

Inteligencia de negocios para la empresa de servicios de la Unión del Níquel*¹

*[Business Intelligence for Company of
Services to the Union of the Nickel]*

VLADIMIR ALBERTO TORRES TORRES², EDGAR NÚÑEZ TORRES³, YANET MOLINA
HERNÁNDEZ⁴, DAYKENIS CABALLERO FERIA⁵, YANET PEÑA GONZÁLEZ⁶,
MAGBYS GONZÁLEZ ROBLEJO⁷, YADIRA ARGUELLES BLANCO⁸

Recibo: 28.04.2015 – Aprobación: 29.09.2015

Resumen

La Inteligencia de Negocios es una estrategia que ha alcanzado un nivel elevado en la competitividad empresarial. Aplicar una solución de Inteligencia de Negocios parte de los sistemas de origen de datos que posee una organización, apoyándose de un conjunto de herramientas encargadas de la extracción, depuración y consolidación de los datos. Esta información será almacenada en

* Modelo para la citación de este artículo:

NÚÑEZ TORRES, Édgar (2015). Inteligencia de negocio para la Empresa de Servicios de la Unión del Níquel. En: Ventana Informática No. 33 (jul-dic). Manizales (Colombia): Facultad de Ciencias e Ingeniería, Universidad de Manizales. p. 39-55. ISSN: 0123-9678

- 1 Reporte de caso proveniente del proyecto *Solución informática integral para geología y minería*, ejecutado en el periodo 09.2013-08.2016, e inscrito en el grupo de investigación *Datawarehouse y Tecnologías*, del Instituto Superior Minero Metalúrgico de Moa.
- 2 Ingeniero Informático. Sistematizador B, Empresa de Servicios a la Unión del Níquel Camilo Cienfuegos Gorriaran (Moa, Holguín, Cuba). Correo electrónico: vtorres@esuni.moa.minem.cu
- 3 Ingeniero Informático, Magíster en Informática para la Gestión Medioambiental. Profesor, Departamento Informática - Facultad Geología y Minas- Instituto Superior Minero Metalúrgico de Moa "Dr. Antonio Núñez Jiménez" (Moa, Holguín, Cuba). Correo electrónico: enunez@ismm.edu.cu
- 4 Ingeniero Informático. Especialista en ciencias informáticas, Almacenes Universales S.A. (Moa, Holguín, Cuba).
- 5 Ingeniero Informático. Especialista B en Gestión de Recursos Humanos, Empresa del Níquel Comandante. Ernesto Che Guevara (Moa, Holguín, Cuba). Correo electrónico: dcaballero@ecg.moa.minem.cu
- 6 Ingeniero Informático, Máster en Informática para la Gestión Medioambiental. Profesor, Instituto Superior Minero Metalúrgico de Moa (Moa, Holguín, Cuba). Correo electrónico: ypgonzalez@ismm.edu.cu
- 7 Ingeniero Informático, Máster en Informática para la Gestión Medioambiental. Profesor, Instituto Superior Minero Metalúrgico de Moa (Moa, Holguín, Cuba). Correo electrónico: mgonzalezr@ismm.edu.cu
- 8 Ingeniero en Ciencias Informáticas. Profesor, Instituto Superior Minero Metalúrgico de Moa (Moa, Holguín, Cuba). Correo electrónico: yarguelles@ismm.edu.cu

un Data Warehouse o en los Data Mart, los cuales son unidades más pequeñas orientadas a áreas específicas o un tema en particular. Esta investigación realiza el diseño e implementación de un Data Mart como solución de Inteligencia de Negocios para los servicios de alimentación prestados por la Empresa de Servicios a la Unión del Níquel (Esuni), radicada en Moa (Cuba). Se emplearon las herramientas Pentaho Bussiness Intelligence, Pentaho Data Integration 4.2.1, Pentaho Schema Workbench, PostgreSQL 9.0 y Embarcadero ERStudio 8.0.que permitieron la construcción del Data Mart y fue seleccionada la metodología Ralph Kimball para el diseño de la arquitectura y Hefesto para el desarrollo del mercado de datos, permitiendo que la información generada por los servicios gastronómicos se encuentre en un lugar específico, depurada y consolidada sirva como soporte a la toma de decisiones en la empresa.

Palabras claves: Servicios Gastronómicos, Mercado de datos, Pentaho.

Abstract

The Intelligence of Business is a strategy that has reached a high level when of managerial competitiveness it is. Applying a solution of Business Intelligence it begin with the systems data origin that it possesses a company, leaning on a tools group in charge of the extraction, purification and consolidation of the data. This information will be stored in a Data Warehouse or in Data Mart which are smaller units guided in to specific areas or a particular topic. In this investigation is carried out the design and implementation of a Data Mart like solution of Business Intelligence for gastronomic services for the Company of Services to the Union of Nickel which resides in the municipality of Moa. Several tools were used that allowed the construction of the Data Mart and Hefesto was the methodology selected for the development of the same. Allowing that all the information generated by the gastronomic services is in a specific place purified and consolidated serves like support to the taking of decisions in the gastronomic services of the Esuni.

Keywords: Food Services, Data Mart, Pentaho

Introducción

En la actualidad, las empresas dedican parte importante de su tiempo y recursos (económicos y humanos) a la información, en cuanto a su obtención, proceso y utilización, ya que ella *«juega un papel decisivo en la empresa y se convierte en su principal patrimonio. Debe ser clara, precisa y que se adapte a la formación y perfil de las personas a la que va dirigida. Tiene que ser rápida y estar disponible en el momento que se la necesite, y de una manera completa y armonizada con otras informaciones»* (Martín, citado por Arrubias, 2000). Así, *«las organizaciones hacen uso de la información para el desarrollo de sus actividades cotidianas; esta información es la parte fundamental de la empresa para que pueda tener un alto nivel de competitividad y posibilidades de desarrollo»* (Calzada & Abreu, 2009, 17).

