

Tipo de artículo: Artículo original
Temática: Desarrollo de aplicaciones informáticas
Recibido: 04/03/17 | Aceptado:10/05/17 | Publicado: dd/mm/aa

Mercado de Datos para el Departamento de Mercadotecnia perteneciente a la Dirección de Transferencia de Tecnología y Conocimiento

Data Mart for the Marketing Department belonging Technology and Knowledge Transfer Directorate

Midalis Rodríguez Rosado^{1*}, Jorge Alfonso Torres², Elian Eduardo Castell Legrá³

¹ Universidad de las Ciencias Informáticas. Carretera San Antonio de los Baños km 2 1/2. mrosado@uci.cu

² Universidad de las Ciencias Informáticas. Carretera San Antonio de los Baños km 2 1/2.

³ Universidad de las Ciencias Informáticas. Carretera San Antonio de los Baños km 2 1/2. eecastell@uci.cu

* Autor para correspondencia: mrosado@uci.cu

Resumen

La Universidad de las Ciencias Informáticas, cuenta con la Dirección de Transferencia de Tecnología y Conocimiento, la que tiene asociado un Departamento de Mercadotecnia que se encarga de la comercialización y la negociación de las diferentes ofertas que brinda la universidad. Actualmente, en el departamento elaboran informes para la toma de decisiones que incluyen tablas resumen, gráficos, listados, cálculos de indicadores complejos y análisis estadísticos. Esta tarea es afectada por el manejo manual de los datos obtenidos desde diferentes fuentes, causando su pérdida y duplicación, además del consumo de tiempo. El presente trabajo de diploma tiene como objetivo general desarrollar un Mercado de Datos para dicho departamento que contribuya al proceso de toma de decisiones. El mismo permitirá obtener, de forma organizada y rápida, los datos a través de listados y gráficos. Para guiar el proceso de desarrollo de la solución se utilizó la Metodología para el desarrollo de proyectos de Almacenes de Datos, apoyándose en las herramientas Visual Paradigm 8.0, PostgreSQL 9.4 y PgAdmin III 1.20. Para la limpieza de datos se emplea el Data Cleaner 3.1 y Data Integration 6.0 para los procesos de extracción, transformación y carga. Para el diseño de cubos, implementación de reportes y presentación de la información se utiliza el Schema Workbench 3.11, Report Designer 6.0 y el BI Server 6.0. Después del análisis, diseño e implementación de la

solución, se obtuvo un Mercado de Datos poblado al que se le aplicaron un conjunto de pruebas para verificar su correcto funcionamiento.

Palabras clave: Almacén de Datos, Mercado de Datos, toma de decisiones

Abstract

The University of Computer Science has the Technology and Knowledge Transfer Directorate, which has a Marketing Department associated with the commercialization and negotiation of the different offers offered by the university. Currently, the department specialists produce decision-making reports which include summary tables, charts, lists, complex indicator calculations and statistical analyzes. This task is affected by the manual handling of the data obtained from different sources, causing its loss and duplication, besides the time consumption. This diploma work is part of the development of a Data Market for this department that contributes to the decision-making process. It will allow to obtain, in an organized and fast, the data through lists and graphs. To guide the process of developing the solution, the Methodology for the development of Data Warehouse projects was used to guide the process, using Visual Paradigm 8.0, PostgreSQL 9.4, PgAdmin III 1.20, Data Cleaner 3.1, Data Integration 6.0, Schema Workbench 3.11, Report Designer 6.0, and BI Server 6.0. The analysis, design and implementation of the solution was performed, resulting in a populated Data Market, which meets the characteristics identified by the client. A set of tests were applied to verify the correct operation of the market, throwing several nonconformities resolved in their entirety.

Keywords: Data Mart, Data Warehouse, decision making.

Introducción

Las Tecnologías de la Información y las Comunicaciones (TIC), han transformado el mundo que nos rodea, siendo piezas fundamentales para la superación y desarrollo de un país. Aportan herramientas específicas, que se incluyen de forma responsable en el día a día. Se puede considerar que la incorporación de las mismas es un proceso imparabile, por lo que hay que estar a la vanguardia para adaptarse a las nuevas tecnologías que vayan surgiendo. Su implantación permite evolucionar la manera de trabajar, agilizando los procesos existentes, de manera que contribuya a la informatización de la sociedad.

Con el uso de computadoras en los distintos sectores de la sociedad, se ha visto un mayor crecimiento en la capacidad de generar y almacenar información. Mientras mayor es la capacidad de los dispositivos de almacenamiento de datos,

mayor es el interés de los humanos por guardar grandes volúmenes de datos. Esto trae consigo dificultades a la hora de extraer conocimiento, realmente útil, de la información almacenada. Las empresas e instituciones se han dado cuenta de la necesidad de procesar dicha información, que es realmente eficaz, si los datos están ordenados, analizados y transformados de modo que permitan resolver problemas específicos.

Cuba cuenta con varias empresas e instituciones encargadas de informatizar la sociedad y exportar productos y servicios a otros países. Una de estas instituciones es la Universidad de las Ciencias Informáticas (UCI), la que tiene como misión, formar profesionales comprometidos con su Patria y altamente calificados en la rama de la Informática, capaces de producir aplicaciones y servicios informáticos, a partir de la vinculación estudio-trabajo, como modelo de formación y servir de soporte a la industria cubana del software.

La UCI cuenta con una Dirección de Transferencia de Tecnología y Conocimiento (DTTC), la que tiene como misión, gestionar la transferencia de tecnologías y conocimientos de la institución hacia la sociedad cubana y al exterior del país, protegiendo la propiedad intelectual de los productos y servicios resultantes de su actividad y estableciendo la política de calidad institucional. Así mismo, coordina y organiza la elaboración y actualización de la planeación estratégica, los manuales de organización y los objetivos de trabajo, controlando su cumplimiento. Cuenta con tres departamentos: el Departamento de Desarrollo y Calidad, el Departamento de Aseguramiento a la Gestión y el Departamento de Mercadotecnia.

