

Universidad de las Ciencias Informáticas

Facultad 6



Título:

“Capa de visualización del mercado de datos de Tecnología para la Sala Situacional de la Universidad de las Ciencias Informáticas.”

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

Autora:

Romy Rodríguez Pino.

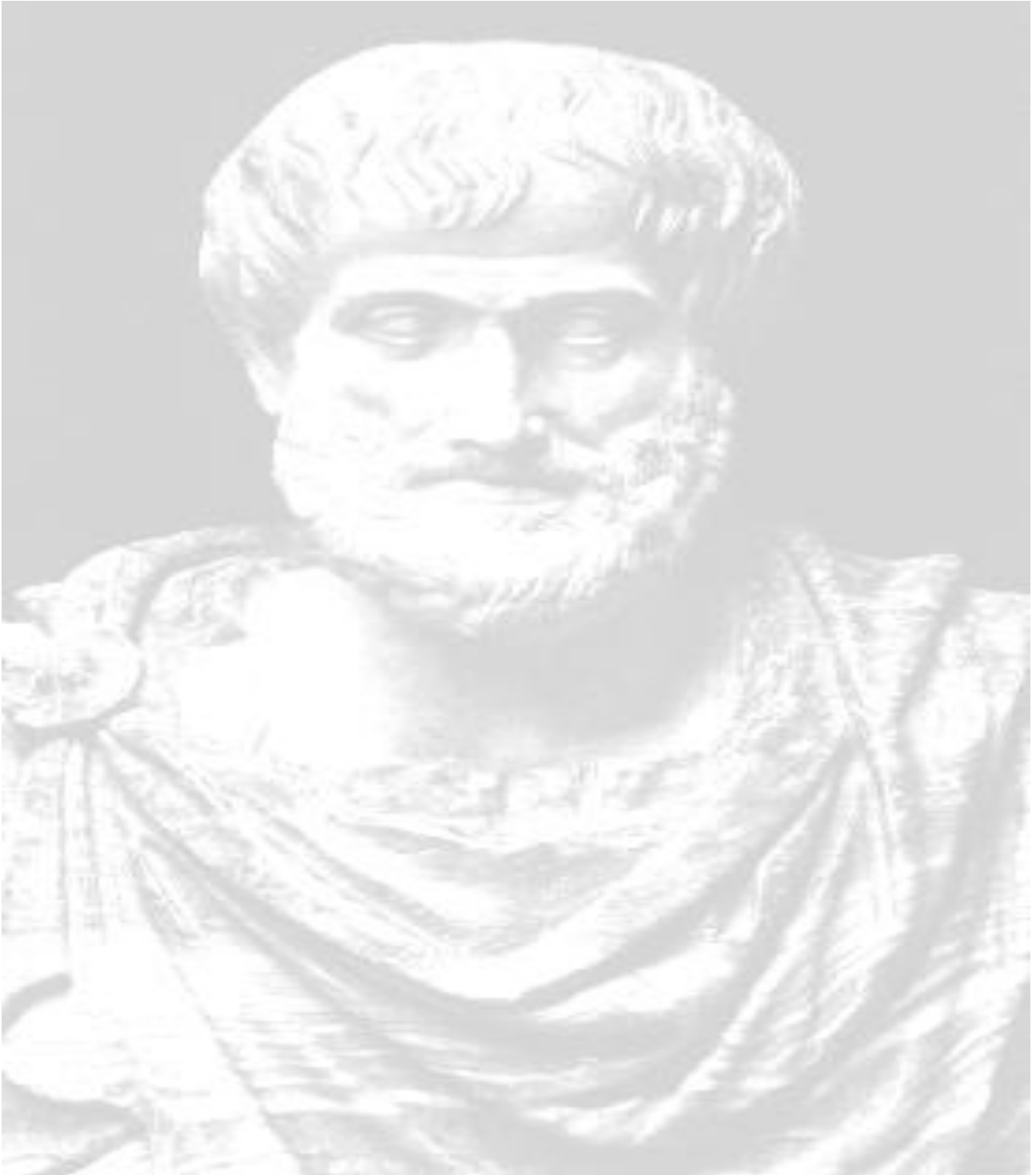
Tutores:

Ing. Haymee Llerena Esperón.

Ing. Fabián López García.

La Habana, junio 2012

“Año 54 de la Revolución”



La inteligencia consiste no sólo en el conocimiento, sino también en la destreza de aplicar los conocimientos en la práctica.

Aristóteles

Declaro ser autora del presente trabajo “Capa de visualización del mercado de datos Tecnología para la Sala Situacional de la Universidad de las Ciencias Informáticas” y reconozco a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo.

Para que así conste firmo la presente a los ____ días del mes de _____ del año 2012.

Autora: _____
Romy Rodríguez Pino.

Tutores: _____
Ing. Fabián López García.

Ing. Haymee Llerena Esperón.

