

MERCADO DE DATOS COOPERACIÓN INTERNACIONAL PARA LA SALA SITUACIONAL DE LA UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS

DATAMART INTERNATIONAL COOPERATION IN THE SITUATION ROOM OF THE UNIVERSITY OF INFORMATIC SCIENCES

Yisel de los Angeles González Pompa¹, María Teresa Rosales González²

1 Universidad de las Ciencias Informáticas Facultad Regional Granma, Cuba, yisel.pompa@grm.uci.cu, Calle 3ra %12 y 14 # 253 Reparto Francisco Vicente Aguilera

2 Universidad de las Ciencias Informáticas, Cuba, mtrosales@uci.cu

RESUMEN: *La Universidad de las Ciencias Informáticas es un centro de estudios universitarios que cuenta con varias direcciones para garantizar su correcto funcionamiento. Una de estas direcciones es Cooperación internacional, cuya información es almacenada en diversos formatos y en períodos variables, convirtiéndose en una tarea compleja su unificación, manejo e interacción. Este conjunto de factores trae como consecuencia, que los directivos de la universidad no cuenten con las herramientas requeridas para ofrecer análisis de los datos que respalden las decisiones tomadas. La presente investigación se enmarca en el tema de los almacenes de datos, los mercados de datos, y su utilización para los análisis estadísticos de la información. El Centro de Tecnologías de Gestión de Datos (DATEC) desarrolló un almacén de datos que contribuye al proceso de toma de decisiones y que está compuesto por mercados de datos que abarcan la totalidad de las áreas de la universidad. La metodología para guiar el desarrollo de la solución es la Propuesta de metodología para el desarrollo de almacenes de datos en DATEC, obteniéndose como resultado del proceso de desarrollo un mercado de datos poblado y funcional, con una capa de inteligencia de negocio que brinda vistas de análisis actualizadas, facilitándole a los especialistas de la Universidad de las Ciencias Informáticas un mejor estudio de la información.*

Palabras Clave: Cooperación internacional, inteligencia de negocio, Mercado de datos.

ABSTRACT: *The University of the Informatic Sciences is a university center that features several directions to ensure proper operation. One of these addresses is international cooperation, whose information is stored in different formats and in varying periods, becoming a complex task the unification, management and interaction. This combination of factors results in that the management of the university don't have the tools required to provide data analysis to support decisions. This research is part of the topic of data warehouses, data marts, and their use for statistical analysis of the information. The Center Data Management Technologies developed a data warehouse that contributes to the process of decision making and that is comprised of market data covering all areas of the University. The methodology to guide the development of the solution is the proposed methodology for developing data warehouses by DATEC, obtained as a result of the process of developing a market populated data and functional with a business intelligence layer that provides views of updated analysis, providing to specialists from the University of Informatic Sciences a better study of information.*

KeyWords: International cooperation, business intelligence, Datamart.

1 INTRODUCCIÓN

El avance alcanzado por las Tecnologías de la Informática y las Comunicaciones ha influido en el desarrollo del ser humano. Rompiendo paradigmas, jugando un papel significativo para las economías de los países destacados en este campo. Continuamente el desarrollo de software amplía su alcance y se gestiona con mayor profundidad. Siendo

hoy en día un proceso de simple y rápida ejecución para los usuarios de sistemas informáticos procesar operaciones muy complejas, donde se manejen enormes volúmenes de información. Relacionado a estos sistemas y dentro de su entorno surgen las bases de datos (BD).

Paralelo a ello, la necesidad de extraer información útil a partir de datos históricos se ve reflejada de

forma más tangible y valiosa en el mundo de los negocios. Con el objetivo de brindarle soporte a este proceso surgen los sistemas de apoyo a la toma de decisiones, creados para medir y controlar el desarrollo de las variables importantes del negocio, buscando identificar, proyectar y predecir tendencias futuras a partir de datos acumulados. La Universidad de las Ciencias Informáticas (UCI), cuenta con el centro de desarrollo DATEC que se especializa en el desarrollo de soluciones, productos y servicios relacionados con las tecnologías de gestión de datos.