Este último es el encargado de la gestión de la comercialización y la negociación de los servicios académicos, editoriales e informáticos, así como de las aplicaciones que son desarrolladas por los centros de producción de la universidad. Para cumplir con dichos objetivos los especialistas que allí laboran son responsables de ejecutar los siguientes procesos:

1. Registro en el Plan de Transferencia y Tecnología del Conocimiento (PTTC).
2. Registro de visitas al salón de exposiciones.
3. Registro de expedientes comerciales.
4. Registro en el Centro Nacional de Derecho de Autor (CENDA).

La información que presenta el departamento data desde 2008 y es almacenada digitalmente en ficheros excel (.xls) con una periodicidad mensual y anual, indistintamente, para cada uno de los procesos. El primero, guarda todo lo referente a los contratos que se llevan a cabo en la institución; el segundo, lo relacionado con las visitas al salón de exposiciones; el tercero, los expedientes comerciales realizados, mientras que el cuarto, archiva los productos u obras literarias a inscribir en el CENDA para el derecho de autor.

A partir de la consulta y el análisis de los datos, los especialistas generan informes mensuales en formato PDF (.pdf) solicitados por el jefe de departamento para la toma de decisiones. Con la utilización de las herramientas de Microsoft Office, la cual posee una licencia privativa y su uso requiere de grandes costos para la institución, crean tablas resumen, gráficos o listados, de manera que la información sea mostrada desde diferentes perspectivas y aristas de análisis.

Realizar estos informes resulta engorroso por parte de los especialistas, ya que, para la creación de cada uno de ellos deben analizar y copiar los datos contenidos en los .xls para un documento .doc con sus tablas y gráficos correspondientes. Se hacen, además, cálculos de indicadores complejos y análisis estadísticos, comparativos y de tendencias de los mismos en diferentes periodos de tiempo. Como consecuencia se genera la pérdida y duplicación de los datos, además de la utilización de tiempo excesivo en la conformación y entrega de la información solicitada.

El jefe de departamento analiza la información enviada de cada uno de los procesos para realizar investigaciones de mercado y análisis de factibilidad; tarea que resulta compleja, al no existir una correspondencia entre los resultados que brindan los especialistas, pues esta no se encuentra centralizada e integrada. Las problemáticas, antes enunciadas, permiten a los autores declarar que no son pocas las dificultades a las que se enfrentan los diferentes especialistas para elaborar los informes, que, de forma mensual, deben rendir al jefe de departamento; lo que genera, afectaciones en el proceso de toma de decisiones en el Departamento de Mercadotecnia perteneciente a la DTTC.

Después de un análisis de la situación problemática, queda conformado el **problema de la investigación** mediante el cuestionamiento de: ¿Cómo contribuir al proceso de toma de decisiones en el Departamento de Mercadotecnia perteneciente a la DTTC?

Para darle solución al problema planteado, se define como **objetivo general**: desarrollar un Mercado de Datos para el Departamento de Mercadotecnia perteneciente a la DTTC que contribuya al proceso de toma de decisiones.

El resultado esperado del presente trabajo es el desarrollo de un Mercado de datos poblado que apoye el proceso de toma de decisiones en el Departamento de Mercadotecnia perteneciente a la DTTC.

Materiales y métodos

A continuación, se explican los métodos de la metodología de la investigación científica empleados para la obtención de información y el futuro desarrollo del Mercado de Datos.

Métodos Teóricos

- ❖ **Histórico – lógico:** para el desarrollo del presente trabajo de diploma se realizó un estudio de los fundamentos teóricos relacionados con los Mercados de Datos en cuanto a definiciones, ventajas, desventajas y características.
- ❖ **Analítico – sintético:** este método se utilizó para el análisis de documentos, materiales y temas relacionados con el desarrollo de los Mercado de Datos y de elementos claves para el Departamento de Mercadotecnia. Permitted definir los conceptos fundamentales del tema y se usó para especificar la metodología a utilizar, así como los elementos que conforman un Mercado de Datos. La síntesis permitió realizar un análisis del impacto de los resultados de la solución desarrollada como parte de la validación y cumplimiento de los objetivos de la presente investigación.

Métodos empíricos

- ❖ **Entrevistas:** se utilizó para la recolección de información mediante las entrevistas no estructuradas a los clientes, lo que permitió la recogida de datos durante el proceso de ingeniería de requisitos y la consulta de materiales y normativas referentes al departamento.

Metodología de desarrollo

Para el desarrollo de la presente investigación se decidió utilizar la “Metodología para el desarrollo de proyectos de Almacenes de datos”, definida por la MCs. Yanisbel González Hernández en su tesis de maestría. La misma es una combinación de varias metodologías. Toma como base la metodología de Kimball para definir los aspectos específicos del desarrollo de AD. Para incorporar los principios básicos que permiten una adecuada gestión del proyecto, utiliza la Guía para los Fundamentos de la Dirección de Proyectos (Guía del PMBOK). Los temas asociados a CMMI (*Capability maturity model integration*) se incorporan a partir del Programa de Mejora, por lo tanto, hereda algunos de sus enfoques, artefactos y actividades. Se adapta a las características de trabajo específicas del centro DATEC, por lo que se escoge para todo desarrollo de AD en dicho centro. (González, 2013)

Esta metodología de desarrollo está dividida en siete fases y un flujo de trabajo. Algunas de estas fases podrán ser implementadas de forma paralela como es el caso de la fase de requisitos y arquitectura, además durante la fase de diseño e implementación, podrán desarrollarse varios componentes al mismo tiempo, lo que permite un desarrollo más ágil. Para el desarrollo de la presente investigación solo se va a tomar en cuenta las cinco primeras fases y el flujo de trabajo gestión de proyecto para guiar la solución de forma paulatina. A continuación, la figura 1 muestra las fases de la metodología con su flujo de trabajo (González, 2013).