Datos de contacto

Tutores:

Ing. Haymee Llerena Esperón

Especialidad en graduación: Ingeniería en Ciencias Informáticas

Categoría docente: Instructora

Categoría científica:

Años de experiencia en el tema:

Años de graduado: 6

Correo electrónico: hllerena@uci.cu

Ing. Fabián López García

Especialidad en graduación: Ingeniería en Ciencias Informáticas

Categoría docente:

Categoría científica:

Años de experiencia en el tema: 2

Años de graduado: 1

Correo electrónico: flgarcia@uci.cu

Dedico este trabajo a toda mi familia, en especial a mi mamá por todo su amor, su dedicación y sacrificio para que este sueño se me hiciera realidad.

Agradezco de todo corazón a las personas más importantes en mi vida, a mi mamá por ser mi motor impulsor, mi guía, por darme fuerzas para seguir, por sus consejos, y por toda su confianza en mí. A mi hermano que lo adoro con mi alma, a mi papá, a mis padrinos por ser mis otros padres, por darme tanto amor todo el tiempo, por ser tan adorables.

A Luis Alejandro por su incondicionalidad, por estar justo en el momento que siempre más necesité. Por todo su cariño, su tiempo, su paciencia, su preocupación, por aguantarme estos últimos años.

A mis amigos de estos 5 años los cuales hemos compartidos tantos momentos inolvidables, por su ayuda a Andry, Damián, Ernesto, Ramiro, Ayito.

A Yohana, Rebeca, Tahimy, Yeilen por ser tan buenas amigas, por darme su confianza, por todas las cosas que hemos vivido juntas, por hacerme reír, por los buenos y malos momentos. A Lianet, que aunque vivimos solamente el último año juntas, has demostrado ser una persona maravillosa.

A Fabián y Rayko por toda su preocupación, dedicación en todo momento, por su tiempo para ayudarme en lo que me hiciera falta, siempre les estaré agradecida.

RESUMEN

Con el propósito de informatizar las áreas de la Universidad de las Ciencias Informáticas, se crea el Almacén de Datos Sala Situacional UCI, que tiene como objetivo apoyar los procesos de toma de decisiones en la universidad. Para ello cuenta con varios mercados de datos, que se encargan de almacenar la información de las diferentes áreas, en los cuales se encuentra el mercado de datos Tecnología. Este contiene los datos del uso y explotación de las computadoras, y cuenta con una capa de presentación de datos que no cumple con las expectativas de los principales directivos de la universidad.

La presente investigación se enmarca en el área de inteligencia de negocio y sus principales herramientas para la realización de la presentación de los datos. Para dar cumplimiento a la problemática planteada se implementaron vistas lógicas, que facilitaron el diseño de los cubos multidimensionales y la creación de las vistas de análisis. Para realizar un análisis adecuado de los indicadores claves del negocio se muestra la información mediante cuadros de mando digitales y gráficos dinámicos. Como resultado de la investigación se obtuvo la capa de visualización del mercado de datos Tecnología, facilitando el análisis de la información a los directivos de laboratorios en la universidad.

Palabras Claves: almacén de datos, capa de visualización, inteligencia de negocio, cuadros de mando digitales, gráficos dinámicos, toma de decisiones.

Índice

ÍNDICE DE TABLASIX

ÍNDICE DE FIGURASX

INTRODUCCIÓN 1

CAPÍTULO 1: Fundamentos teóricos de la inteligencia de negocio..... 4

 Introducción.....4

 1.1 Almacenes de datos operacionales 4

 1.2 Inteligencia de negocio5

 1.2.1 Ejemplos prácticos de la inteligencia de negocio6

 1.3 Metodología de desarrollo 6

 1.4 Técnicas de inteligencia de negocio7

 1.4.1 Cuadro de mando digital8

 1.4.2 Modos de almacenamiento de datos.....8

 1.5 Herramientas de inteligencia de negocio 10

 1.5.1 IBM Cognos Business Intelligence 10

 1.5.2 MicroStrategy 11

 1.5.3 Oracle Business Intelligence Enterprise11

 1.5.4 Pentaho BI Suite 12

 1.6 Herramientas de desarrollo de la solución 14

 1.7 Indicadores y métricas 16

Conclusiones del capítulo 16

CAPÍTULO 2: Análisis y diseño..... 17

 Introducción..... 17

 2.1 Arquitectura de la solución..... 17

 2.2 Especificación de requisitos..... 18

 2.2.1 Requisitos de información 18

 2.2.2 Requisitos funcionales..... 19

 2.2.3 Requisitos no funcionales..... 19

 2.3 Reglas del negocio 20

 2.4 Casos de usos del sistema 21

 2.4.1 Actores del sistema 22

 2.4.2 Casos de usos de información 22

 2.4.3 Casos de usos funcionales..... 24

 2.5 Diagrama de casos de uso 26

 2.6 Reglas de navegación 27

 2.6.1 Agrupaciones de información y jerarquías 27

 2.7 Elementos de la regla de navegación 28

 2.8 Diseño de las vistas lógicas 29

 2.9 Diseño del subsistema de visualización..... 30

 2.9.1. Diseño de los cubos OLAP 30

2.9.2.	Diseño de la arquitectura de información	30
2.9.3.	Descripción de los libros de trabajo	31
2.10	Esquema de seguridad	32
	Conclusiones del capítulo	32
	CAPÍTULO 3: Implementación y prueba	34
	Introducción.....	34
3.1	Implementación de los cubos OLAP	34
3.2	Arquitectura de información	35
3.3	Implementación de las vistas de análisis	36
3.4.1	Tipos de gráficos	38
3.4	Pruebas	42
3.5	Herramientas para la realización de las pruebas	43
3.6	Resultado de las pruebas	45
	Conclusiones del capítulo	48
	Conclusiones generales	49
	Recomendaciones	50
	Referencia bibliográfica	51
	Bibliografía	53
	Anexos	55
	Glosario de términos	69