El desarrollo de la industria minera, se caracteriza por la transformación de la minería, en un proceso de producción planificada para mejorar las condiciones de vida del pueblo. En Cuba, esta industria constituye una fuente estable de divisa, necesaria para el desarrollo económico del país. En su parte oriental, específicamente en la zona Nicaro-Moa-Baracoa, cuenta con grandes yacimientos minerales, siendo el municipio de Moa el de mayor concentración. Para mantener esta industria estable se necesitan servicios importantes donde la Empresa Servicios a la Unión del Níquel (Esuni) presta los servicios de alimentación, hospedaje y transporte entre otros, a los trabajadores de tal sector. Se cuenta con un grupo de control de los servicios gastronómicos, desarrollado por los especialistas, siendo una tarea agotadora que puede durar días cuando en ocasiones urgen resultados en periodos de tiempo extremadamente cortos.

Mediante entrevistas al personal involucrado en el proceso, la revisión de informes y reportes administrativos se determinaron algunas de las deficiencias que atentan en el proceso de gestión de la información puesto que la misma procede desde varios procesos en distintos formatos atentando contra la homogenización, consolidación y posterior procesamiento, lo que conlleva a las siguientes dificultades: Descentralización de la información, Grandes cantidades de datos almacenados en hojas de cálculo Excel, Generación de informes muy extensos con hojas adheridas, Generación de informes no homogéneos, Duplicidad de Información entre el sistema informático y los archivos Excel, Desgaste físico de los especialistas, Dificultad en la toma de decisiones y Los procesos de análisis y elaboración de informes consumen tiempo excesivo.

Lo anterior da cuenta de la necesidad urgente de una solución inteligente para la empresa. La tecnología *DataWarehouse* (DW) parece ofrecer una opción interesante y bien adaptada a estas necesidades, pues se sustenta en la integración de datos provenientes de múltiples fuentes externas que son extraídos, filtrados y almacenados en un repositorio central que según su alcance se denomina *DataWarehouse* o mercado de datos (Datamart) que soporta los procesos de toma de decisiones administrativas. A partir de las necesidades del cliente y teniendo en cuenta las características de esta tecnología en este trabajo se pretende desarrollar un mercado de datos que facilite la centralización, análisis y procesamiento de la información para soporte a toma de decisiones en el proceso de control de servicios de alimentación en la Esuni.

1. Antecedentes

1.1 Inteligencia del negocios

Este término de uso creciente, puede definirse como:

- «*La Inteligencia de Negocios, (Business Intelligence, BI, por sus siglas en inglés), tiene como objetivos la generación de información estratégica-gerencial e histórica, así como su despliegue y difusión entre los usuarios. Las aplicaciones de BI son herramientas de soporte de decisiones que permiten en tiempo real, acceso interactivo, análisis y manipulación de información crítica para la empresa*» (Thornthwaite citado por Valdés, Ramos & Gómez, 2006, 4).
- «*Un proyecto de “Inteligencia de Negocios” o “Business Intelligence” posee una serie de etapas bien definidas, empezando por la identificación de los distintos orígenes de datos, para luego pasar a la etapa de extracción, transformación y carga en el almacén de datos o Warehouse, para finalizar con la explotación de la información mediante de una herramienta OLAP (Análisis o Reportes)*» (Krauthamer, 2011).
- «*Las seis etapas que comprende un proceso completo de inteligencia de negocio: 1. Identificar los problemas de negocio y los datos que se necesitan para tratarlos. 2. Identificar la ubicación de todos los datos necesarios y extraerlos de sus fuentes. 3. Transformar los datos que provienen de varias fuentes, consolidándolos y resolviendo las heterogeneidades (estructurales y semánticas) que se presenten mediante el diseño de procesos ETL específicos. 4. Cargar los datos transformados dentro de una ubicación centralizada. 5. Construir el almacén de datos (DataWarehouse) o DataMart, tomando los datos de la ubicación centralizada. 6. Instalar en el lugar*

productos comerciales o aplicaciones personalizadas que brinden acceso a los datos» (Sáez & Naranjo, 2011, 22).

Las bases de datos son una herramienta fundamental de control y manejo de las operaciones comerciales. En grandes empresas y negocios existía considerable volumen de información almacenada en diferentes fuentes, por lo que se requirió gestionarla en un único lugar al que solo se le incorporaría aquella relevante, sobre la base de una estructura organizada, integrada, lógica, dinámica y de fácil explotación, obteniendo los Almacenes de Datos o *DataWarehouse*.

Para hacer un uso eficiente de la información histórica almacenada en un DW para la ayuda a la toma de decisiones, era vital garantizar que estos datos fueran fáciles de obtener, estandarizados y confiables. *«Un Almacén de Datos o DataWarehouse (DW) es un almacén de información temática orientado a cubrir las necesidades de aplicaciones de los sistemas de Soporte de Decisiones (DSS) y de la Información de Ejecutivos (EIS), que permite acceder a la información corporativa para la gestión, control y apoyo a la toma de decisiones»* (Fernández citado por Alba, 2009, 7).

Date (2000, 710), define un Mercado de Datos como *«un almacén de datos especializado, orientado a un tema, integrado, volátil y variante en el tiempo para apoyar un subconjunto específico de decisiones de administración»*, cuyas características son: *«Temático: Los datos están almacenados por materias o temas. Integrado: Todos los datos almacenados en el DW están integrados. No volátil: Únicamente hay dos tipos de operaciones en el DW: la carga de los datos procedentes de los entornos operacionales y la consulta de los mismos. Histórico: El tiempo debe estar presente en todos los registros contenidos en un DW»* (Fernández et al., 2010, 3-4).

