entidad	dim_entidad_id	integer
dim_nae	dim_nae_id	integer

dim_temporal_mes	dim_temporal_mes_id	integer
dim_indicador_general	dim_indicador_general_id	integer
Medidas		Tipo de dato
Real Año Actual		float

Tabla 6 Hecho indicadores generales

➤ Indicadores seleccionados de la contabilidad

Atributos	ID	Tipo de dato
dim_subordinacion	dim_subordinacion_id	integer
dim_organismo	dim_organismo_id	integer
dim_dpa	dim_dpa_id	integer
dim_entidad	dim_entidad_id	integer
dim_nae	dim_nae_id	integer
dim_temporal_mes	dim_temporal_mes_id	integer
dim_indicador_general	dim_indicador_general_id	integer
Medidas		Tipo de dato
Saldo Inicial Año Actual		float
Saldo Final Año Actual		float

Tabla 7 Hecho indicadores seleccionados de la contabilidad

➤ Cumplimiento del plan económico

Atributos	ID	Tipo de dato
-----------	----	--------------

dim_subordinacion	dim_subordinacion_id	integer
dim_organismo	dim_organismo_id	integer
dim_dpa	dim_dpa_id	integer
dim_entidad	dim_entidad_id	integer
dim_nae	dim_nae_id	integer
dim_temporal_mes	dim_temporal_mes_id	integer
dim_indicador_general	dim_indicador_general_id	integer
Medidas		Tipo de dato
Real Acumulado		float

Tabla 8 Hecho cumplimiento del plan económico

➤ Estado de situación

Atributos	ID	Tipo de dato
dim_subordinacion	dim_subordinacion_id	integer
dim_organismo	dim_organismo_id	integer
dim_dpa	dim_dpa_id	integer
dim_entidad	dim_entidad_id	integer
dim_nae	dim_nae_id	integer
dim_temporal_mes	dim_temporal_mes_id	integer
dim_indicador_balance	dim_indicador_balance_id	integer
Medidas		Tipo de dato

Apertura	float
En esta Fecha	float

Tabla 9 Hecho estado de situación

➤ Disponibilidad de fondo

Atributos	ID	Tipo de dato
dim_subordinacion	dim_subordinacion_id	integer
dim_organismo	dim_organismo_id	integer
dim_dpa	dim_dpa_id	integer
dim_entidad	dim_entidad_id	integer
dim_nae	dim_nae_id	integer
dim_temporal_mes	dim_temporal_mes_id	integer
dim_indicador_balance	dim_indicador_balance_id	integer
Medidas		Tipo de dato
Importe		float

Tabla 10 Hecho disponibilidad de fondo

➤ Estado de gastos

Atributos	ID	Tipo de dato
dim_subordinacion	dim_subordinacion_id	integer
dim_organismo	dim_organismo_id	integer
dim_dpa	dim_dpa_id	integer

dim_entidad	dim_entidad_id	integer
dim_nae	dim_nae_id	integer
dim_temporal_mes	dim_temporal_mes_id	integer
dim_indicador_balance	dim_indicador_balance_id	integer
dim_nae_gastos	dim_nae_gastos_id	integer
dim_forma_gastos	dim_forma_gastos_id	integer
Medidas		Tipo de dato
Real		float
Plan		float

Tabla 11 Hecho estado de gastos

➤ Inversiones y donaciones

Atributos	ID	Tipo de dato
dim_subordinacion	dim_subordinacion_id	integer
dim_organismo	dim_organismo_id	integer
dim_dpa	dim_dpa_id	integer
dim_entidad	dim_entidad_id	integer
dim_nae	dim_nae_id	integer
dim_temporal_mes	dim_temporal_mes_id	integer
dim_indicador_balance	dim_indicador_balance_id	integer
Medidas		Tipo de dato

Plan	float
Apertura	float
En esta Fecha	float

Tabla 12 Hecho inversiones y donaciones

➤ Estado de resultado

Atributos	ID	Tipo de dato
dim_subordinacion	dim_subordinacion_id	integer
dim_organismo	dim_organismo_id	integer
dim_dpa	dim_dpa_id	integer
dim_entidad	dim_entidad_id	integer
dim_nae	dim_nae_id	integer
dim_temporal_mes	dim_temporal_mes_id	integer
dim_indicador_balance	dim_indicador_balance_id	integer
Medidas	Tipo de dato	
Hasta Esta Fecha Año Anterior	float	
Plan Anual	float	
Hasta la Fecha	float	

Tabla 13 Hecho estado de resultado

2.8 Esquema de seguridad

El esquema de seguridad estará respaldado por los niveles de acceso, específicamente por los roles y usuarios definidos.

La arquitectura del sistema de seguridad y alta disponibilidad debe ser de tres capas:

- Funcionamiento: Dispositivos de seguridad (firewall).
- Servidores: Servidor de gestión de administración y de BD.
- Presentación: Consolas de administración.

La configuración y prueba de todos los equipos deberá efectuarse en las instalaciones de la ONE en Ciudad de La Habana, conjuntamente con personal técnico especializado, de acuerdo con las especificaciones de configuración y prueba.

El mantenimiento preventivo y correctivo deberá efectuarse según las especificaciones del contrato que se establezca entre la ONE y la Universidad de las Ciencias Informáticas.

2.9 Política de respaldo y recuperación

La política de respaldo y recuperación que utiliza la solución es sencilla, pero a la vez sólida, por ello se miden los siguientes puntos esenciales:

- Periodicidad de las salvas: Mensualmente se realizan las salvas de toda la información contenida en la BD en un servidor. La organización lo tiene definido de esta manera, certificando en todo momento la existencia de una copia escrita de la información que está presente en el servidor. Además, existe un grupo especial denominado Banco de Datos, perteneciente al área de la dirección de informática que también recoge toda la información.
- Cada seis meses antes de período vacacional (julio y diciembre), se hace una salva general de la información, y se traslada fuera de la ONE en un DVD-RW, siendo responsabilidad del grupo de administración de redes trasladarla hacia otro edificio correspondiente a la oficina de estadística de La Habana en tres versiones abuelo, padre e hijo, donde abuelo es la primera versión de la información, padre la segunda, e hijo la última y luego al traer la próxima salva se sustituye por la versión más antigua.