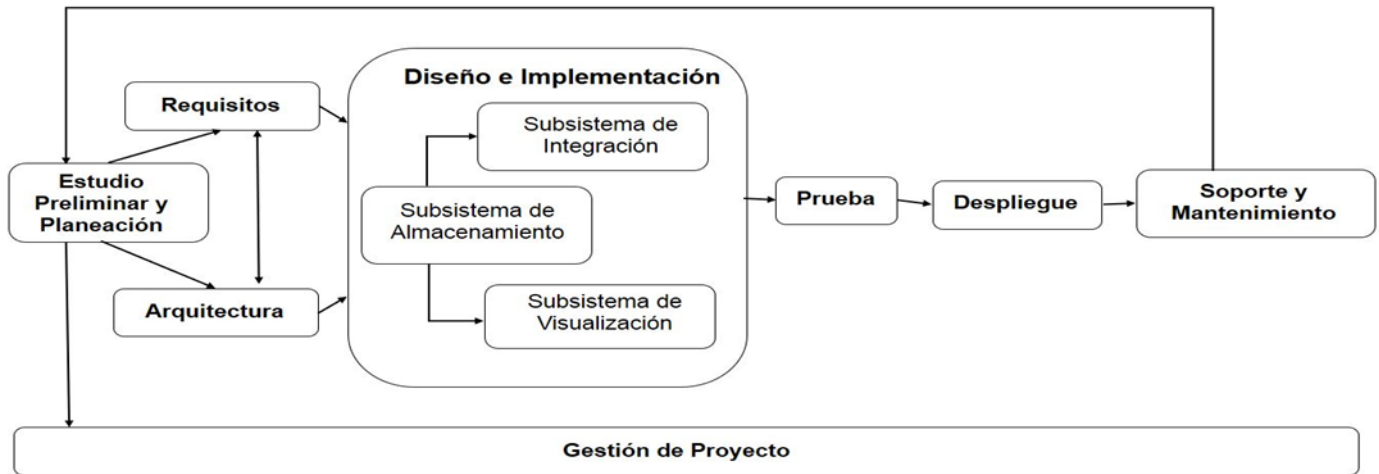


Fig. 1: Metodología de desarrollo (Elaboración Propia).

Herramientas de desarrollo

- ❖ **Visual Paradigm for UML Community Edition 8.0:** es una herramienta para desarrollo de aplicaciones utilizando modelado UML (Lenguaje Unificado de Modelado) ideal para el desarrollo orientado a objetos. Visual Paradigm ofrece navegación intuitiva entre la escritura del código y su visualización. Potente generador de informes en formato PDF/HTML. Ambiente visualmente superior de modelado. Sofisticado diagramador automático de layout y sincronizador de código fuente en tiempo real. (Targetware, 2016)
- ❖ **PostgreSQL 9.4:** es un sistema de gestión de base de datos relacional orientado a objeto, distribuido bajo licencia BSD (Berkeley Software Distribution) y su código fuente está disponible libremente. Es el sistema de gestión de base de datos de código abierto más potente del mercado, sus últimas versiones son tan poderosas como las de los sistemas de gestión de base de datos comerciales. Utiliza un modelo cliente/servidor y usa multiprocesos en vez de multihilos para garantizar la estabilidad del sistema. Un fallo en uno de los procesos no afectará el resto y el sistema continuará funcionando. (Martínez, 2010)
- ❖ **PgAdmin III 1.20:** es una aplicación de diseño y gestión de base de datos PostgreSQL. Está escrito en C++, multiplataforma lo que permite que se ejecute en múltiples plataformas. La aplicación se puede utilizar para manejar PostgreSQL 7.3 y superiores y funciona sobre casi todas las plataformas. Este software fue diseñado para responder a las necesidades de todos los usuarios, desde la escritura de simples consultas SQL a la elaboración de base de datos complejas. (PostgreSQL, 2016)

- ❖ **Data Cleaner 3.1:** esta herramienta permite descubrir y analizar la calidad de sus datos. Encuentra los patrones, valores faltantes, conjuntos de caracteres y otras características de los datos. Realiza comparaciones de tablas y columnas, lo que permite verificar la consistencia y veracidad de los datos. Esta herramienta permite analizar la calidad de la información que se encuentra almacenada en fuentes como: base de datos Oracle, MySQL, PostgreSQL y ficheros excel, XML. (Sorensen, 2012)
- ❖ **Pentaho Data Integration (PDI) 6.0:** herramienta de código abierto y multiplataforma, prepara y combina datos para crear una imagen completa del negocio y generar información útil. La plataforma proporciona datos precisos y preparados para el análisis de los usuarios finales desde cualquier fuente, permite la lectura y escritura de base de datos, el mapeo de valores y la normalización de los datos en distintas filas y viceversa con herramientas visuales para eliminar la codificación y la complejidad. (Pentaho, 2016)
- ❖ **Pentaho Schema Workbench (PSW) 3.11.0.0:** es una herramienta de desarrollo multiplataforma que permite crear, modificar y publicar un esquema de Mondrian1. Permite crear todos los objetos tales como, esquemas, cubos, dimensiones y métricas. Tiene dos áreas: la zona donde se encuentra la estructura jerárquica del esquema OLTP y la zona de edición de las propiedades de cada elemento. Presenta un menú superior para crear cubos, dimensiones, métricas, miembros calculados subconjuntos y roles, así como operaciones estándar como cortar, copiar y pegar. (Díaz, 2010)
- ❖ **Pentaho Report Designer (PRD) 6.0:** es una herramienta de diseño desktop, que permite definir y construir informes, y luego publicarlos en el portal de BI para que puedan ser ejecutados por los usuarios. Trabaja con múltiples orígenes de datos (JDBC, Olap4J, Pentaho Análisis, Pentaho Data Integration, XML). PRD es un generador de informe que se divide en secciones o grupos de datos en los que los elementos del informe pueden ser posicionados. El resultado de los informes que se vayan diseñando se puede ver con las opciones de previsualización, y permite la salida de resultados en diferentes formatos como PDF, HTML, XLS, RTF y CSV. (Espinosa, 2010)
- ❖ **Pentaho BI Server 6.0:** provee de servicios críticos incluyendo programación, seguridad, integración automatización y flujo de trabajo, proporcionando habilidades a los usuarios finales de Pentaho y proporcionando un lugar central para administrar y mantener el despliegue de la empresa BI. Posee integración con procesos de negocio, administra y programa reportes, administra seguridad de usuarios asegurando la escalabilidad, integración y portabilidad. Puede trabajar sobre las BD: vía JDBC, IBM DB2, Microsoft SQL Server,