DATEC se encuentra colaborando con el proceso de informatización que se está llevando a cabo en la universidad, entre sus proyectos se encuentra el denominado Sala Situacional UCI que abarca varias áreas de la universidad, una de ellas es Cooperación internacional, dentro de esta la información detallada es almacenada en diversos formatos y en períodos variables, convirtiéndose en una tarea compleja su unificación, manejo e interacción. Por lo cual la realización de los reportes del área es de forma manual y con un gran costo de tiempo, lo que trae consigo una demora en el análisis y estudio de la totalidad de la información. Este conjunto de factores trae como consecuencia que los directivos de la UCI no cuenten con las herramientas requeridas para ofrecer datos que respalden las decisiones tomadas, lo que puede incidir de manera negativa en la ejecución de sus procesos regulares y en la prevención de comportamientos futuros que pueden afectar el adecuado funcionamiento de la misma.

Para darle solución al problema antes descrito se define como objetivo general: Desarrollar el mercado de datos Cooperación internacional para la Sala Situacional de la Universidad de las Ciencias Informáticas, que contribuya a la toma de decisiones.

2 CONTENIDO

2.1 Almacenes de datos

Los almacenes de datos (AD) constituyen un escalón superior en la evolución de las BD hacia mayor inteligencia y mejores funcionalidades. Los fundamentos o características de los AD según W. H. Inmon, padre de la disciplina, se basan en que estos sistemas consisten en una colección de datos orientado a materias, integrado, no volátil, que varía con el tiempo y que se diseñan para dar apoyo a los procesos de toma de decisiones [1].

Se puede decir que un AD es orientado a materias porque sus funcionalidades se concentran en un tema en específico o en base a los aspectos que son de interés para la empresa. Además es integrado debido a que se construye mediante la unificación de fuentes de datos múltiples y heterogéneas, eliminando las inconsistencias en el nombrado, estructuras codificadas y medidas de los atributos que puedan surgir entre ellas. Una vez que se en-

cuentra poblado el almacén, los datos no cambian, sus actualizaciones no ocurren en su propio entorno, sino en las fuentes de datos de las que se nutre, siendo esta una de las razones por las que un AD es no volátil. La variación en el tiempo se ve evidenciada en un AD primeramente porque cada clave contiene una referencia a la fecha explícita o implícitamente. Igualmente el espacio de tiempo en él, es significativamente más largo que el de los sistemas operacionales debido a que la información es mostrada desde una perspectiva histórica [1; 2].

Ventajas

A continuación se refieren algunas ventajas que trae la implementación de un AD [1]:

- Proporciona una herramienta para el proceso de toma de decisiones estratégicas y tácticas en cualquier área funcional, teniendo en cuenta información integrada y global del negocio.
- Simplifica la aplicación de técnicas estadísticas de análisis y modelación para encontrar relaciones ocultas entre los datos del almacén; obteniendo metadatos como valor añadido para el negocio sobre dicha información.
- Aporta la capacidad de aprender de los datos del pasado y de predecir situaciones futuras en diversos escenarios.

Desventajas

Algunas de las desventajas de los AD son las que se presentan posteriormente:

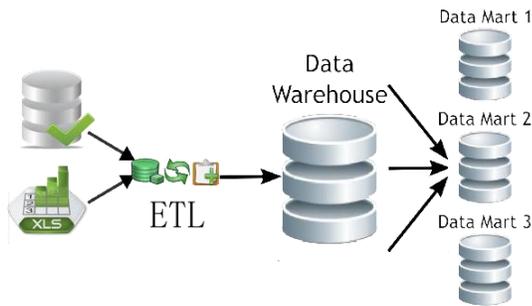
- Subestimación de los recursos imprescindibles para la captura, carga y almacenamiento de los datos [3].
- Incremento continuo de los requerimientos de información del AD [3].
- Subestimación de las funcionalidades que ofrece la correcta utilización del AD [1].