2.10 Conclusiones

Con el desarrollo de este capítulo se definen pautas importantes para realización del análisis y diseño del MD tales como:

- Se define el negocio así como la identificación de las reglas.
- Se especifican los requerimientos tanto funcionales como no funcionales, determinando estos los aspectos fundamentales a tener en cuenta para el desarrollo e instalación del MD.
- Se modelan los CU y se explican de forma tal, que puedan dar respuesta a las necesidades de los usuarios.
- Se definieron las dimensiones, medidas y hechos asociados, realizándose después el modelado tanto lógico como físico y generándose el script, necesario para llevar a cabo la propuesta de solución.



CAPÍTULO 3: IMPLEMENTACIÓN Y PRUEBAS DEL MERCADO DE DATOS BALANCE ECONÓMICO FINANCIERO

3.1 Introducción

En este capítulo se define la implementación y las pruebas de la capa de visualización del MD, además se incluirán los resultados de las pruebas realizadas al sistema, permitiendo corregir estas últimas los posibles errores para proveer una solución con calidad al cliente. En la implementación se utilizó el Schema Workbench para la modelación de los cubos OLAP, para la realización y visualización de las vistas de análisis se utilizó el Pentaho BI Server. Se realizan los casos de pruebas basados en casos de uso y la lista de chequeo para verificar si el producto final satisface o no las necesidades del cliente.

3.2 Implementación del modelo de datos físico

Para el desarrollo del sistema propuesto se definieron dos esquemas: dimensiones, que contiene todas las dimensiones comunes en el AD, y el esquema mart_balance, que agrupa todos los hechos de dicho mercado y las dimensiones propias del MD. La solución cuenta con 18 tablas en total, diez dimensiones y ocho hechos, distribuidas por los dos esquemas propuestos con anterioridad quedando de la siguiente manera:

Esquemas	Tablas
dimensiones	dim_nae
dimensiones	dim_entidad
dimensiones	dim_dpa
dimensiones	dim_subordinacion
dimensiones	dim_organismo

dimensiones	dim_temporal_mes
dimensiones	dim_indicador_general
mart_balance	dim_forma_gastos
mart_balance	dim_indicador_balance
mart_balance	dim_nae_gastos
mart_balance	hech_cumplimiento_plan_economico
mart_balance	hech_estado_situacion
mart_balance	hech_estado_resultado
mart_balance	hech_estado_gastos
mart_balance	hech_inversiones_donaciones
mart_balance	hech_disponibilidad_fondo
mart_balance	hech_indicadores_generales
mart_balance	hech_indicadores_seleccioandos_contabilidad

Tabla 14 Esquemas y tablas

Índices

El índice de una BD es una estructura de datos que está formada por las llaves primarias de las dimensiones, y se usan principalmente en campos de búsquedas frecuentes, debido a que aumentan la velocidad de acceso a la información.

Índice	Esquema	Tabla	Tipo de índices
PK54	dimensiones	dim_nae	Btree
PK47	dimensiones	dim_entidad	Btree
PK53	dimensiones	dim_dpa	Btree
PK55	dimensiones	dim_subordinacion	Btree

PK56	dimensiones	dim_organismo	Btree
PK58	dimensiones	dim_temporal_mes	Btree
PK59	dimensiones	dim_indicador_general	Btree
PK19	mart_balance	dim_forma_gastos	Btree
PK52_1_1_1	mart_balance	dim_indicador_balance	Btree
PK57	mart_balance	dim_nae_gastos	Btree
PK16	mart_balance	hech_cumplimiento_plan_economico	Btree
PK10	mart_balance	hech_estado_situacion	Btree
PK11	mart_balance	hech_estado_resultado	Btree
PK12	mart_balance	hech_estado_gastos	Btree
PK27	mart_balance	hech_inversiones_donaciones	Btree
PK13	mart_balance	hech_disponibilidad_fondo	Btree
PK14	mart_balance	hech_indicadores_generales	Btree
PK15	mart_balance	hech_indicadores_seleccionados_contabilidad	Btree

Tabla 15 Tabla de índices

3.3 Mapa de navegación

Un mapa de navegación representa como se organiza la información dentro de una estructura web. Expresa todas las relaciones de jerarquía y secuencia, que permiten elaborar escenarios de comportamiento de los usuarios. A continuación se representa el mapa de navegación que será presentado en la capa de visualización del MD Balance económico financiero.