1.2 Metodología de Ralph Kimball para el diseño de la arquitectura

Según Date (2000, 710), *«algunas instalaciones han seguido un enfoque de “primero el datamart”, donde los datamarts son creados conforme van siendo necesarios y el datawarehouse general es creado, finalmente, como una consolidación de los diversos datamarts»*, siendo además una copia de los datos transaccionales, estructurados de una forma especial para realizar su análisis, de acuerdo al modelo dimensional no normalizado.

En el sistema, los procesos ETL extraen la información de los sistemas operacionales y los procesan realizando posteriormente el llenado de cada uno de los *DataMart* de una forma individual, aunque siempre respetando la estandarización de las dimensiones. Este enfoque es

eficaz y conduce a una solución completa en un corto período de tiempo, como se muestra en la Figura 1. Para comprender mejor qué es un sistema de almacén de datos, es interesante considerar los procesos que intervienen en su construcción y uso. A continuación se describe cada uno de ellos, según Acosta & Jorge (2013, 5):

Extracción de los Datos: Se obtiene toda la información de las distintas fuentes tanto internas como externas. Se cargan los datos de los archivos dbf o excel, para el área temporal y así adaptarlos al modelo relacional que se ha establecido.

Transformación y limpieza: Una vez terminado el proceso de extracción, se realiza la limpieza de los datos provenientes de las diferentes fuentes, porque los mismos pueden ser incoherentes, tener errores o estar incompletos. Con esto se busca obtener datos precisos, completos, y lo más accesibles posibles. Después se efectúa la integración de los datos con el propósito de eliminar problemas de redundancia e identificar las fuentes más fiables. La transformación y limpieza garantizan el resultado final de cómo se muestran los datos, se aplican las reglas del negocio y se detectan otras posibles deficiencias de la fuente y se corrigen.

Carga: Los datos son cargados al MD, organizados y actualizados, para ser usados por el cliente de forma satisfactoria.

1.3 Metodología Hefesto para el desarrollo del mercado de datos

De acuerdo con Darío (2010, 85) «HEFESTO es una metodología propia, cuya propuesta está fundamentada en una muy amplia investigación, comparación de metodologías existentes y experiencias propias en procesos de confección de almacenes de datos. Cabe destacar que HEFESTO está en continua evolución, y se han tenido en cuenta, como gran valor agregado, todos los feedbacks que han aportado quienes han utilizado esta metodología en diversos países y con diversos fines». Es una de las más difundidas y utilizadas por su fácil implementación y aporte práctico, aunque no propone de forma explícita los artefactos y entregables a generar en cada fase.

«La metodología presenta algunas características: - Se basa en los requerimientos de los usuarios, por lo cual su estructura es capaz de adaptarse con facilidad y rapidez ante los cambios en el negocio. - Utiliza modelos conceptuales y lógicos, los cuales son sencillos de interpretar y analizar.- Es independiente del tipo de ciclo de vida que se emplee para contener la metodología.- Es independiente de las herramientas que se utilicen para su implementación.- Es independiente de las estructuras físicas que contengan el DW

y de su respectiva distribución.- Cuando se culmina con una fase, los resultados obtenidos se convierten en el punto de partida para llevar a cabo el paso siguiente.- Se aplica tanto para DataWarehouse como para DataMart» (Darío, 2010, 88)

La metodología comprende los siguientes pasos, planteados por Darío (2010, 87): - Análisis de requerimientos, que implica Identificar preguntas, Identificar indicadores y perspectivas, y Modelo conceptual; - Análisis de los OLTP, que abarca Conformar indicadores, Establecer correspondencias, Nivel de granularidad y Modelo conceptual ampliado; - Modelo lógico del DW, señalado por Tipo de modelo lógico del DW, Tablas de dimensiones, Tablas de hechos y Uniones; e – Integración de datos, que comprende Carga inicial y Actualización.

1.4 Herramientas para la construcción del mercado de datos

Uno de los grandes obstáculos a los que se enfrenta un proceso de implantación de BI es el alto costo que va asociado al mismo, por lo que cada vez más empresas optan por una solución de código abierto frente a las soluciones privativas. Para la implementación de mercado de datos se emplearon herramientas de código abierto, a partir del análisis de sus características y el tipo de licencia de software sobre la cual se distribuye.

1.4.1 Sistema gestor de bases de datos. *«Los Sistemas Gestores de Base de Datos (SGBD) fueron diseñados para gestionar grandes volúmenes de información, tanto la definición de estructuras para el almacenamiento como los mecanismos para la gestión de los datos. Estos permiten a los usuarios definir, crear, mantener la BD y proporcionar un acceso controlado en todo momento» (Daniel & López, 2007, 9).* De igual manera, señalan que el propósito principal de proporcionar a los usuarios una visión abstracta de los datos se logra mediante la definición de tres niveles de abstracción: *«El nivel físico es el nivel más bajo de abstracción, es el que describe cómo se almacenan los datos, a su vez, el nivel lógico describe que datos se almacenan realmente en la Base de Datos y que relaciones existen entre estos. El nivel más alto de abstracción es el nivel de vistas, el cual sólo presenta una determinada porción de la Base de Datos, dependiendo del tipo de usuario que la consulta, así, el sistema puede proporcionar muchas vistas para la Base de Datos» (Daniel & López, 2007, 9).*

Por su parte, Mato (2005, 4), considera que un sistema gestor de bases de datos, es *«el software que permite la utilización y/o la actualización de los datos almacenados en una o varias bases de datos por uno o varios usuarios desde diferentes puntos de vista y a la vez, se denomina sistema de gestión de bases de datos».*

«PostgreSQL es un sistema de base de datos relacional, basado en Postgres Versión 4.2, que fue desarrollado en la Universidad de California en Berkeley por el Departamento de Ciencias de la computación. PostgreSQL es un descendiente de código abierto del código original de Berkeley. Es compatible con una gran parte del estándar SQL y ofrece las siguientes características: Consultas complejas, Claves externas, Disparadores, Integridad de las transacciones.