2.1.1 Mercado de datos

Un MD es la implementación de un almacén con alcance limitado a un área funcional, problema en específico, departamento, tema o grupo de necesidades [1]. Este se puede integrar a un AD de dos formas, dependiendo del enfoque de la arquitectura que se decida implementar:

Enfoque de Inmon

Define previamente el AD y posteriormente delimita sobre él los MD. Esta implementación tiene como principal inconveniente que se requiere una inversión muy costosa en tiempo y recursos [1].



**Figura 1 Arquitectura descendente
Enfoque de Kimball**

Plantea crear previamente los MD departamentales y posteriormente integrarlos en un AD para la organización, la fundamental ventaja de esta implementación es que se desarrolla y despliega en un período de tiempo menor que una realizada con el enfoque de Inmon y con menos recursos. El principal problema que se presenta al aplicar esta solución está dado en tener que sincronizar los hechos cuando se realice la consolidación del AD [1].

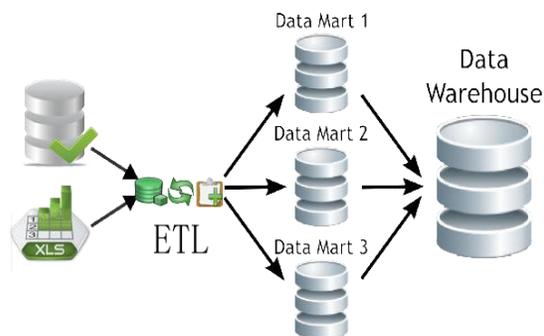


Figura 2 Arquitectura ascendente

Modelo multidimensional

En un AD se representan los datos usando modelos multidimensionales para facilitar su análisis. Un modelo multidimensional proporciona dos conceptos fundamentales: hecho y dimensión. Constituyendo el cubo el objeto central en el modelado de datos multidimensional.

Un cubo consiste en un conjunto de celdas, de tal manera que cada una está identificada por la combinación de los miembros de las diferentes dimensiones y contiene el valor de la medida analizada para dicha combinación de dimensiones [4].

Contenida en estas tablas, en muchos casos aparece la relación lógica entre dos o más atributos, es decir las jerarquías de datos, conjuntamente, de acuerdo a las dimensiones del negocio, se refleja la granularidad, que representa el nivel de detalle en que se desea almacenar la información en el modelo [4].

2.1.2 Procesos básicos de un almacén de datos

Para el desarrollo exitoso de un AD se requiere seguir una serie de procesos en cada una de las etapas del mismo. Estas etapas van desde el análisis de la organización hasta el despliegue del AD y

básicamente estos procesos se resumen en dos: la integración de los datos y el BI.

Integración de datos

Aunque la creación del ETL constituye una actividad que pasa desapercibida en muchas ocasiones al usuario final, fácilmente consume el 70 por ciento de los recursos requeridos para la implementación y el mantenimiento de un típico AD [5]. En el proceso ETL se añaden valores significativos a los datos. Este es mucho más que profundizar para obtener datos de los sistemas fuentes y del AD. Fundamentalmente el ETL permite realizar funcionalidades como:

- Eliminar errores y corregir datos faltantes.
- Capturar el flujo de datos transaccionales para su protección.
- Concordar los datos de múltiples fuentes para ser usadas conjuntamente.
- Estructurar los datos para ser usables por las herramientas del usuario final [6].

Inteligencia de negocio

En la actualidad la información se ha convertido en fundamento para conseguir superioridad competitiva en los negocios. Las funcionalidades de las herramientas de BI están encaminadas a generar reportes, gráficos en diferentes perspectivas, vistas de análisis y todo lo que se necesite para que el usuario final pueda analizar y razonar los datos con mayor facilidad, con el objetivo de arribar a conclusiones correctas con mayor prontitud.