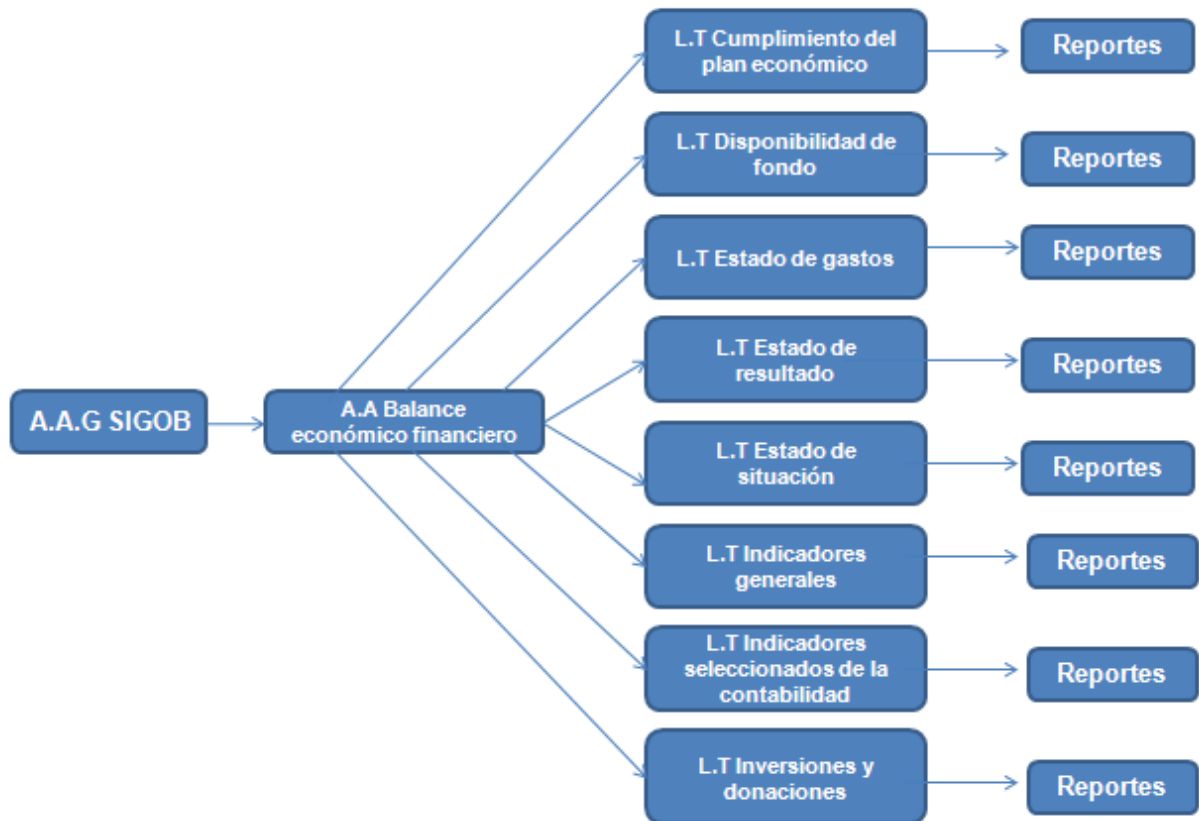


Figura 4 Mapa de navegación

3.4 Implementación de los cubos OLAP

Durante la implementación de los cubos OLAP, es necesaria la creación de cubos multidimensionales, los cuales se realizan utilizando la herramienta Schema Workbench la misma permite definir dimensiones, jerarquías, niveles de jerarquía, medidas, hechos y se define la conexión con la BD, obteniéndose esta configuración en un fichero XML. Se modelaron en el esquema ocho cubos multidimensionales correspondientes a cada una de las tablas de hechos identificadas en el MD, diez dimensiones y 15 medidas que dan solución al problema planteado.

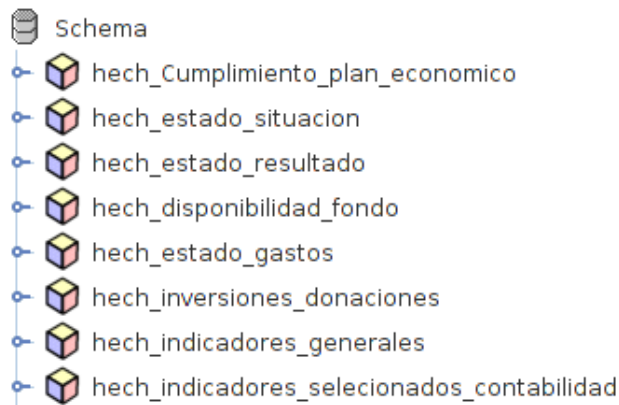


Figura 5 Diseño de los cubos utilizando Schema Workbench

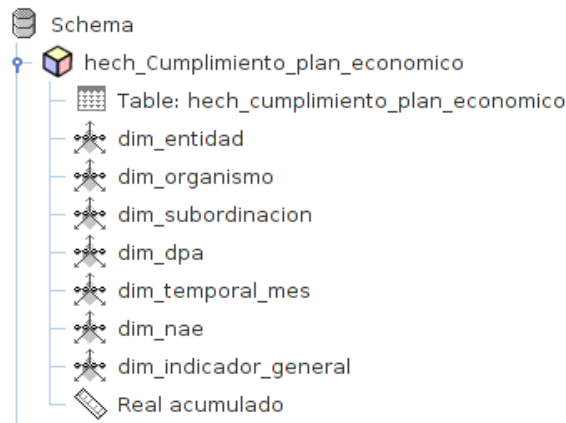


Figura 6 Elementos que componen los cubos

3.5 Arquitectura de información

A continuación se detallan los elementos que componen las estructuras de navegación de la información, que será presentada en la capa de visualización del MD Balance económico financiero; la misma contiene un Área de Análisis (A.A), ocho Libros de Trabajos (L.T) y 38 vistas de análisis, asociados a los libros de trabajos correspondientes.

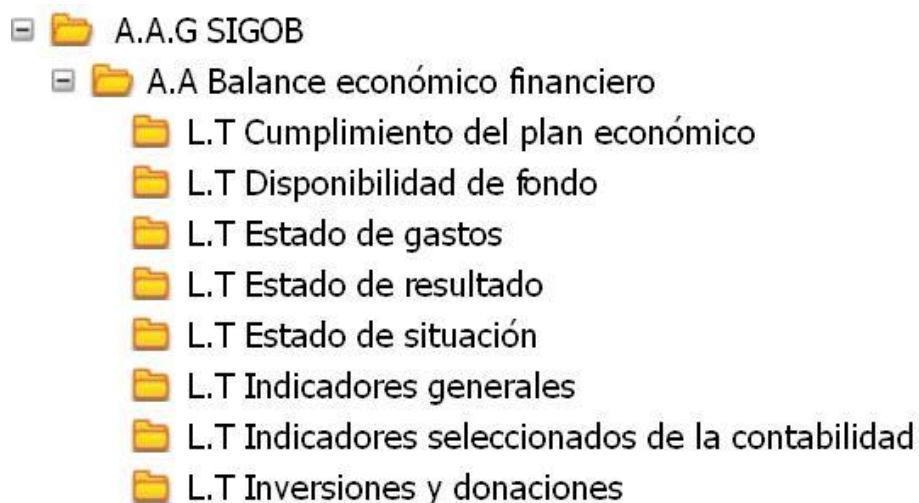


Figura 7 Arquitectura de información

3.6 Implementación de los reportes candidatos

En el siguiente epígrafe se detallan los reportes candidatos o tablas de salidas, como también se les denomina.