¹ Es un servidor/motor OLAP escrito en java que está licenciado bajo la Eclipse Public License y fue adquirido por Pentaho en 2005.

MySQL, Oracle, PostgreSQL sin importar el sistema operativo. Repositorio de datos basado en XML. Todos los componentes están expuestos vía Web Services para facilitar la integración con Arquitecturas Orientadas a Servicios. (Gravitar, 2016)

Base de datos multidimensional

Las BD multidimensionales, proveen una estructura que permite tener acceso flexible a los datos, para explorar y analizar sus relaciones, y resultados consiguientes. Estas se pueden visualizar como un cubo multidimensional, donde las variables asociadas existen a lo largo de varios ejes o dimensiones, y la intersección de las mismas representan la medida, indicador o el hecho que se está evaluando. Para la implementación de las mismas, según (Bernabeu, 2010) existen variantes de modelamiento y de implementación.

Variantes de modelamiento

- ❖ **Esquema en estrella (*Star Scheme*):** el esquema en estrella consta de una tabla de hechos central y de varias tablas de dimensiones relacionadas a esta, a través de sus respectivas claves.
- ❖ **Esquema copo de nieve (*Snowflake Scheme*):** este esquema representa una extensión del modelo en estrella cuando las dimensiones se organizan en jerarquías de dimensiones. Consta de una tabla de hechos central que está relacionada con una o más tablas de dimensiones, quienes, a su vez, pueden estar relacionadas o no con una o más tablas de dimensiones.
- ❖ **Esquema constelación o copo de estrellas (*Starflake Scheme*):** un esquema de constelación está compuesto por una serie de esquemas en estrella. Consta de varias tablas de hechos que comparten dimensiones iguales.

Variantes de implementación

- ❖ **ROLAP (*Relational On Line Analytic Processing*):** este tipo de organización física se implementa sobre tecnología relacional, pero disponen de algunas facilidades para mejorar el rendimiento. Es decir, ROLAP cuenta con todos los beneficios de una Sistema Gestor de Base de Datos (SGBD) relacional a los cuales se les provee extensiones y herramientas para poder utilizarlo como un Sistema Gestor de AD.
- ❖ **MOLAP (*Multidimensional On Line Analytic Processing*):** el objetivo de los sistemas MOLAP es almacenar físicamente los datos en estructuras multidimensionales, de manera que la representación externa y la interna coincidan. Para ello, se dispone de estructuras de almacenamiento específicas (*Arrays*) y técnicas de compactación de datos que favorecen el rendimiento del depósito de datos.
- ❖ **HOLAP (*Hybrid On Line Analytic Processing*):** constituye un sistema híbrido entre MOLAP y ROLAP, que combina estas dos implementaciones para almacenar algunos datos en un motor relacional y otros en un BD multidimensional.

Resultados y discusión

Para construir la base del Mercado de Datos para el Departamento de Mercadotecnia fue necesario identificar las necesidades de información de cada especialista para poder apoyar al proceso de toma de decisión en la dirección. Para esto se separó la información en áreas de análisis según la temática abordada, cada una de ellas, representa un proceso en específico ya que los cuatro son totalmente diferentes. A continuación, se muestran las áreas de análisis y su proceso correspondiente. Ver tabla 1.

Tabla 1: Distribución de las áreas de análisis.

Áreas de análisis	Procesos
PTTC	Registro de contratos en el PTTC.
Visitas al Salón de Exposiciones	Registro de Visitas al Salón de Exposiciones.
Expedientes Comerciales	Registro de Expedientes Comerciales.
Productos en el CENDA	Registro de Productos en el CENDA.

Fueron identificados un total de 2 requisitos funcionales, 13 no funcionales y 16 requisitos de información agrupados en las cuatro áreas de análisis antes mencionadas. Como ejemplos de estos últimos se tiene:

1. Obtener listado de contratos por centro, estado legal, encargado y tiempo dado el centro, el tiempo (año y mes), el estado legal y el encargado.
2. Obtener el monto total en CUC por centro, estado legal, encargado y tiempo dado el centro, el tiempo (año y mes), el estado legal y el encargado.

Variables de entrada: Estado legal, Encargado, Centro, tiempo (año y mes)

Variable de salida: Cantidad de contratos, Monto total en CUC, Monto total en CUP, Monto total en MT, Listado de contratos, Por ciento de contratos.

El proceso de obtención de las variables de salida será guiado a través de la arquitectura definida para la implantación del Mercado de datos, la misma consta de tres subsistemas fundamentales (Fig. 2).



Fig. 2: Arquitectura del MD.