El BI constituye la habilidad corporativa para tomar decisiones. Esto se logra mediante el uso de metodologías, herramientas y tecnologías que permiten reunir, depurar, transformar datos, y aplicar en ellos técnicas analíticas de extracción de información [7]. Como principales ventajas de la utilización de BI se tienen [7]:

- Mayor rapidez en la obtención y lectura de la información, y en consecuencia para la toma de decisiones.
- Visibilidad inmediata y automática de los nuevos datos incorporados.
- Se genera un flujo constante de información válida, organizada y visualizada de tal forma que contribuye a minimizar la incertidumbre.

2.2 Metodología para el desarrollo de almacenes de datos

Una metodología es un conjunto de pasos o acciones definidas para guiar un proceso determinado. En el caso del desarrollo del software, su definición no se aleja de este marco, a esto se le añade la definición de las etapas, los roles y las tareas para

la generación de los artefactos finales de productos informáticos [8].

Con el transcurso de la madurez de los AD, se han convertido en referente en cuanto a sistemas para el apoyo en la toma de decisiones partiendo de información histórica. Las más reconocidas en el mundo son la metodología de Kimball y la metodología de Inmon ambas abundantemente documentadas y definidas.

Para las soluciones de esta rama DATEC propone una la Metodología para el desarrollo de almacenes de datos en DATEC que abarca todas las fases del desarrollo de un AD, esta toma como fundamento la metodología de Kimball en los siguientes teniendo en cuenta que esta propone los conceptos de hechos y dimensiones, lo cual apoya la toma de decisiones y también en el proceso de desarrollo. Plantea que la construcción del AD sea a medida que se construyan las BD departamentales, lo que da la impresión de organización en la empresa, entidad u organización. Abundante documentación sobre la misma, posibilitando así la aclaración de las dudas que vayan surgiendo relacionadas a ese tema. Amplio reconocimiento por los especialistas que se desempeñan en esta disciplina, pues constituye una metodología madura y que tiene bien definidas las etapas, actividades, artefactos y roles [6]. Además esta metodología se ajusta a las condiciones y características de la producción en DATEC y la UCI y está constituida por las fases estudio preliminar y planeación, levantamiento de requisitos, arquitectura, diseño e implantación y prueba, esta última incorporada a partir de las peculiaridades de la universidad.

RESULTADOS Y DISCUSIÓN

En la presente sección se muestran los resultados de la aplicación de la metodología y los elementos que se tuvieron en cuenta en cada una de las fases de desarrollo. Así como los resultados obtenidos al probar el mercado de datos una vez poblado y funcional. Además se evidencian las experiencias adquiridas y las ventajas que puede ofrecer la utilización de la aplicación.

2.3 Análisis, diseño e implementación del MD Cooperación internacional.

En el proceso de análisis se llevó a cabo la traducción de las necesidades de los usuarios, es decir lo que precisan que el sistema les facilite, en requisitos. De la calidad con la cual se realiza la extracción de estas necesidades depende en gran medida el éxito o fracaso del software.

El análisis comprende una serie de tareas entre las que se encuentra la definición de las reglas de negocio que describen el nivel adecuado de detalle hasta el que se va a llegar en la implementación. Para el caso del MD Cooperación internacional fueron identificadas las siguientes RN.

1. Una vez que los datos estén cargados en el almacén, no pueden existir campos nulos.

2. El código de los atributos en cada una de las dimensiones no puede tomar valores repetidos.

3. No pueden existir campos con valores negativos.

Otra de las actividades es la definición de los requisitos de información, estos consisten en el conjunto de funcionalidades que se deben implementar para analizar los datos en cuanto a parámetros previamente definidos por el cliente.