Además, PostgreSQL puede ser ampliado por el usuario en muchos aspectos, por ejemplo mediante la adición de nuevos tipos de datos, funciones, operadores, funciones de agregación, métodos de índice y lenguajes de procedimiento» (PostgreSQL, 2013).

1.4.2 Herramientas de integración de datos. Para Naranjo & Sáez (2012), *Pentaho*⁹ proporciona un espectro completo de herramientas de inteligencia de negocio, reportes, análisis, *dashboards*, minería de datos e integración de datos. Ofrece además, una serie de servicios críticos entre los que están la autenticación, programación de tareas, seguridad y servicios Web. Este conjunto de herramientas y servicios forman una plataforma integral de inteligencia de negocio, convirtiendo a *Pentaho* en el proveedor líder de soluciones BI de código abierto.

1.4.3 Tecnología OLAP (On-Line Analytical Processing o Procesamiento Analítico en línea). Esta tecnología es definida por Peña & Suárez (2005, 7), como *«la técnica que permite ver y manipular los datos por dimensiones, que provee a los gerentes y analistas fácil acceso a la información con el fin de soportar el proceso de toma de decisión. En esta técnica de análisis, en lugar de ejecutar múltiples consultas, los datos son estructurados para permitir un acceso rápido y fácil a las respuestas de las preguntas que son típicamente formuladas».*

Ibarra (2006, 16), considera que OLAP pertenece al Sistema de Información para Ejecutivos, EIS, que proporcionan, al nivel estratégico, información útil tendiente a la toma de decisiones. En él, los datos son clasificados en diferentes dimensiones que pueden verse unas con otras en cualquier combinación y así obtener diferentes análisis de los datos contenidos. Además señala que en un modelo OLAP, *«la información*

⁹ De acuerdo con Naranjo & Sáez (2012), el servidor *Pentaho* es una colección de programas que trabajan en conjunto para proveer un número de funciones de la *Suite BI Pentaho*. Además, sus programas que no son servidores pueden ser clasificados como programas de escritorio, algunos de ellos solo actúan como clientes y necesitan interactuar con el servidor *Pentaho*, pero otros pueden ser usados de forma independiente, aprovechadas por los desarrolladores como herramientas de diseño.

es vista en cubos, los cuales consisten de categorías descriptivas (dimensiones) y valores cuantitativos (medidas). El modelo de datos multidimensional simplifica a los usuarios formular complejas, arreglar datos en un reporte, cambiar de datos resumidos a datos detallados y filtrar o rebanar en subconjuntos significativos».

Lo anterior es ratificado por Peña & Suárez (2005, 7), al asegurar que *«su objetivo fundamental es proveer al usuario final el fácil análisis de los datos, con la potencia y confiabilidad de una base de datos corporativa, y con la posibilidad de ver los datos desde diversos puntos de vista o dimensiones. Permite vistas reformateadas y calculadas sin el riesgo de perder o corromper los datos originales y hace que la información pueda ser compartida por varios usuarios sin tener que duplicar archivos».* En la tecnología OLAP se cuenta con herramientas de tres tipos: ROLAP¹⁰, MOLAP y HOLAP, siendo usada en el presente proyecto, la primera de ellas.

2. Aplicación de Hefesto

A continuación se describe paso a paso cada una de las fases de la metodología Hefesto que incluye desde el análisis de los requerimientos, pasando por la identificación de indicadores y perspectivas, dimensiones y medidas hasta llegar al modelo lógico.

2.1 Análisis de los requerimientos

Primeramente se deben identificar los requerimientos del negocio a partir de las preguntas generadas por los usuarios, para definir los indicadores o perspectivas (dimensiones) que permiten el análisis para el soporte a la toma de decisiones y el diseño del *DataWarehouse*. Los usuarios visualizan el resultado mediante un modelo conceptual. La investigación se orienta hacia un desarrollo a partir de la identificación de información de alto nivel para cumplir los objetivos y estrategias trazadas por la empresa. Es necesario formular preguntas complejas sobre el negocio en las que influyan variables de análisis, lo cual posibilitará el análisis profundo de la información en diferentes perspectivas.

Con el objetivo de obtener una visión profunda sobre las necesidades de la empresa, se revisaron informes administrativos, se observó la manera de realizar las tareas por los diferentes responsables, y no

¹⁰ «Herramientas 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, metadata, etc.» (Anchatuña & Núñez, 2007, 35-36).

menos importante fueron las entrevistas a los especialistas encargados de mantener un control absoluto de los servicios gastronómicos. A partir de dichas entrevistas se pudo conformar una serie de 10 aspectos que ayudan a conocer los requerimientos del negocio:

- Se desea conocer la cantidad de servicios de alimentación solicitados por clientes en un tiempo determinado.
- Se desea conocer la cantidad de servicios de alimentación elaborados por unidades en un tiempo determinado.
- Se desea conocer la cantidad de servicios de alimentación prestados por unidades y por clientes tiempo determinado.
- Se desea conocer la cantidad de recursos utilizados por servicios y por clientes tiempo determinado.
- Se desea conocer la cantidad de recursos utilizados por servicios, unidades y clientes tiempo determinado.
- Se desea conocer la relación de clientes que reciben tiempo determinado los mismos productos en sus servicios.
- Se desea conocer la relación de servicios prestados por unidades y clientes.
- Se desea conocer en qué tiempo determinado se prestan más servicios agrupados por tipos.
- Se desea conocer el informe de un tiempo determinado por total de servicios prestados agrupados por tipos por unidades.
- Se desea conocer el informe de un tiempo determinado por total de servicios prestados agrupados por tipo por cliente.