El **Subsistema de Almacenamiento** es el encargado de almacenar toda la información recogida de las fuentes de datos correspondientes a los procesos que se llevan a cabo en el Departamento de Mercadotecnia. Después de los análisis correspondientes y la identificación de los RI según las necesidades de los clientes, fueron creadas cuatro tablas de hechos agrupadas en el esquema **mart_mercadotecnia**, y 19 tablas de dimensiones agrupadas en el esquema **dimension**. (Fig. 3)

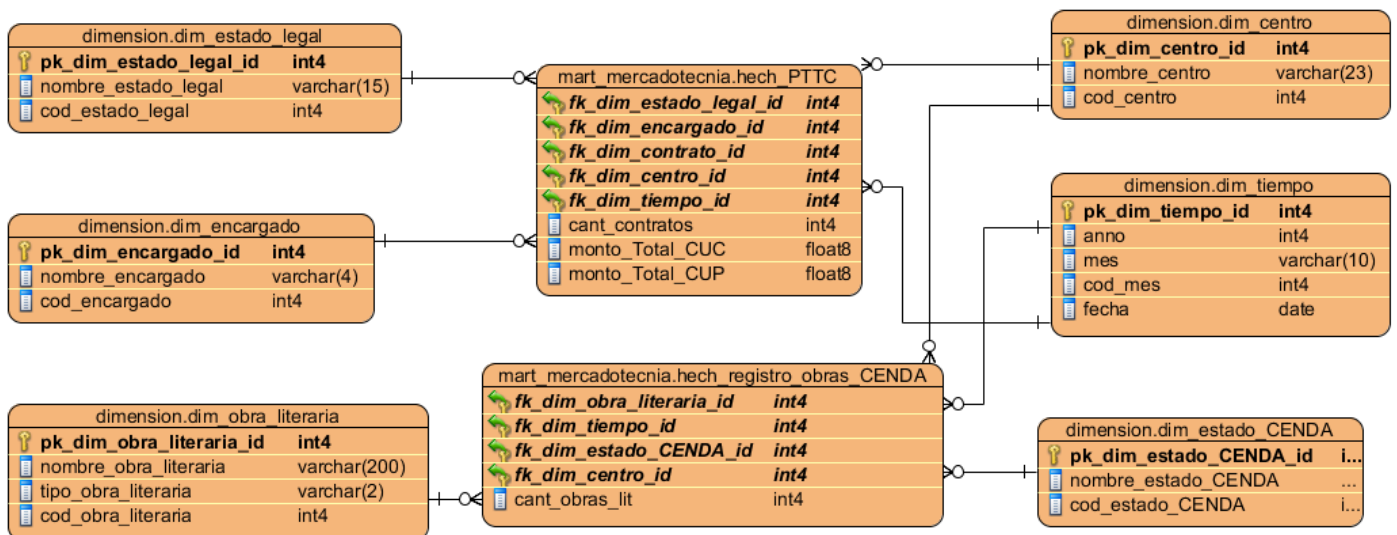


Fig. 3: Modelo de datos reducido.

En el **Subsistema de Integración**, para poder extraer los datos desde los ficheros fuentes, para luego manipularlos, integrarlos, transformarlos y cargarlos fue identificado ETL (Extracción, Transformación y Carga) como estrategia de integración y se realizaron 19 transformaciones para las dimensiones y 4 para los hechos. En la figura 4 se muestra un ejemplo de transformación.

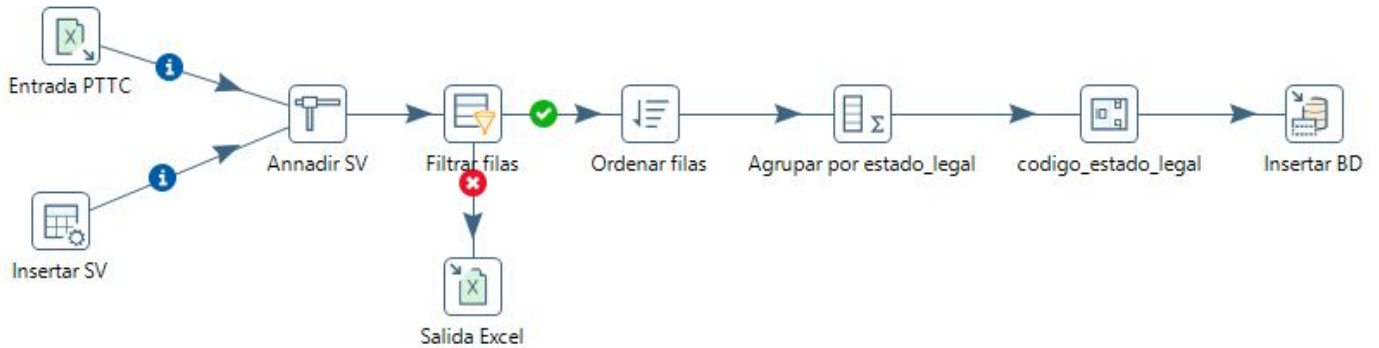


Fig. 4: Proceso de integración de datos para la dimensión “estado legal”.

Para el MD del Departamento de Mercadotecnia se realizaron tres trabajos, cuyo objetivo consiste en evitar cargar individualmente las transformaciones. La figura 5 (Fig. 5) muestra el trabajo para la carga de los hechos mientras que la figura 6 (Fig. 6) lo hace para la carga de las dimensiones.

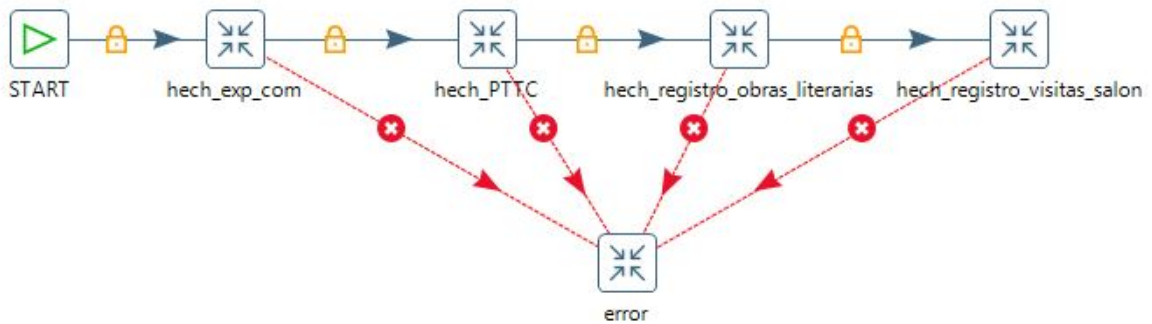


Fig. 5: Trabajo para la carga de los hechos.