A.A.G SIGOB

A.A Balance económico financiero

L.T Cumplimiento del plan económico

- Por División Política Administrativa y entidad los indicadores calculados del plan económico.
- Por División Política Administrativa y entidad los indicadores del plan económico.
- Por Nomenclador de Actividades Económicas los indicadores calculados del plan económico.
- Por Nomenclador de Actividades Económicas los indicadores del plan económico.
- Por organismo los indicadores calculados del plan económico.
- Por organismo los indicadores del plan económico.
- Por subordinación los indicadores calculados del plan económico.
- Por subordinación los indicadores del plan económico.

L.T Disponibilidad de fondo

- Por División Política Administrativa y entidad los indicadores de disponibilidad de fondo.

- Por Nomenclador de Actividades Económicas los indicadores de disponibilidad de fondo.
- Por organismo los indicadores de disponibilidad de fondo.
- Por subordinación los indicadores de disponibilidad de fondo.

L.T Estado de gastos

- Por División Política Administrativa y entidad los indicadores del estado de gastos.
- Por Nomenclador de Actividades Económicas los indicadores del estado de gastos.
- Por forma de gastos los indicadores del estado de gastos.
- Por organismo los indicadores del estado de gastos.
- Por subordinación los indicadores del estado de gastos.
- Según gastos por actividad económica los indicadores del estado de gastos.

L.T Estado de resultado

- Por División Política Administrativa y entidad los indicadores del estado de resultado.
- Por Nomenclador de Actividades Económicas los indicadores del estado de resultado.
- Por organismo los indicadores del estado de resultado.
- Por subordinación los indicadores del estado de resultado.

L.T Estado de situación

- Por División Política Administrativa y entidad los indicadores del estado de situación.
- Por Nomenclador de Actividades Económicas los indicadores del estado de situación.
- Por organismo los indicadores del estado de situación.
- Por subordinación los indicadores del estado de situación.

L.T Indicadores generales

- Por División Política Administrativa y entidad los indicadores generales.
- Por Nomenclador de Actividades Económicas los indicadores generales.
- Por organismo los indicadores generales.
- Por subordinación los indicadores generales.

L.T Indicadores seleccionados de la contabilidad

- Por División Política Administrativa y entidad los indicadores seleccionados de la contabilidad.
- Por Nomenclador de Actividades Económicas los indicadores seleccionados de la contabilidad.
- Por organismo los indicadores seleccionados de la contabilidad.
- Por subordinación los indicadores seleccionados de la contabilidad.

L.T Inversiones y donaciones

- Por División Política Administrativa y entidad los indicadores de inversiones y donaciones.
- Por Nomenclador de Actividades Económicas los indicadores de inversiones y donaciones.
- Por organismo los indicadores de inversiones y donaciones.
- Por subordinación los indicadores de inversiones y donaciones.

3.7 Pruebas

El desarrollo de sistemas de software implica la realización de una serie de pruebas, que permiten identificar posibles fallos y relevar la calidad del producto. Existen diversas pruebas aplicables a los MD, en la presente investigación se utilizaron los casos de prueba basados en casos de uso y la lista de chequeo, para validar el funcionamiento del mismo.

3.7.1 Diseño de casos de pruebas

Los casos de prueba son un conjunto de condiciones o variables, bajo las cuales se determinará si los requisitos de una aplicación se han cumplido de forma parcial o completa. A continuación se muestra un ejemplo de un caso de prueba, aplicado al caso de uso de información “**Analizar información del estado de situación**”, para ver los restantes casos de prueba, remitirse al Anexo 5.

Escenario	Descripción	Perfiles de análisis	Indicadores a medir	Respuesta del sistema	Flujo central
EC 1.1: Por División Política Administrativa y entidad los indicadores del estado de situación	Permite visualizar el reporte con las variables presentes en el mismo.	DPA Entidad Indicadores Tiempo	Apertura En esta Fecha	Se muestra la tabla con los valores correspondientes a cada escenario.	Se abre la aplicación. Se autentifica. Se entra al sistema. Se despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Se selecciona el área de análisis de A.A Balance económico financiero. Se selecciona el libro de trabajo L.T Estado de situación. En la parte inferior izquierda se selecciona el reporte deseado. En el área de trabajo se visualiza la tabla correspondiente al
EC 1.2: Por Nomenclador de Actividades Económicas los indicadores del estado de situación		Nomenclador de Actividades Económicas Indicadores Tiempo	Apertura En esta Fecha		
EC 1.3: Por organismo los indicadores del estado de situación		Organismo Indicadores Tiempo	Apertura En esta Fecha		
EC 1.4: Por subordinación los indicadores del estado de situación		Subordinación Indicadores Tiempo	Apertura En esta Fecha		

					reporte.
--	--	--	--	--	----------

Tabla 16 Caso de prueba Analizar información del estado de situación

No	Nombre de campo	Clasificación	Valor nulo	Descripción
1	Subordinación	Lista desplegable	No	Muestra todas las subordinaciones existentes.
2	Organismo	Lista desplegable	No	Muestra todos los organismos existentes.
3	DPA	Lista desplegable	No	Muestra las provincias y municipios existentes de acuerdo con la nueva división política administrativa.
4	Nomenclador de Actividades Económicas	Lista desplegable	No	Muestra todos los nomencladores de actividades económicas existentes.
5	Entidad	Lista desplegable	No	Muestra todas las entidades existentes.
6	Indicadores	Lista desplegable	No	Muestra todos los indicadores existentes.
7	Tiempo	Lista desplegable	No	Muestra todos los meses.
8	Apertura	Valor fijo	No	Debe ser un valor entero.
9	En esta Fecha	Valor fijo	No	Debe ser un valor entero.

Tabla 17 Descripción de las variables

3.7.2 Lista de chequeo

Las listas de chequeo tienen el objetivo de examinar cuidadosamente áreas importantes, para lograr identificar problemas comunes, encontrar errores y solucionarlos en un espacio breve de tiempo. La lista de chequeo definida para la evaluación del MD, cuenta con varios puntos a medir, los mismos son: estructura del documento, indicadores definidos en el desarrollo y semántica del documento. La misma fue evaluada por un especialista; recoge 14 indicadores a evaluar, de ellos ocho críticos (Ver Tabla 18).