A continuación se declaran los RI identificados para esta investigación:

4. Obtener el total de misiones realizadas por país, tipo de misión y/o período.

5. Obtener el total de prórrogas por estado de prórroga, nombre de misión, tipo de misión y/o período.

6. Obtener el total de pasaportes UCI por tipo de pasaporte, período y/o estado del pasaporte.

7. Obtener el total de trámites realizados por nombre de misión, tipo de misión, país y/o período.

8. Obtener el total de becas intergubernamentales por país y/o período.

9. Obtener el total de misioneros en el exterior por país, nombre de misión, tipo de misión y/o período.

10. Obtener el total de misioneros con misión cumplida por país, tipo de misión, nombre de misión y/o período.

11. Obtener el total de extranjeros con residencia temporal en Cuba por país y/o período.

12. Obtener el total de acuerdos de colaboración por país, entidad, directivo de la UCI firmante y/o período de firma.

13. Obtener el total de profesores invitados por país, actividad y/o período.

14. Obtener el total de visitas por país, clasificación y/o período.

15. Obtener el total de personal que ha visitado la UCI por país y/o período.

16. Obtener el total de regalos entregados por tipo de artículo y/o período.

17. Obtener el presupuesto del área por epígrafe, tipo de moneda y/o período.

18. Obtener el total de gastos autorizados por el CAD por período.

19. Obtener cantidad de materiales gastados por tipo de epígrafe y/o período.

Por último se definen los requisitos funcionales (RF) y los requisitos no funcionales (RNF), según la Ayuda del Proceso Unificado de Software (RUP) un RF es una capacidad de software necesaria para que el usuario solucione un problema al alcanzar un objetivo [9]. A continuación se exponen los RF identificados para esta investigación:

1. Validar datos y contraseña

2. Adicionar rol

3. Eliminar rol

4. Visualizar rol

5. Modificar rol

6. Adicionar usuario

7. Eliminar usuario
8. Visualizar usuario
9. Modificar usuario
10. Adicionar reporte
11. Visualizar reporte
12. Eliminar reporte
13. Modificar reporte
14. Extraer datos de la fuente
15. Realizar transformación de los datos
16. Cargar los datos al MD Cooperación internacional
17. Personalizar reporte
18. Exportar a otro formato

Estos son el punto de partida para elaborar el diagrama de caso de uso del sistema este a su vez agrupa un conjunto de funcionalidades por un patrón determinado que benefician a uno o varios actores.

Los actores del sistema son personas que interactúan con la aplicación y se benefician de ella [9]. Para el MD Cooperación internacional los actores son Administrador, que realiza todas las actividades de administración del MD, este se puede comportar como el actor analista debido a la relación de generalización/especialización que ellos tienen, evidenciándose de esta manera el patrón múltiples actores rol común. El analista realiza las actividades de extracción de los datos de las fuentes, así como la transformación y la carga de estos hacia el MD y el administrador de ETL consulta la información de los reportes.

La interacción entre los CU y los actores del sistema se modela en un DCUS. Para la elaboración de este se tomó como base los RI y RF anteriormente identificados agrupados en 11 CU, de ellos tres son agrupados por el patrón CRUD, que plantea la agrupación de las funcionalidades crear, leer, modificar y eliminar. Este se ve reflejado en los CU Gestionar reporte, Gestionar usuario y Gestionar rol. Otro patrón de CU reflejado en este diagrama es concordancia por adición, el cual se pone de manifiesto en la relación de extensión que existe entre los CU de información y el CU Realizar operaciones sobre reportes.

2.3.1 Arquitectura del mercado de datos

La arquitectura de la solución se define a partir de los RNF obtenidos en el levantamiento de requisitos. Esta comprende aspectos como, la comunicación entre los subsistemas y la tecnología a utilizar, entre otros aspectos de gran importancia [10]. Para el caso del MD Cooperación internacional la arquitectura definida consta de tres subsistemas: integración de datos, almacenamiento y visualización.

Dimensiones, hechos y medidas

Los términos hecho, dimensión y medida son muy comunes en la disciplina de los AD, debido a que estos constituyen uno de los principales aspectos que la caracterizan y diferencian sobre las BD. Un hecho representa la ocurrencia de un proceso es-

pecífico en el interior de la organización; por su parte una dimensión constituye una perspectiva mediante la cual se puede llevar a cabo el análisis sobre el hecho y una medida constituye un valor o indicador de análisis del hecho [2].