Luego de haber establecido las preguntas se procedió a su descomposición para descubrir los indicadores que se utilizarán y las perspectivas de análisis que intervendrán:

- Perspectivas Identificadas: Clientes, Tiempo, Unidades, Servicios, Recursos, Productos y Servicios por tipo.
- Indicadores Identificados: Cantidad de servicios solicitados, Cantidad de servicios elaborados, Cantidad de servicios prestados, Cantidad de recursos, Relación de clientes, Relación de servicios y Total de servicios prestados por tipo.

Luego, se realizó el modelo conceptual, el que engloba cada perspectiva con los campos pertenecientes a la misma, y por otro lado se visualizaran los indicadores y los cálculos que les corresponde para lograr consultarlos para observar con claridad cuáles son los alcances del proyecto, para luego trabajar sobre ellos, además al poseer un alto nivel de definición de los datos, permite que pueda ser presentado ante los usuarios y explicado con facilidad.

2.2 Análisis de los OLTP

El análisis de los OLTP (Procesamiento Transaccional en Línea), permite tener una idea de cómo identificar los indicadores. Para lograr lo mencionado anteriormente es necesario adquirir conocimiento de los elementos o hechos que proporcionará realizar cálculos y los tipos de operaciones que se debe aplicar a los hechos para obtener el valor final. Se realiza un mapeo de los sistemas transaccionales y el modelo conceptual para establecer e identificar las correspondencias existentes entre ellos. Los usuarios son los encargados de definir los campos que integrarán cada dimensión lo que representará el nivel de granularidad, una vez obtenida toda la información necesaria se realizará una ampliación al modelo conceptual.

Con este análisis, se examinan los OLTP disponibles que contengan la información requerida, como también sus características, y así poder identificar las correspondencias entre el modelo conceptual y las fuentes de datos. La idea es, que todos los elementos del modelo conceptual estén correspondidos en los OLTP. La figura 1 muestra los atributos pertenecientes al sistema operacional objeto de estudio.

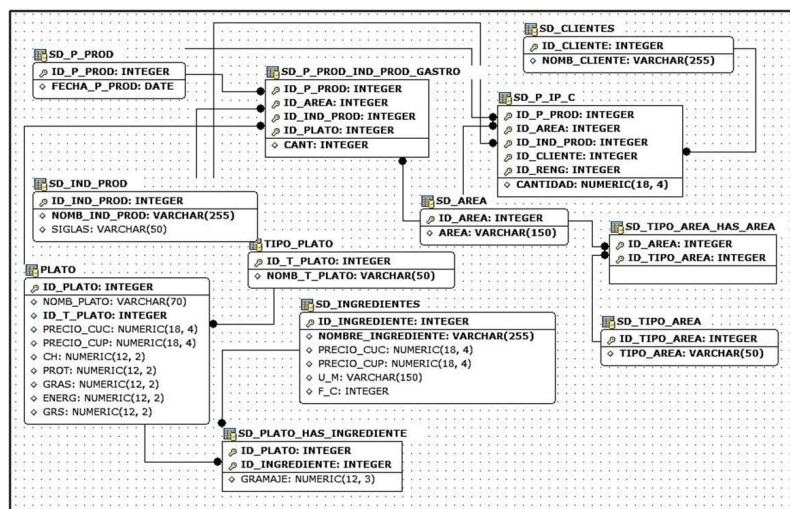


Figura 1. Representación del sistema operacional

2.3 Modelo lógico del DataMart

A continuación, se confecciona el modelo lógico de la estructura del DW, teniendo como base el modelo conceptual que ya ha sido creado. Para ello, primero se define el tipo de modelo que se utilizará y luego se llevan a cabo las acciones propias al caso, para diseñar las tablas

de dimensiones y de hechos. Finalmente, se realizan las uniones pertinentes entre las tablas.

El esquema utilizado fue estrella, ya que provee beneficios como son la simplificación del entendimiento para la navegación en los metadatos por parte del desarrollador y los usuarios finales. Esta estructura es considerada la más básica y sencilla de diseñar con respecto a las existentes, y está definida por una tabla central o tabla de hechos y un conjunto de tablas organizadas alrededor de ella, denominadas tablas de dimensiones, brindando la sensación de una estrella.

En este paso se diseñaron las tablas de dimensiones que forman parte del DW. Cada perspectiva definida en el modelo conceptual constituye una tabla de dimensión. Estas tablas de dimensión contienen los atributos de las aperturas que interesan al negocio y se pueden utilizar como criterios de filtro, siendo relativamente pequeñas. Cada tabla de dimensión se vincula con la tabla de hechos por un identificador, por lo que se tienen tablas dimensiones: *Dim_Cliente*, *Dim_Producto*, *Dim_Tipo_Producto*, *Dim_Recurso*, *Dim_Tiempo*, *Dim_Servicios*, *Dim_Unidades* y la tabla de hechos: *Fact_Gastronomia*

Una vez definido en el negocio las dimensiones, las medidas y la granularidad, se procedió a la estructuración del modelo lógico del mercado de datos. En tal sentido, se puede destacar que por las necesidades actuales del negocio solo existe un modelo que unifica las dimensiones definidas y la medida que se ha especificado hasta el momento. En la figura 2 se muestra el modelo lógico definido para este caso.