Estructura del documento					
Peso	Indicadores a evaluar	Eval	(NP)	Cantidad de elementos afectados	Comentarios
crítico	1. ¿Los entregables contienen las secciones obligatorias de la plantilla estándar definidas para un expediente de proyecto? (portada, control de versiones, reglas de confidencialidad, tabla de contenidos y contenido)	0		0	
Indicadores definidos en el desarrollo					
Peso	Indicadores a evaluar	Eval	(NP)	Cantidad de elementos afectados	Comentarios
	1. ¿Se utilizó un lenguaje cuyas sentencias son expresables mediante una	0		0	

	sintaxis bien definida?				
crítico	2. ¿Los reportes son configurables a través de la interfaz del sistema?	0		0	
	3. ¿La interfaz está orientada a facilitar el uso de las funciones del sistema por parte de los usuarios?	0		0	
crítico	4. ¿No existen restricciones para construir cubos OLAP con dimensiones y niveles de agregación ilimitados?	0		0	
crítico	5. ¿Los usuarios son capaces de manipular los resultados de manera que se ajusten a sus necesidades, conformando nuevos reportes?	0		0	
	6. ¿El sistema responde de una forma rápida a la información que le sea solicitada por el usuario?	0		0	
	7. ¿El sistema refleja cualquier lógica del negocio para poder responder a preguntas específicas?	0		0	
crítico	8. ¿El sistema garantiza la confidencialidad y seguridad de acceso a los datos por rol de los usuarios?	0		0	

	9. ¿Los datos e información derivados del proceso de análisis realizado mediante la aplicación, apoyan la toma de decisiones en la Institución?	0		0	
crítico	10. ¿Los cambios en los datos se reflejan automáticamente en los reportes de forma instantánea?	0		0	
Semántica del documento					
Peso	Indicadores a evaluar	Eval	(NP)	Cantidad de elementos afectados	Comentarios
crítico	1. ¿Se han identificado errores ortográficos en los entregables?	0		0	
crítico	2. ¿Se entiende claramente lo que se ha especificado en el documento?	0		0	
	3. ¿El número de página que aparece en el índice coincide con el contenido que se refleja realmente en dicha página?	1		1	El número de páginas que aparece en las reglas de confidencialidad no coincide con el número de páginas del documento: Evaluación de las áreas de la

					organización.
--	--	--	--	--	---------------

Tabla 18 Lista de chequeo

Como resultado, después de aplicada la lista de chequeo se genera el siguiente gráfico, donde se visualiza el comportamiento de los 14 indicadores identificados, fue detectada una no conformidad la cual fue resuelta satisfactoriamente.

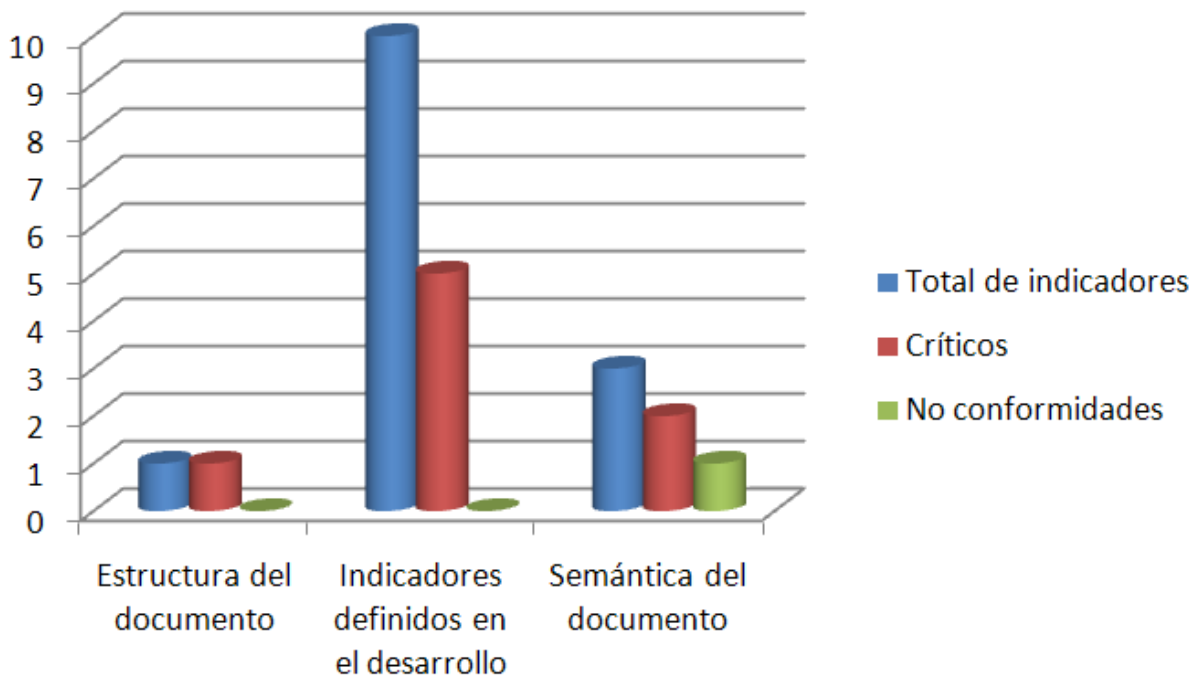


Figura 8 Resultado de la lista de chequeo

3.7.3 Resultados y discusión de las pruebas

Fueron realizadas tres iteraciones en la revisión de los casos de prueba, internas, del centro y calidad UCI, detectándose dos no conformidades las cuales fueron resueltas satisfactoriamente. De manera general el producto es evaluado de Bien debido a que no existieron indicadores críticos evaluados de mal en la aplicación de la lista de chequeo y se obtuvo la carta de aceptación por parte del cliente, quedando certificado que el sistema cumple con las necesidades del mismo.