Partiendo de lo identificado en el análisis y en conjunto con el estudio de las fuentes de datos se elaboró el modelo de datos en el cual se identificaron 20 tablas de dimensiones, 11 tablas de hecho y 14 medidas.

2.3.2 Diseño del subsistema de integración de datos

Para garantizar que la implementación del subsistema de integración de datos se realice de forma exitosa primero se debe concebir un diseño que lo soporte, teniendo en cuenta la mayor cantidad de aspectos y con el mayor nivel de detalle posible, este comprende aspectos como la definición de la arquitectura de integración para definir como se debe realizar el proceso de ETL, donde se incluyen aspectos como los tipos de archivos fuente y su interacción con las herramientas dentro del entorno de trabajo y por último la gestión de metadatos para controlar la ejecución de los procesos de integración de datos.

Otro aspecto importante a tener en cuenta en este proceso el perfilado de datos, este constituye una de las primeras tareas en el proceso de calidad de datos, consiste en ejecutar un primer análisis sobre los datos fuentes, usualmente sobre las tablas, con el objetivo de comenzar a conocer su estructura, formato y nivel de calidad. Al realizar el perfilado de los datos a las fuentes del MD Cooperación internacional se identificó que los tipos de datos de las fuentes son en su mayoría varchar, float e integer, pero además existe la presencia de fechas con formato timestamp. Además se encontraron 612 valores nulos, lo que permitió crear nuevas RN que serán solucionadas en la implementación de las transformaciones. Igualmente no se hallaron valores vacíos ni negativos.

2.3.3 Diseño del subsistema de visualización

El subsistema de visualización constituye un elemento fundamental para el MD, debido a que el usuario final interactúa con él. Por ello realizar un correcto y detallado diseño del mismo brinda muchas facilidades en la implementación. Es por ello que definir una arquitectura de información que responda a las particularidades del área brinda la posibilidad de organizar los contenidos para que el acceso a ellos se pueda realizar de forma más simple y directa.

2.3.4 Esquema de seguridad

El desarrollo de un completo esquema de seguridad

para una aplicación informática, garantiza en gran medida que los datos no se dañen o alteren por circunstancias o factores externos, brindando de esta forma protección, resguardo y recuperación a ellos. La seguridad para el MD Cooperación internacional se maneja en gran medida por los niveles de acceso al sistema, teniendo como apoyo los roles creados para la interacción de los usuarios con el sistema.

Seguridad en la base de datos

Para acceder a la BD se definieron 3 roles, el Administrador ETL, cuenta con permiso sólo para la realización de las funciones del proceso de ETL, el Administrador posee acceso total a la BD y ejecuta las funciones de administración en la misma, y por último el Analista que cuenta con acceso de sólo lectura a la BD y accede a la misma solo mediante la aplicación.

Seguridad en el proceso de ETL

En el proceso de ETL se garantiza que los datos de las fuentes, las transformaciones y los trabajos no sean manipulados mediante la seguridad que ofrece el sistema operativo, al marcar el atributo de sólo lectura en las propiedades de la carpeta donde se almacenan todos ellos. De esta forma sólo puede modificar estos ficheros el usuario que tenga permisos de escritura para ello.

Seguridad en la aplicación

Con el objetivo de garantizar que accedan al sistema sólo los usuarios definidos por el cliente, es importante la creación de un esquema de seguridad que permita la sostenibilidad en la gestión de la misma en este servidor. A continuación se muestran los roles delimitados con el objetivo de que el acceso a la aplicación se haga de forma organizada y con los permisos correspondientes. El administrador tiene acceso total al sistema Sala Situacional UCI, realiza las operaciones de administración sobre el mismo, por su parte el analista tiene acceso de sólo lectura al MD Cooperación internacional además de analizar la información contenida en los reportes.