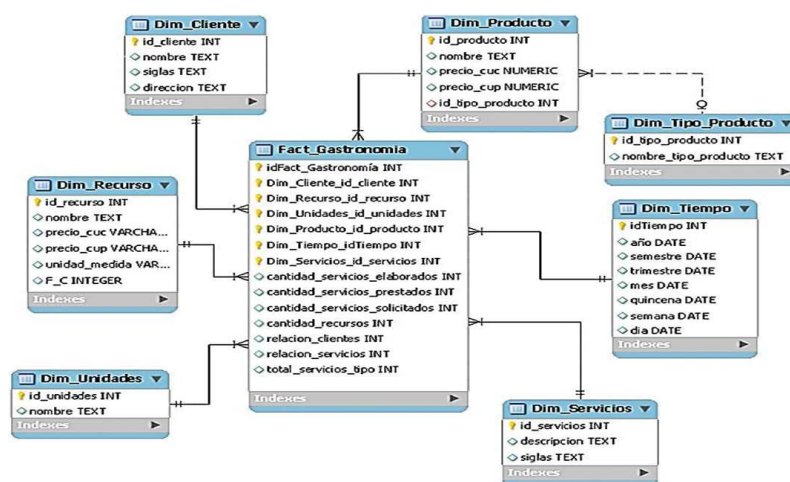


Figura 2. Modelo lógico del Mercado de datos

Luego de construido el modelo lógico, se realizó el poblado con datos, a través de procesos ETL, desde los sistemas operacionales que se desean integrar. Existen varios software que facilitan estas tareas, por lo cual se ejecutó el proceso de extracción, transformación y carga de los datos con la herramienta descrita anteriormente de la *suite Pentaho*. Cada vez que existan jerarquías de dimensiones, se comienza el cargado de las tablas de dimensiones del nivel general al detallado. Cuando se haya cargado en su totalidad el DW, se deben establecer sus políticas de actualización o refresco de datos.

2.4 Integración de datos

Para implementar el sistema primeramente debe ser poblado el mercado de datos, que implica la aplicación de las técnicas de procesos de extracción y transformación, para luego cargar los datos siendo estos lo más consistentes posible, utilizando la herramienta *Pentaho Data Integration*, que permite extraer los datos de los sistemas operacionales, transformarlos y luego cargarlos en el mercado de datos.

Una vez poblado el *DataMart* se hizo el diseño físico del mismo, teniendo en cuenta la estructura multidimensional y haciendo uso de los cubos de datos. El servidor OLAP requiere la especificación lógica de los esquemas y tablas que formarán estos cubos, así como las jerarquías y miembros que forman las dimensiones. Este diseño es elaborado con la herramienta de modelado *Pentaho Schema Workbench*, la cual permite confeccionar los cubos multidimensionales y hacer las consultas que responden a las exigencias de los clientes, a través de un archivo de esquema en XML. A partir de este esquema el servidor OLAP traduce las consultas del usuario a SQL y las ejecuta.

3. Resultados

El resultado principal consiste en la implementación de un mercado de datos en MySQL compuesto por ocho tablas que se encuentra estructurado y cargado de tal manera que pueda mediante el servidor *Pentaho* responder a las preguntas del negocio formuladas (Figura 3).

Se realizó un proceso de prueba de cada consulta y procedimientos con el fin de obtener resultados correctos. Para conocer la eficiencia de las herramientas y técnicas empleadas se efectuó una consulta, donde su funcionamiento correcto se comprobó en que los resultados arrojados fuesen los esperados. Las entradas de las consultas fueron MDX devolviendo así los datos que se consultaban. Realizada la prueba se observó que las respuestas a las consultas y procedimientos fueron los correctos.

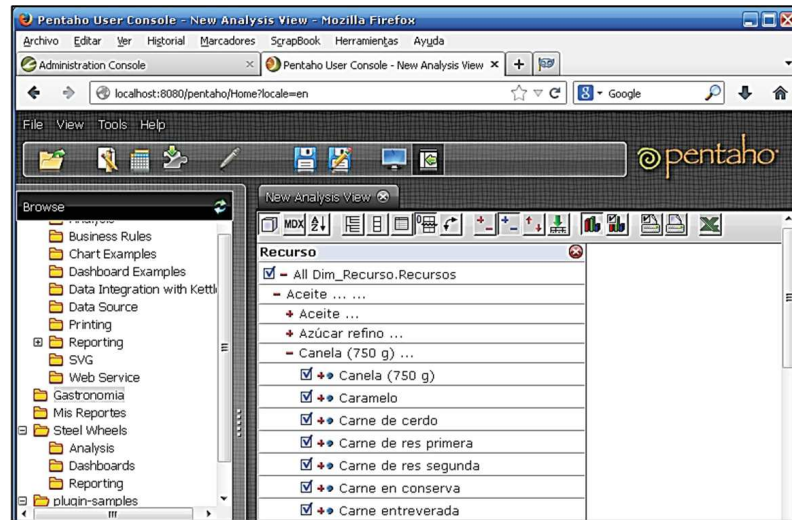


Figura 3. Mercado de datos publicado en el servidor de Pentaho

Para las pruebas de validación se hizo una consulta a Gastronomía con la finalidad de conocer si los resultados eran correctos. La medida consultada fue Cantidad de servicios utilizando el navegador OLAP el cual presenta la información de los cubos y las medidas cuando se despliegan para seleccionarlasy. Luego se ejecuta un código MDX para conocer si el sistema responde a la consulta, y una vez ejecutada la consulta, los resultados muestran que la prueba fue correcta (figura 4).