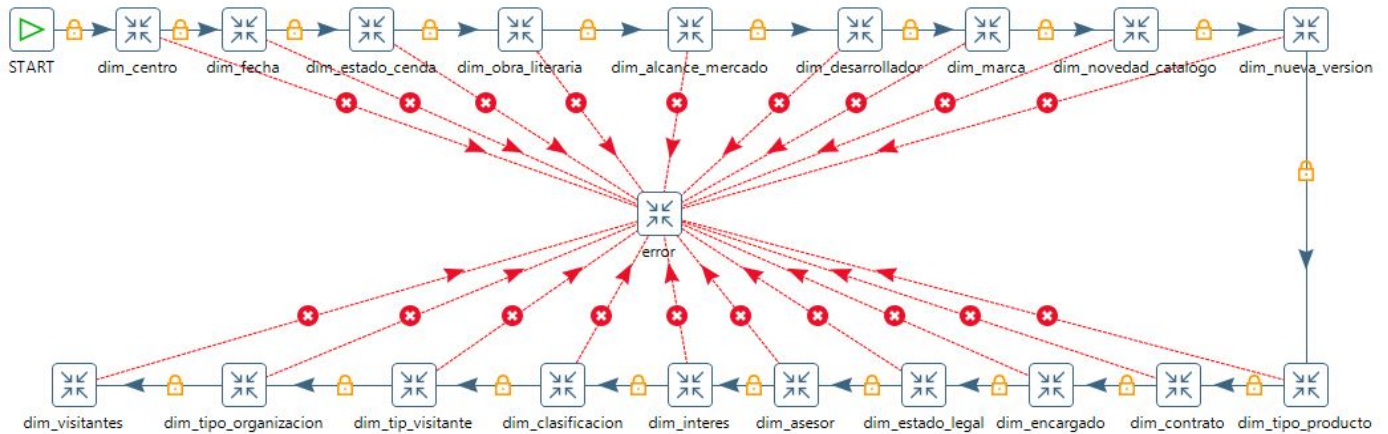


Fig. 6: Trabajo para la carga de las dimensiones.

Luego de haber concluido con los subsistemas anteriores, se procede a la implementación del **Subsistema de Visualización**, que es el responsable de mostrar el resultado final a los usuarios. Para esto se realizan los cubos OLAP, donde se definieron las dimensiones, medidas y niveles de jerarquía que conforma el esquema **mart_mercadotecnia**, creándose un cubo por cada temática (Fig. 7), la capa de visualización que cuenta con un A.A.G, cuatro A.A y cuatro LT, estos últimos contienen tres carpetas referentes a las vistas de análisis, reportes personalizados y cuadros de mando (Fig. 8) y por último, los reportes candidatos, los cuales contienen la información referente a los pedidos de información establecidos por el cliente. Estos reportes fueron elaborados mediante consultas SQL y MDX, utilizando los términos hechos, dimensiones y medidas (Fig. 9).

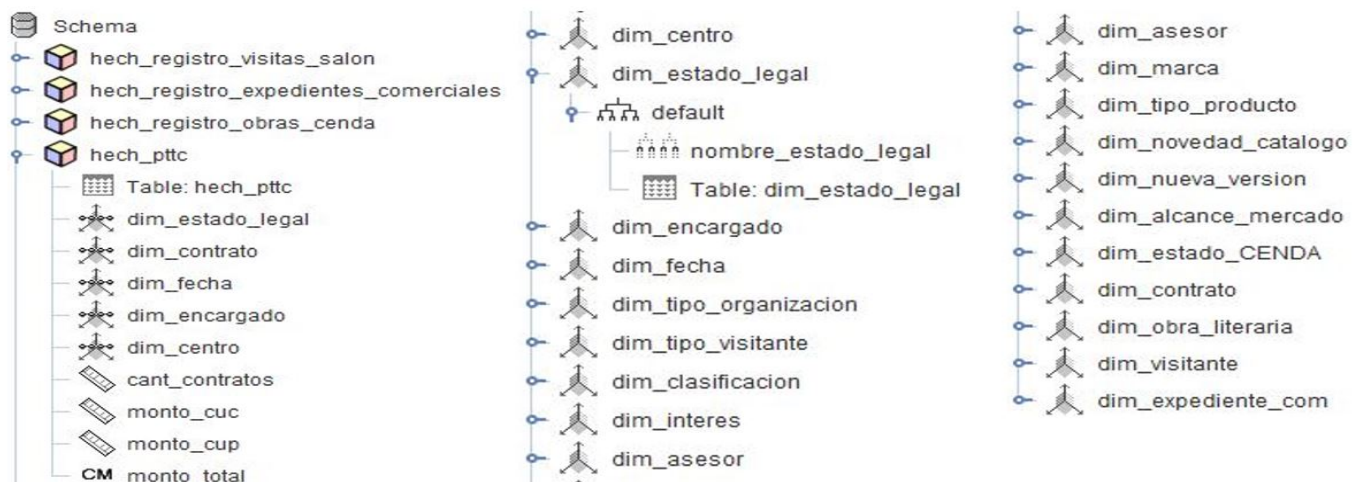


Fig. 7: Implementación de los cubos OLAP.