3.8 Conclusiones

En el presente capítulo se describieron los elementos de implementación y pruebas necesarios para la construcción del MD tales como:

- Definición de la estructura de los datos a partir del modelo de datos físico, contando con dos esquemas: dimensiones y mart_balance.
- Diseño e implementación de los cubos OLAP quedando definidos ocho cubos, diez dimensiones y 15 medidas, necesarias para llevar a cabo la solución.
- Desarrollo del subsistema de visualización, determinándose ocho L.T con un total de 38 vistas de análisis.
- Elaboración de la lista de chequeo y casos de prueba, obteniéndose tres no conformidades resueltas de forma satisfactoria.

CONCLUSIONES GENERALES

Al culminar el desarrollo de la presente investigación se llegan a las siguientes conclusiones:

- Se realizó el análisis y diseño del mercado de datos Balance económico financiero, identificándose los requisitos que dieron solución a las necesidades de los usuarios. Quedando definida la estructura del mercado de datos.
- Se implementó el subsistema de visualización para el área de análisis Balance económico financiero, determinándose las vistas de análisis que apoyan el proceso de toma de decisiones en la dirección de Cuentas nacionales.
- Las pruebas realizadas a la aplicación permitieron validar la solución propuesta, se elaboraron los casos de prueba basados en casos de uso y la lista de chequeo, los mismos fueron aplicados por los especialistas, obteniéndose resultados satisfactorios en cada uno de ellos.

RECOMENDACIONES

Se recomienda para el mejoramiento del mercado de datos:

- Restringir el acceso a la información a los especialistas de las unidades presupuestadas y especialistas de organismo, debido a la sensibilidad que tiene la información de los Balances económicos financieros del país.

REFERENCIAS BIBLIOGRÁFICAS

- [1] **Kimball, Ralph.** *The Data Warehouse Lifecycle Toolkit. Segunda Edición.* 2008.
- [2] **Inmon, Willian H.** *Building the Data Warehouse.3rd Edition.* New York, USA : s.n., 2002. ISBN 0471081302.
- [3] **Casales Cabrera, Evelia María (Maestría en Ciencias e Ingeniería de la Computación).** Data Warehouse (Almacenes de Datos). [En línea] 2009. <http://hp.fciencias.unam.mx/~alg/bd/dwh.pdf>.
- [4] **Díaz Morales, Themis Patricia y Bermúdez Rodríguez, Jose Salvador.** *Diseño de un Datawarehouse para los Ensayos Clínicos que se gestionan en el Centro de Inmunología Molecular.* Universidad de las Ciencias Informáticas : s.n., 2010.
- [5] **Ortiz Sierra, Julio Ernesto.** *Diseño e Implementación de un Mercado de Datos para la Oficina Nacional de Estadísticas.* 2010.
- [6] El Portal de las Estadísticas Dominicanas. *ONE pone en línea Almacén Central de Datos del Sistema Estadístico Nacional.* [En línea] <http://www.one.gob.do/index.php?module=articles&func=display&aid=1377>.
- [7] **Galdiano, José Lázaro.** Datatur. Almacén de Datos para el Análisis y Difusión de la Información Estadística del Turismo en España. *Almacén de Datos para el Análisis y Difusión de la Información Estadística del Turismo de España.* [En línea] http://www.iet.tourspain.es/es-ES/fondodocumental/publicaciones/otras/Documents/FichaTecnicaDataturEsp_actual.pdf.
- [8] Consulta Interactiva de Datos en la Página de INEGI en Internet.Antecedentes e importancia. [En línea] <http://unstats.un.org/unsd/dnss/docViewer.aspx?docID=1552>.
- [9] **Cerezal Tamargo, Lic. Lourdes.** La revista del empresario cubano.Desarrollo computacional en el mundo de la Sanología. [En línea] http://www.betsime.disaic.cu/secciones/tec_mj_05.htm.
- [10] *REVISTA DE ARQUITECTURA E INGENIERÍA.* *Almacén de Datos para la Gestión Contable de la EMPAI.* **González Hidalgo, MSc.Ing. Gisel y Rizo Rizo, MSc.Lic. Emma.**
- [11] **Hernández López, Asnioby, Cárdenas Tandrón, Iván M y Peña Vazquez, Yanet.** Almacenes de datos aplicados a la seguridad ciudadana. [En línea] <http://www.uci.cu/files/investigaciones/vol1/SC%20Num%204.pdf>.
- [12] ETL-Tools.Info. *Arquitectura y concepto de un Almacén de Datos.* [En línea] http://etl-tools.info/es/bi/almacenedatos_esquema-constelacion.htm.

- [13] *Modelamiento multidimensional*. **Wolff, Carmen Gloria**. 2002.
- [14] **Moral Rubia, Jorge**. Introducción a los almacenes de datos. [En línea] marzo-abril 2009. https://www.icaei.es/publicaciones/anales_get.php?id=1653.
- [15] *La revista de tecnología y estrategia de negocio en Internet. La empresa multidimensional: OLAP*. **Borja, Ronda**.
- [16] *El impacto de las herramientas de inteligencia de negocios en la toma de decisiones de los ejecutivos*. **Calzada, Leticia y Abreu, José Luis**. septiembre 2009. ISSN 1870-557X.
- [17] **Fernández, Alarcón Vicenç**. *Desarrollo de sistemas de información. Una metodología basada en el modelado*. s.l. : Edicions UPC, 2006. ISBN 84-8301-862-4.
- [18] *COMPARACION ENTRE SISTEMAS DE GESTION DE BASES DE DATOS (SGBD) BAJO LICENCIAMIENTO LIBRE Y COMERCIAL*. **GONZALEZ CASTELLANOS, Ma. ARGENIS y ROJAS PABON, WILSON**. UNIVERSIDAD CATOLICA DE COLOMBIA. FACULTAD DE INGENIERIA DE SISTEMAS BOGOTA : s.n., 2005.
- [19] **Quiñones, Ernesto A**. INTRODUCCION A POSTGRESQL. [En línea] http://www.postgresql.org.pe/articles/introduccion_a_postgresql.pdf.
- [20] **Casanova, Jaime**. PostgreSQL. [En línea] 1996-2008. <http://archives.postgresql.org/pgsql-esfomento/2009-07/msg00000.php>.
- [21] **Alinga Ibarra, Antonio y Miami Flores, Marcos Agustín**. PostgreSQL-I.E.S.San Vicente. [En línea] 2008. <http://www.iessanvicente.com/colaboraciones/postgreSQL.pdf>.
- [22] *INGENIERÍA DEL SOFTWARE I Práctica 1. Trabajando con Visual Paradigm for UML*. **Sierra, María**. Univ. Cantabria - Fac de Ciencias : s.n.
- [23] **Fernández Fernández, Karina y Navarro Fonseca, Carlos**. *Análisis, diseño e implementación del mercado de datos para los indicadores sobre el Plan Turquino para la Oficina Nacional de Estadísticas*. s.l. : Universidad de las Ciencias Informáticas, 2010.
- [24] *Comparativa BI Open Source*. [En línea] junio 2010. http://www.dataprix.com/files/Comparativa_OSBI%20%281%29.pdf.
- [25] *Business Intelligence: Negocios Inteligentes para Empresas Inteligentes*. **Campazzo, Eduardo Nicolás y Santos, Virginia Inés**. Universidad Nacional de La Rioja : s.n.