2.3.5 Implementación del mercado de datos Cooperación internacional

Como parte de la implementación del subsistema de almacenamiento se desarrolló el modelo físico, este constituye una colección integrada de entidades que describe las estructuras de los datos. Dicho modelo se genera a partir del modelo lógico dimensional; conteniendo las relaciones entre las tablas de hechos y dimensiones que conforman el MD.

En el subsistema de integración de datos incluye aspectos como la gestión del cambio en las dimensiones para los datos que tienden a ser modificados en el transcurso del tiempo. Fueron utilizados los metadatos de proceso con el objetivo de obtener la información de las transformaciones y los trabajos

pertinentes a los subprocesos de ETL. Se realizaron 11 transformaciones para la carga de los hechos, en las cuales se utilizaron un conjunto de subsistemas que Kimball propone para la implementación del proceso de ETL y los trabajos.

Por último se implementó el subsistema de visualización, proceso muy importante en el desarrollo de un MD, debido a que define qué información se presenta para los usuarios y cómo se muestra la misma para el análisis. Como parte de la implementación del mismo se desarrollaron un total 11 cubos OLAP, uno por cada hecho. Estos permiten estructurar los datos para que concuerden con el modo que tienen los usuarios de analizarlos. Luego se realizaron los reportes definiéndose una conexión única para el AD nombrada ssuci_dw en la cual se establecieron los parámetros necesarios para acceder a la BD.

2.4 Pruebas realizadas al mercado de datos Cooperación internacional

Para validar el correcto funcionamiento de un software es importante ir probándolo desde el inicio de su construcción hasta después de aceptado por el cliente. Las actividades de control de la calidad del mismo deben tener un orden de ejecución lógico y funcional, por lo cual el centro DATEC define el uso del modelo V con estos fines.

Este modelo tiene mucha similitud con el Modelo en cascada pues constituye una evolución del mismo. Lo que diferencia uno del otro es que el modelo V toma las actividades de prueba como un proceso que corre en paralelo con las actividades de análisis y diseño, en lugar de establecer una fase independiente al concluir el proyecto como define el Modelo en cascada [10].

Como parte de la estrategia de prueba para aplicar los tipos de pruebas mencionados anteriormente al MD Cooperación internacional, se emplearon como herramientas los Casos de prueba (CP) para CU y reglas de transformación, las listas de chequeo a los artefactos de ETL y el perfilado de los datos destino para probar la calidad de los mismos.

Los CP por CU se diseñan en correspondencia con las funcionalidades detalladas en los CU de información. Este diseño se confecciona antes de realizar las pruebas funcionales de la aplicación. Para el MD Cooperación internacional fueron diseñados cuatro CP. En el proceso de ETL también se aplicaron CP, en este caso para las reglas de transformación con la finalidad de comprobar si luego de ejecutada cada transformación, los datos cargados sean los esperados, encontrándose un grupo de inconformidades resueltas satisfactoriamente.

Las listas de chequeo son un mecanismo para controlar los riesgos, su función básica consiste en detectar condiciones peligrosas que puedan acarrear problemas al software. Constituyen un documento con un conjunto de parámetros a medir de un aspecto específico, sea documentación o aplicación.

Estos parámetros son básicamente una serie de preguntas referentes al atributo de calidad que se está probando y a las individualidades del documento en el caso de la documentación. A cada pregunta se le asocia una evaluación en una escala que da como resultado una medida del grado de cumplimiento y disponibilidad de lo evaluado, de esta manera se determina la evaluación del elemento probado.

2.4.1 Pruebas unitarias y de integración

Al concluir la implementación del sistema se le aplicaron pruebas unitarias a los flujos de integración de datos y a la capa de visualización, arrojando como resultado tres no conformidades en cada caso, para un total de seis, las cuales fueron resueltas satisfactoriamente.

2.4.2 Pruebas de sistema

Las pruebas del sistema al MD Cooperación internacional se aplicaron por parte del grupo de calidad del departamento de AD, encontrándose tres no conformidades de complejidad alta, todas resueltas satisfactoriamente. Quedando el sistema liberado para su despliegue en el área.