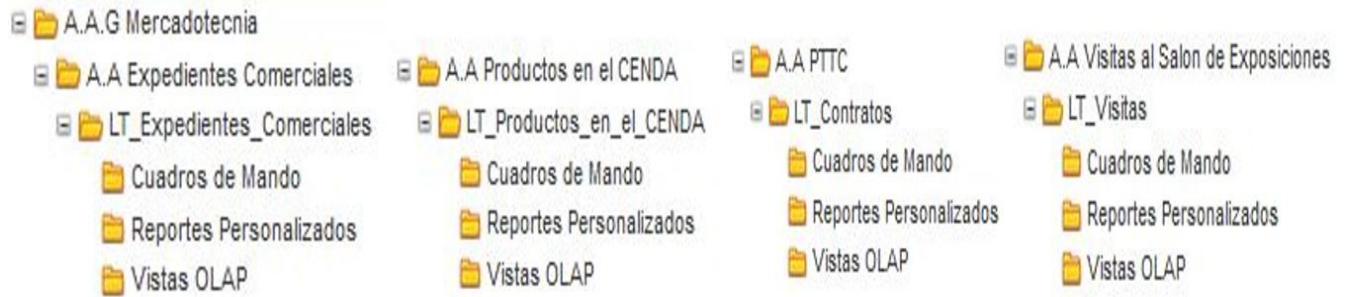


Fig. 8: Implementación de la capa de visualización.

		Fecha				
		+ Todos				
		Estado Legal				
		Detenido		En Ejecución		Facturado
		Encargado		Encargado		Encargado
		SCT3	SCT1	SCT3	SCT2	SCT1
		Medidas	Medidas	Medidas	Medidas	Medidas
Contrato	Centro	• Cant_contratos	• Cant_contratos	• Cant_contratos	• Cant_contratos	• Cant_contratos
+ Todos	+ Todos	1	11	16	7	1
	CEGEL					1
	CEIGE	1		1	1	
	CESOL				1	
	CIDI				1	
	CIGED		2		1	
	CISED		1			
	FORTES			2		
	Soporte		8	13	3	

Fig. 9: Vista de análisis referente al PTTC (Elaboración propia).

Para presentar la información a los usuarios finales fueron creados, mediante gráficos, los dashboard, los que representan en alto nivel el contenido de BI. La figura 10 corresponde a un dashboard realizado para el A.A PTTC. Para que este muestre su información, primero se tiene que seleccionar el rango de fecha (por año) deseado. A continuación, se muestra un ejemplo que evidencia la cantidad de contratos en el año 2015, por centro, por encargado y según su estado legal.



Fig. 10: Dashboard correspondiente al PTTC (Elaboración propia).

Pruebas MD para el Departamento de Mercadotecnia

La realización de las pruebas al MD para el Departamento de Mercadotecnia se sustentó en el Modelo V (figura 11), utilizado por el centro DATEC para garantizar el buen funcionamiento y la calidad de los productos de AD. Dicho modelo cuenta con un vértice común en la codificación, una rama izquierda donde se muestran las actividades relacionadas con el análisis y diseño del producto en cuestión y una rama derecha que muestra las actividades relacionadas con el aseguramiento de la calidad mediante los tipos de pruebas aplicadas.

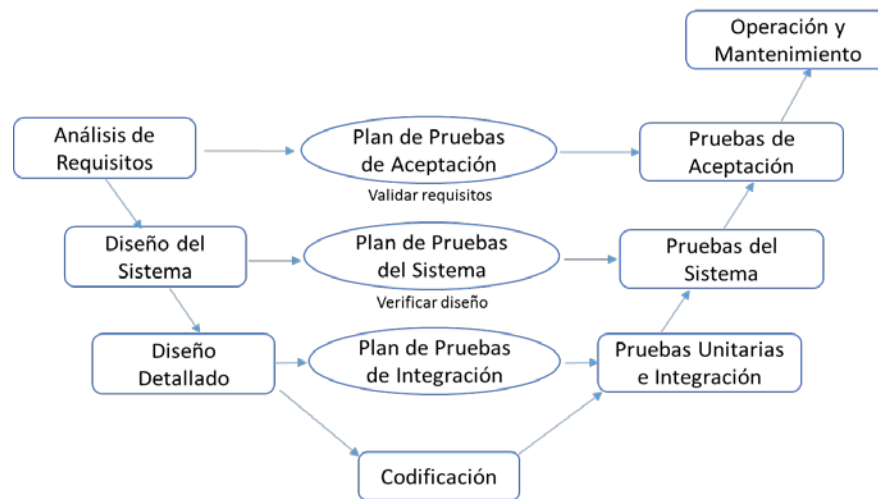


Fig. 11: Modelo V.

Tipos de Pruebas

Pruebas unitarias: tienen como objetivo verificar la funcionalidad y estructura de cada componente, de forma individual. Estas pruebas son realizadas por el programador a medida que implementa los componentes.

Pruebas de integración: verifica el correcto ensamblaje entre dos o más componentes una vez que han sido probados, unitariamente.

Pruebas de sistemas: estas buscan discrepancias entre el programa y sus objetivos o requerimientos, enfocándose en los errores hechos durante la transición del proceso al diseñar la especificación funcional.

Pruebas de aceptación: el objetivo de las pruebas de aceptación es comprobar si el software cumple con las expectativas del cliente.

La figura 12 muestra un resumen de las NC detectadas en los niveles de pruebas: unidad e integración, sistema y aceptación, en las que se detectaron un total de, cinco, 11 y cero NC, respectivamente. Estas NC fueron resueltas satisfactoriamente.