[26] *Control de versiones con Subversion*. **Ben Collins, Sussman, W, Brian y Pilato, Michael C.** 2002-2003-2004.

BIBLIOGRAFÍA

Alberto Límia Navarro, Mabel Medina Rodriguez, Daulemys Rigo Portillo, Madelys Cuesta Villa, Yamila Mateu Romero, Julio Ernesto Ortiz, Osniel Hernandez Calvo, Marisleydis Socas Alvez , Anisley Delfino Rodríguez, Yonelbys Iznaga González, y otros. Metodología para el desarrollo de Soluciones de Almacenes de Datos e Inteligencia de Negocio.

Alinga Ibarra, Antonio y Miami Flores, Marcos Agustin. 2008. PostgreSQL-I.E.S.San Vicente. [En línea] 2008. <http://www.iessanvicente.com/colaboraciones/postgreSQL.pdf>.

Aumentar la productividad con las tecnologías innovadoras y productivas. [En línea] 2010. www.visual-paradigm.com.

Berríos, Br.Ericka Graciela Sevilla. GUÍA METODOLÓGICA PARA LA DEFINICIÓN Y DESARROLLO DE UN DATA WAREHOUSE. [En línea] 2003. <http://www.scribd.com/doc/39978465/GUIA-METODOLOGICA-PARA-LA-DEFINICION-Y-DESARROLLO-DE-UN-DATAWAREHOUSE>.

Business Intelligence: Negocios Inteligentes para Empresas Inteligentes. **Campazzo, Eduardo Nicolás y Santos, Virginia Inés.** Universidad Nacional de La Rioja : s.n.

Casales Cabrera, Evelia María (Maestría en Ciencias e Ingeniería de la Computación). Data Warehouse (Almacenes de Datos). [En línea] 2009. <http://hp.fciencias.unam.mx/~alg/bd/dwh.pdf>.

Casanova, Jaime. PostgreSQL. [En línea] 1996-2008. <http://archives.postgresql.org/pgsql-es-fomento/2009-07/msg00000.php>.

Cerezal Tamargo, Lic. Lourdes. La revista del empresario cubano. Desarrollo computacional en el mundo de la Sanología. [En línea] http://www.betsime.disaic.cu/secciones/tec_mj_05.htm.

COMPARACION ENTRE SISTEMAS DE GESTION DE BASES DE DATOS (SGBD) BAJO LICENCIAMIENTO LIBRE Y COMERCIAL. **GONZALEZ CASTELLANOS, Ma. ARGENIS y ROJAS PABON, WILSON.** UNIVERSIDAD CATOLICA DE COLOMBIA. FACULTAD DE INGENIERIA DE SISTEMAS BOGOTA : s.n., 2005.

Comparativa BI Open Source. [En línea] junio 2010. http://www.dataprix.com/files/Comparativa_OSBI%20%281%29.pdf.

Consulta Interactiva de Datos en la Página de INEGI en Internet. Antecedentes e importancia. [En línea] <http://unstats.un.org/unsd/dnss/docViewer.aspx?docID=1552>.

Control de versiones con Subversion. **Ben Collins, Sussman, W, Brian y Pilato, Michael C. 2002-2003-2004.**

DATA WAREHOUSING: Investigación y sistematización de conceptos. Hefesto: Metodología propia para la construcción de un Datawarehouse. **Ricardo Dario, Ing. Bernabeu. 2009.**

Díaz Morales, Themis Patricia y Bermúdez Rodríguez, Jose Salvador. *Diseño de un Datawarehouse para los Ensayos Clínicos que se gestionan en el Centro de Inmunología Molecular*. Universidad de las Ciencias Informáticas : s.n., 2010.

El impacto de las herramientas de inteligencia de negocios en la toma de decisiones de los ejecutivos. **Calzada, Leticia y Abreu, José Luis.** septiembre 2009. ISSN 1870-557X.

El Portal de las Estadísticas Dominicanas. *ONE pone en línea Almacén Central de Datos del Sistema Estadístico Nacional*. [En línea] <http://www.one.gob.do/index.php?module=articles&func=display&aid=1377>.

ETL-Tools.Info. *Arquitectura y concepto de un Almacén de Datos*. [En línea] http://etl-tools.info/es/bi/almacenedatos_esquema-constelacion.htm.

ETL-Tools.Info. [En línea] 2006-2011. http://etl-tools.info/es/bi/almacenedatos_arquitectura.htm.

Fernández Fernández, Karina y Navarro Fonseca, Carlos. 2010. *Análisis, diseño e implementación del mercado de datos para los indicadores sobre el Plan Turquino para la Oficina Nacional de Estadísticas*. s.l. : Universidad de las Ciencias Informáticas, 2010.

Fernández, Alarcón Vicenç. *Desarrollo de sistemas de información. Una metodología basada en el modelado*. s.l. : Edicions UPC, 2006. ISBN 84-8301-862-4.