MDX Query Editor		Consulta MDX	
<pre>select {[Measures].[Relacion_Recurso]} ON COLUMNS, {[Dim_Recurso].[Precio_CUC].Members,[Dim_Recurso].[Precio_CUP].Members,[Dim_Recurso].[Nombre_Recurso].Members,[Dim_Recurso].[Unidad_Medida].Members} ON ROWS from [Gastronomia]</pre>			
Respuesta de la consulta			
Recurso	Nombre_Recurso	Precio_CUC	Precio_CUP
Aceite	Aceite	1.8500	0.0000
	Aceite	0.0000	0.0000
Acelga	Acelga	0.3500	0.0000
	Acelga	0.0000	0.0000
Agua	Agua	0.6000	0.0000
	Agua	0.0000	0.0000

Figura 4. Consulta MDX y respuesta de consulta

4. Conclusiones

La investigación presentada consistió en el desarrollo un mercado de datos que facilite la centralización, análisis y procesamiento de la información para el soporte a la toma de decisiones en el proceso de control de servicios de alimentación en la Empresa de Servicios a la Unión del Níquel, Esuni, con el fin de aportar una ayuda al proceso de toma de decisiones a partir de datos que se almacenan en el mismo de forma integrada y segura y posibilitan la creación de informes, mejorando los tiempos de análisis y trabajo con dichos datos. Para ello se diseñó el mercado de datos sobre la base de la metodología Hefesto, la cual fue muy cómoda de aplicar por la simplicidad y lo bien descrita que se encuentra cada una de sus fases lo que contribuye a un mejor entendimiento del desarrollador.

A su vez la metodología de *Kimball* permitió plantear el enfoque *Bottom-up* para el desarrollo del mercado de datos y en futuro su integración con otros en la empresa y así construir el *DataWarehouse* o repositorio central. Cabe destacar que se realizaron procesos ETL que permitieron lograr la integración, limpieza, y consolidación de los datos para poderlos cargar luego al repositorio de datos. Se validó el mercado de datos a través de pruebas realizadas a las funcionalidades, las que arrojaron resultados favorables, dando cumplimiento a las necesidades y exigencias previstas para su desarrollo hasta la etapa propuesta para esta investigación. Además, se investigó y se estudiaron las tendencias, tecnologías y herramientas a utilizar para el desarrollo de la solución propuesta, donde finalmente se optó por emplear las tecnologías libres por encontrarse libres de costo y haber obtenido muy buena popularidad por sus potentes funcionalidades para las comunidades de desarrollo.

5. Referencias bibliográficas

- ACOSTA MÉNDEZ, Geidy & JORGE CHACÓN, Disnayle (2013) SLD080 Mercado de Datos para una Dirección de Salud en Cuba [en línea]. En: Informática 2013, XV Convención y Feria Internacional - IX Congreso Internacional de Informática en la Salud 2013 (18-22/03/2013), La Habana (Cuba): Ministerio de Salud Pública. 7 p. ISBN: 978-959-7213-02-4. <<http://www.informatica2013.sld.cu/index.php/informaticasalud/2013/paper/view/195/184>> [consulta: 18/10/2013]
- ALBA MÉNDEZ, Ana Laura (2009). Arquitectura, Diseño, Construcción, Mantenimiento y Consulta de un Almacén de Datos [en línea]. Monografía (Licenciado en Sistemas Computacionales Administrativos). Xalapa-Enríquez (Veracruz, México): Universidad Veracruzana, Facultad de Contaduría y Administración. 105 p. <<http://cdigital.uv.mx/bitstream/123456789/29385/1/AnaLauraAlbaMendez.pdf>> [consulta: 16/09/2013]