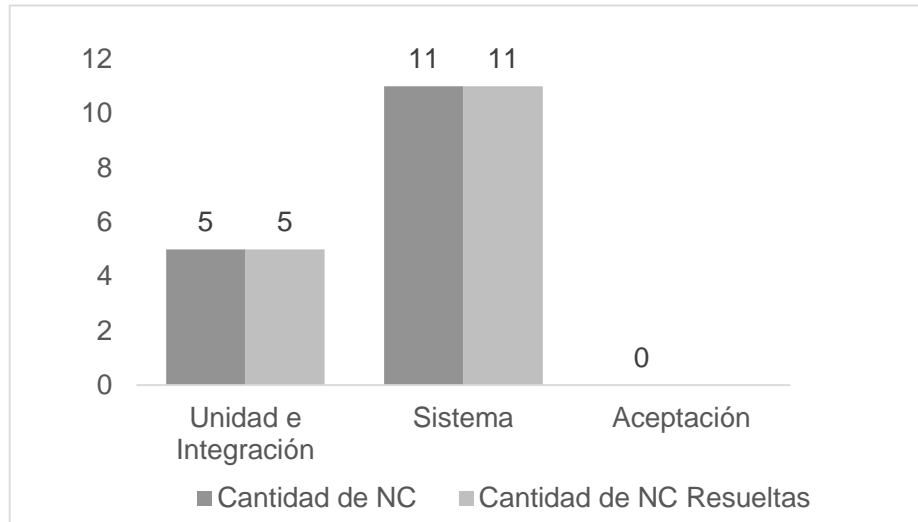


Fig. 12: Resultados de las pruebas.

Valoración económica y aporte social

Una vez aplicada la solución en el Departamento de Mercadotecnia perteneciente a la DTTC, el mismo podrá obtener su información de forma organizada a través de listados y gráficos, facilitando así la obtención de información referente a sus procesos. Esto permitirá obtener numerosas ventajas en cuanto a la veracidad de la información y la entrega de los informes a los superiores.

A continuación, se muestran los elementos que fundamentan el impacto que tendrá la solución una vez aplicada:

- ❖ Disminución del tiempo empleado para la creación de los informes.
- ❖ La información obtenida de los reportes posee alto nivel de fiabilidad.
- ❖ Apoya el proceso de toma de decisiones en la entidad.
- ❖ Mayor comprensión en cuanto a las necesidades de los clientes.
- ❖ Facilita el trabajo de los especialistas en cuanto a la creación de los reportes con solo un par de clic.
- ❖ Avances en cuanto la informatización del departamento en el uso de tecnologías para mejorar la calidad del almacenamiento multidimensional, procesos de extracción, transformación, calidad y carga de datos, así como para la visualización y navegación multidimensional de la información a través de herramientas OLAP y de creación de reportes.

Conclusiones

Al finalizar el proceso de desarrollo del MD para el Departamento de Mercadotecnia perteneciente a la DTTC se puede afirmar que se le ha dado cumplimiento de forma satisfactoria al objetivo general del presente trabajo de diploma, obteniéndose los resultados que se esperaban. Es por ello que se llega a las siguientes conclusiones:

- ❖ El estudio de los fundamentos teóricos permitió comprender mejor el tema tratado. Se seleccionó para guiar el proceso de desarrollo, la Metodología para el desarrollo de proyectos de AD, la que se adapta a las características de trabajo específicas del centro DATEC, lo que permite un desarrollo más ágil. La selección de las herramientas basadas en software libre permitió la correcta implementación de la solución con interfaces amigables para el usuario.
- ❖ El proceso de entrevistas con el cliente permitió definir los requisitos de la solución, identificándose así 16 RI agrupados en cuatro CUI, dos RF agrupados en dos CUF, 13 RNF y 26 RN. Se definió la arquitectura de la solución identificándose tres subsistemas. El subsistema de almacenamiento permitió identificar cuatro tablas de hechos y 19 tablas de dimensiones dando paso al diseño del modelo lógico de datos que permitió la implementación del modelo físico. El subsistema de integración permitió definir las estrategias que deben seguirse para el desarrollo de los procesos de integración de datos, en el que fueron realizadas 23 transformaciones y tres trabajos. El subsistema de visualización permitió diseñar los cubos OLAP, las vistas de análisis, los reportes operacionales y dashboard que fueron implementados como parte de la solución.
- ❖ Una vez culminado el proceso de implementación se obtuvo como resultado un MD poblado. El producto fue sometido a un conjunto de pruebas, arrojando como resultado final un total de 16 NC distribuidas en sus diferentes niveles. Estas NC fueron resueltas, garantizando así, el correcto funcionamiento del mercado y la obtención de la carta de aceptación por parte del cliente.

Referencias bibliográficas

BERNABEU, Ricardo Dario. 2010. Metodología propia para la construcción de un Data Warehouse. [En línea] 2010. [Citado el: 16 de marzo de 2017.] <http://www.dataprix.com/category/business-intelligence/business-intelligencedata-mart>.

DÍAZ, Josep Curto. 2010. *Introducción al Business Intelligence*. 2010.

ESPINOSA, Roberto. 2010. Descubriendo el Business Intelligence. [En línea] 15 de julio de 2010. [Citado el: 11 de abril de 2017.] <https://churriwifi.wordpress.com/2010/07/15/17-4-reporting-en-pentaho-con-pentaho-report-designer-otras-posibilidades-de-reporting-birt-y-jasperreports/>.

GONZÁLEZ, Hernández , Yanisbel. 2013. Serie Científica de la Universidad de las Ciencias Informáticas. [En línea] 2013. <http://publicaciones.uci.cu>.

GRAVITAR. 2016. Gravitator Información sin límites. [En línea] 2016. [Citado el: 24 de noviembre de 2016.] <http://gravitar.biz/pentaho/>.

INMON, W. H. 2005. *Building the Data Warehouse*. Fourth Edition. s.l. : Wiley Publishing, Inc., 2005. págs. 29-130.

KIMBALL, Ralph. 2006. *The Data warehouse Lifecycle Toolkit*. 2006.

PENTAHO. 2016. Pentaho. [En línea] 2016. [Citado el: 24 de noviembre de 2016.] <http://www.pentaho.com/product/data-integration#data-integration->.

POSTGRESQL. 2016. [En línea] 2016. [Citado el: 23 de noviembre de 2016.] <https://www.postgresql.org/ftp/pgadmin3/>.

SORENSEN, Kasper. 2012. DataCleaner. [En línea] 2012. [Citado el: 23 de noviembre de 2016.] <https://datacleaner.org/news>.