Galdiano, José Lázaro. *Datatur. Almacén de Datos para el Análisis y Difusión de la Información Estadística del Turismo en España. Almacén de Datos para el Análisis y Difusión de la Información Estadística del Turismo de España*. [En línea] http://www.iet.tourspain.es/es-ES/fondodocumental/publicaciones/otras/Documents/FichaTecnicaDataturEsp_actual.pdf.

García, Gerardo Clemente. *Un Sistema para el Mantenimiento de Almacenes de Datos. Tesis doctoral.* Valencia, España : s.n., junio 2008.

Hernández López, Asnioby, Cárdenas Tandrón, Iván M y Peña Vazquez, Yanet. Almacenes de datos aplicados a la seguridad ciudadana. [En línea] <http://www.uci.cu/files/investigaciones/vol1/SC%20Num%204.pdf>.

INGENIERÍA DEL SOFTWARE I Práctica 1. Trabajando con Visual Paradigm for UML. **Sierra, María.** Univ. Cantabria - Fac de Ciencias : s.n.

Inmon, W.H. *Building the Data Warehouse. 2nd Edition.* New York : s.n., 1996.

Inmon, Willian H. *Building the Data Warehouse.3rd Edition.* New York, USA : s.n., 2002. ISBN 0471081302.

Kimball, Ralph. *The Data Warehouse LifecycleToolkit. Segunda Edición.* 2008.

Kimball, Ralph. *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling. 2nd Edition.* 2002.

La revista de tecnología y estrategia de negocio en Internet. La empresa multidimensional: OLAP. **Borja, Ronda.**

Microsoft Business Intelligence. España. [En línea] 2009. <http://www.microsoft.com/spain/bi/default.mspx>.

Modelamiento multidimensional. **Wolff, Carmen Gloria.** 2002.

Modelos Avanzados de Bases de Datos ALMACENES DE DATOS Y BASES DE DATOS XML. [En línea] <http://alarcos.inf-cr.uclm.es/doc/bbddavanzadas/08-09/FUNCIONALIDAD%204.pdf>.

Moral Rubia, Jorge. Introducción a los almacenes de datos. [En línea] marzo-abril 2009. https://www.icaei.es/publicaciones/anales_get.php?id=1653.

Oficina Nacional de Estadísticas. Cuba. [En línea] <http://www.one.cu>.

Ortiz Sierra, Julio Ernesto. *Diseño e Implementación de un Mercado de Datos para la Oficina Nacional de Estadísticas.* 2010.

Pentaho Open Source Business Intelligence. [En línea] <http://www.pentaho.com/>.

PostgreSQL Global Development Group. [En línea] <http://www.postgresql.org/about/>.

Quiñones, Ernesto A. INTRODUCCION A POSTGRESQL. [En línea] http://www.postgresql.org.pe/articles/introduccion_a_postgresql.pdf.

REVISTA DE ARQUITECTURA E INGENIERÍA. Almacén de Datos para la Gestión Contable de la EMPAI. **González Hidalgo, MSc.Ing. Gisel y Rizo Rizo, MSc.Lic. Emma.**

Servente, Magdalena y García Martínez, Dr. Ramón. Algoritmos TDIDT aplicados a la Minería de Datos Inteligente. Facultad de Ingeniería. Tesis de Grado en Ingeniería Informática. [En línea] 2002.

Sinnexus. [En línea] http://www.sinnexus.com/business_intelligence/datamart.aspx..

Sitio de PgAdmin. [En línea] <http://www.pgadmin.org/>.

The Apache Software Foundation. [En línea] 1999-2010. <http://tomcat.apache.org/>.

Zepeda Sánchez, Leopoldo Zenaido. 2008. *Metodología para el Diseño Conceptual de Almacenes de Datos.* 2008.

GLOSARIO DE TÉRMINOS

A continuación se presentan los términos que podrían resultar de difícil comprensión, nuevos al lector o de diversos significados dependiendo del contexto que se analice. Esta sección tiene como objetivo facilitar la comprensión del contenido expuesto en el documento.

AD: Almacén de datos también conocido como Data Warehouse (DWH). Un Almacén de datos reúne la información histórica generada por todos los distintos departamentos de una organización, orientada a consultas complejas y de alto rendimiento.

BI: Business Intelligence. La inteligencia de negocios une los datos, la tecnología, el análisis y el conocimiento humano para optimizar las decisiones empresariales y, finalmente, impulsar el éxito de una empresa.

Centros informantes: Los centros informantes son las empresas u organismos que suministran información a las oficinas de estadísticas en sus diferentes niveles.

Cuentas Nacionales: Es una de las direcciones de la Oficina Nacional de Estadística, se encarga de analizar la información referente a los Balance económicos financieros de todo el país y trabaja en conjunto con el Ministerio de Finanzas y Precios.

Dimensión: Característica de un hecho que permite su análisis posterior en el proceso de toma de decisiones y brinda una perspectiva adicional a un hecho dado.

Hecho: Operación que se realiza en el negocio la cual está estrechamente relacionada con el tiempo y es objeto de análisis para la toma de decisiones.

Indicador: Permiten clarificar y definir, de forma más precisa, objetivos e impactos. Son medidas verificables de cambio o resultado, diseñadas para contar con un estándar contra el cual evaluar, estimar o demostrar el progreso.

JDBC: Java DataBase Connectivity. Protocolo de conexión de Java a base de datos.

Lista de chequeo: Instrumento de medición y evaluación que consiste básicamente en un formulario de preguntas referentes al atributo de calidad que se está probando.

MD: Mercado de datos. Un mercado de datos es un subconjunto de un almacén de datos, tienen integrado el concepto de almacén de datos. Los mercados de datos tienen mayor utilidad para grupos más pequeños que deseen acceder a datos detallados.

MDX: Multi-Dimensional Expressions. Es un lenguaje de consulta para bases de datos multidimensionales sobre cubos OLAP, se utiliza en Inteligencia de negocios para generar reportes que apoyen la toma de decisiones basados en datos históricos.

NAE: Es un nomenclador, emitido a nivel mundial, para la organización de las actividades económicas de las empresas.

No conformidad: Defecto, error o sugerencia que se le hace al equipo de desarrollo una vez encontrada alguna dificultad en lo que se está evaluando.