2.4.3 Pruebas de aceptación

En conjunto con el cliente se realizaron las pruebas de aceptación de la solución, las cuales arrojaron resultados satisfactorios, quedando comprobado que el sistema cumple con sus necesidades y que están satisfechos con el producto elaborado.

3 CONCLUSIONES

Con la realización de la investigación se logró cumplir con los objetivos planteados, desarrollándose un mercado de datos poblado y funcional que ofrece soporte a la toma de decisiones en el área de Cooperación internacional de la Universidad de las Ciencias Informáticas. Por lo que se concluye que:

1. El estudio de los fundamentos teóricos de la investigación, permitió seleccionar una metodología para organizar de manera estructurada el proceso de desarrollo de software.
2. El análisis y diseño del mercado de datos, posibilitó un mejor entendimiento del proceso de negocio de la dirección de Cooperación internacional permitiendo la implementación de una solución que responde a las necesidades del cliente.
3. La implementación del mercado de datos contribuye al proceso de apoyo a la toma de decisiones mediante el desarrollo de los diferentes subsistemas, garantizando la correcta organización, carga y visualización de los datos.
4. La validación de la solución mediante la aplicación de los casos de pruebas, el perfilado de los datos y las listas de chequeo permitió obtener

un mercado funcional que cuenta con una correcta calidad de los datos y que cumple con los requisitos especificados por el cliente.

4 REFERENCIAS BIBLIOGRÁFICAS

1. **Bernabeu, Ricardo Dario.** HEFESTO: Metodología propia para la Construcción de un Data Warehouse. Córdoba, Argentina : s.n., 2007.
2. **Giménez, Matilde Celma.** Almacenes de Datos. Departamento de Sistemas Informáticos y Computación, Universidad Politécnica de Valencia. España : s.n. 2008.
3. **Zorrilla, Marta.** Data warehouse y OLAP. Cantabria, España : s.n., 2007 – 2008.
4. **Dapena Bosquet, Isabel. Muñoz.** Sistemas de Información Orientados a la Toma de Decisiones: el enfoque multidimensional. Madrid, España: s.n., 2005.
5. **Kimball, Ralph and Caserta, Joe.** The Data Warehouse ETL Toolkit. Canadá : Wiley Publishing, 2004.
6. **Kimball, Ralph and Ross, Margy.** The Datawarehouse Toolkit. Second Edition. s.l.: John Wiley & Sons, Inc, 2002. Vol. The Complete Guide to Dimensional Modeling.
7. **Inmon, W.** Building the Data Warehouse. 2nd edition. John Wiley & Sons. 1996.
8. **González Hernández, Ing. Yanisbel.** PROPUESTA DE METODOLOGIA PARA EL DESARROLLO DE ALMACENES DE DATOS EN DATEC. 2012.
9. **Corporation, IBM.** Ayuda de Rational Unified Process. [Online] 7.0.1, 1987-2006.
10. **Pressman, Roger S.** Ingeniería del software. Un enfoque práctico. s.l. : McGrawHill. Vol. 6ta edición. ISBN-970-10-5473-3.
11. **González Pompa, Yisel de los Angeles y Rosales González, Maria Teresa.** 2012. Mercado de datos Cooperación internacional para la Sala Situacional de la Universidad de las Ciencias Informáticas. : s.n., 2012.

5 SÍNTESIS CURRICULARES DE LOS AUTORES

Fecha de nacimiento: 3 de mayo de 1988,
Lugar de nacimiento: Veguitas, Yara, Granma.
Nivel educacional: Universtaria. Graduada de la Universidad de las Ciencias Informáticas, 2012.
Ocupación Actual: Recién graduada en adiestramiento.
Centro de trabajo: Universidad de las Ciencias Informáticas Facultad Regional Granma.
Publicaciones anteriores: Ninguna.
Dirección particular: Calle 3ra % 10 y 12 #253 Reparto Francisco Vicente Aguilera. Bayamo. Granma. Cuba.
Dirección electrónica: yisel.pompa@grm.uci.cu