- ANCHATUÑA GUANOLISA, Luz América & NÚÑEZ PADILLA, Maritza Germania (2007). Creación de un web site e implantación y desarrollo de datawarehousing aplicada a la gestión del Hospital Docente de Riobamba [en línea]. Proyecto de grado (Ingeniero de Sistemas e Informática). Latacunga (Ecuador): Escuela Politécnica del Ejército, Facultad de Ingeniería de Sistemas e Informática. 140 p + anexos. <<http://repositorio.espe.edu.ec/bitstream/21000/35011/1/T-ESPEL-0036.pdf>> [consulta: 18/10/2013]
- ARRUBIAS URRUTIA, Amaia (2000). Comunicación en la empresa: La Importancia de la Comunicación Interna en la Empresa [en línea]. En: Revista Latina de Comunicación Social, Vol. 3, No. 27 (may). La Laguna (Tenerife, Canarias; España): Universidad de La Laguna, Laboratorio de Tecnologías de la Información y Nuevos Análisis de Comunicación Social. ISSN: 1138-5820 <<http://www.ull.es/publicaciones/latina/aa2000tma/127amaia.html>> [consulta: 18/10/2013]
- CALZADA, Leticia & ABREU, José Luis (2009). El impacto de las herramientas de inteligencia de negocios en la toma de decisiones de los ejecutivos [en línea]. En: Daena: International Journal of Good Conscience, Vol. 4, No 2 (sep). Monterrey (NL, México): Spenta University Mexico. p. 16-52. ISSN: 1870-557X. <[http://www.spentamexico.org/v4-n2/4\(2\)%2016-52.pdf](http://www.spentamexico.org/v4-n2/4(2)%2016-52.pdf)> [consulta: 18/10/2013]
- DANIEL DÍAZ, Marisol & LÓPEZ GUZMÁN, Verónica (2007). Soluciones de software libre para el desarrollo de aplicaciones de bases de datos [en línea]. Tesis (Licenciado en Sistemas Computacionales). Pachuca (Hidalgo, México): Universidad Autónoma del Estado de Hidalgo, Instituto de Ciencias Básicas e Ingeniería. 82 p. <<http://www.uaeh.edu.mx/docencia/Tesis/icbi/licenciatura/documentos/Soluciones%20de%20software%20libre%20para%20el%20desarrollo%20de%20aplicaciones.pdf>> [consulta: 16/09/2013]
- DARÍO, Bernabeu Ricardo (2010). HEFESTO. Data Warehousing: Investigación y Sistematización de conceptos. HEFESTO: Metodología para la Construcción de una Data Warehouse [en línea]. Córdoba (Argentina): Licencia de Documentación Libre de GNU. 148 p. <<http://www.businessintelligence.info/docs/hefesto-v2.pdf>> [consulta: 16/09/2013]
- DATE, Christopher J. (2000). Introducción a los sistemas de bases de datos, 7 ed. México (México): Addison-Wesley. 936 p. ISBN: 9789684444195.
- FERNÁNDEZ ALMODÓVAR, Eduardo; GÓMEZ ORTIZ, Miguel; LÓPEZ LÓPEZ, Quiteria; MORILLO BALCÁZAR Alberto; NOVALVOS NOVALVOS, Nicolás & ROSADO QUINTANILLA, Javier (2010). Funcionalidad 4: Almacenes de Datos y Bases de Datos XML [en línea]. En: MORAGA, Ma. Ángeles (2010). Asignatura: Modelos Avanzados de Bases de Datos, curso 2009-2010. Cuenca (España): Universidad de Castilla-La Mancha, Escuela Superior de Informática. 18 p. <<http://alarcos.inf-cr.uclm.es/doc/bbddavanzadas/07-08/Funcionalidad4.pdf>> [consulta: 16/09/2013]
- IBARRA, María de los Ángeles (2006). Procesamiento Analítico en Línea (OLAP) [en línea]. Trabajo de Adscripción (Licenciada en Sistemas de Información). Corrientes (Argentina): Universidad Nacional del Nordeste, Facultad de Ciencias Exactas, Naturales y Agrimensura, Materia de Diseño y Administración de Datos. 46 p. <<http://exa.unne.edu.ar/depar/areas/informatica/SistemasOperativos/OLAPMonog.pdf>> [consulta: 17/10/2013]
- KRAUTHAMER, Diego (2011). Ventajas del uso de herramientas de ETL sobre ANSI SQL [en línea]. En: VI Jornadas de Docentes Universitarios de Sistemas y Tecnologías de Información de Facultades de Ciencias Económicas, DUTI 2011 (18-20/08/2011). Rosario (Argentina): Asociación de Docentes Universitarios de Sistemas y Tecnologías de Información de Facultades de Ciencias Económicas, DUTI. <<https://duti2011.files.wordpress.com/2011/10/krauthamerdiego-uai-ventajas-del-uso-herramientas-etl-sobre-ansi-sql1.pdf>> [consulta: 28/06/2015]
- MATO GARCÍA, Rosa María (2005). Sistemas de bases de datos. La Habana (Cuba): Editorial Pueblo y Educación. 165 p. ISBN: 978-959-13-1273-0
- NARANJO HERNÁNDEZ, Esther & SÁEZ MOSQUERA, Inty (2012). Pentaho: software líder de Inteligencia de Negocio de código abierto [en línea]. En: Telem@tica: Revista digital de las tecnologías de la información y las comunicaciones, Vol. 10, No. 2 (may-ago). La Habana (Cuba): Departamento de Telemática, Instituto Superior Politécnico. ISSN: 1729-3804 <<http://revistatelematica.cujae.edu.cu/index.php/tele/article/view/44>> [consulta: 10/10/2013]

- PEÑA RIVIERA, Juan David & SUÁREZ DAZA, Jesús Armando (2005). Utilización de información histórica para decisiones empresariales. Trabajo de grado (Ingeniero de Sistemas). Santa Fé de Bogotá D.C (Colombia): Pontificia Universidad Javeriana. 128 p.
- POSTGRESQL (2013). What is PostgreSQL? [online]. Toronto (ON, CA): The PostgreSQL Global Development Group. <<http://www.postgresql.org/docs/8.4/static/intro-whatsis.html>> [consult: 10/10/2013]
- SÁEZ MOSQUERA, Inty & NARANJO HERNÁNDEZ, Esther Lidia (2011). Adición de capacidades de inteligencia de negocio: un acercamiento [en línea]. En: GiGa: Revista cubana de computación, No. 3. La Habana (Cuba): Copextel. p. 22-26. ISSN: 1028-270X. <<http://www.copextelvc.co.cu/index.php/site-map/descargas/finish/39/16>> <<http://www.copextel.com.cu/sites/default/files/revistas/GiGA-3-2011.pdf>> [consulta: 18/10/2013]
- VALDÉS RODRÍGUEZ, Yisel; RAMOS MAESTRE, Mayrin & GÓMEZ RODRÍGUEZ, Jandy Miguel (2013). Mercado de datos para la dirección de cuadros de la Administración Provincial de Artemisa [en línea]. En: Revista Cubana de Ciencias Informáticas, Vol. 7, No. 1 (ene-mar). La Habana (Cuba): Universidad de las Ciencias Informáticas, Grupo Editorial 'Ediciones Futuro'. p. 1-8. e-ISSN: 2227-1899 <<http://rcci.uci.cu/index.php/rcci/article/view/291/201>> <http://scielo.sld.cu/scielo.php?pid=S2227-18992013000100001&script=sci_arttext> [consulta: 17/10/2013]