ÍNDICE DE TABLAS

Tabla 1: Actores del sistema 22

Tabla 2: Descripción textual del CUI Obtener información de explotación..... 22

Tabla 3: Descripción textual del CUF Administrar roles. 24

Tabla 4: Seguridad en la aplicación 32

Tabla 5: Diseño del caso de prueba inventario 43

Tabla 9: Descripción textual del CU: Obtener información de procesos 61

Tabla 10: Descripción textual del CU: Obtener información de inventario 62

Tabla 11: Diseño del caso de prueba explotación..... 63

ÍNDICE DE FIGURAS

Figura 1: Arquitectura de la solución..... 17

Figura 2: Diagrama de CUS..... 26

Figura 3: Arquitectura de la información..... 31

Figura 4: Representación de los cinco cubos del mercado de datos Tecnología..... 34

Figura 5: Representación de los elementos que componen el cubo métricas de disco duro. Se visualiza las dimensiones y medidas asociadas. 35

Figura 6: Representación de las dimensiones y la jerarquía que la compone. 35

Figura 7: Arquitectura de información: el AA contiene los LT, en los que se ubican las vistas de análisis, cuadros de mandos digitales y gráficos dinámicos..... 36

Figura 8: Vista de análisis correspondiente al espacio libre y espacio usado del disco duro..... 37

Figura 9: Vista de análisis correspondiente a la ejecución de procesos por categorías. 37

Figura 10: Vista de análisis correspondiente a la cantidad de computadoras por tipo de RAM..... 38

Figura 11: Cuadro de mando digital correspondiente al aprovechamiento de las computadoras. 40

Figura 12: Cuadro de mando digital correspondiente al consumo de RAM de las computadoras. 40

Figura 13: Gráfico dinámico correspondiente a la cantidad de paquetes recibidos y enviados por la red. 41

Figura 14: Gráfico dinámico correspondiente al promedio de las observaciones por laboratorios..... 41

Figura 15: Modelo V. Permite realizar pruebas a lo largo del ciclo de desarrollo del software..... 43

Figura 16: Resultados de las pruebas unitarias y de integración. 46

Figura 17: Resultados de la aplicación de los casos de prueba basados en CU al sistema. 47

Figura 18: Resultado general de las pruebas aplicadas a la solución..... 47

Figura 19: Vista lógica métricas estáticas 55

Figura 20: Vista lógica métricas de disco duro..... 55

Figura 21: Vista lógica métricas utilización..... 56

Figura 22: Vista lógica procesos por usuarios..... 56

Figura 23: Representación del cubo métricas de RAM y red. 57

Figura 24: Representación del cubo métricas estáticas. 57

Figura 25: Representación del cubo procesos por usuario. 58

Figura 26: Representación del cubo métricas de utilización. 58

Figura 27: Carta de aceptación del cliente..... 59

Figura 28: Acta de liberación interna del producto de software..... 60

INTRODUCCIÓN

La asombrosa velocidad con la que avanza la tecnología, ha dejado prácticamente en el olvido los tiempos en los que el hombre apenas soñaba con herramientas tan útiles como las que se tienen hoy día. El auge que ha tenido la implantación y utilización de los avances tecnológicos para elevar el nivel de vida de la población ha sido un acontecimiento importante para el desarrollo económico y social a nivel mundial.

Muchas empresas basan sus movimientos en los resultados que arrojan los estudios desarrollados por sus analistas. Estos estudios se realizan sobre la información generada en el quehacer cotidiano de las empresas, haciendo énfasis en el almacenamiento y uso de los datos. De esta forma se originaron estructuras potentes como los almacenes de datos, que brindan la posibilidad de consolidar contenidos relevantes de fuentes dispersas de información en un único repositorio. Esta solución presenta la capacidad de generar conocimientos a partir de la información recogida en ella mediante el uso de la inteligencia de negocio. La automatización de los procesos que se llevan a cabo en estas tecnologías facilita la comparación de resultados obtenidos por una institución, ofreciendo una estrategia a seguir para la toma de decisiones.

Cuba, no se encuentra exenta de la modernización tecnológica, y se ha propuesto el reto de informatizar el país. El desarrollo de esta actividad se ha efectuado rápidamente, surgiendo así, la Universidad de las Ciencias Informáticas (UCI). Esta universidad cuenta con múltiples centros productivos, que aportan grandes ganancias a la economía del país y a la formación de profesionales. Dentro de estos centros se encuentra el Centro de Tecnologías de Gestión de Datos (DATEC), el cual cuenta con cuatro líneas de desarrollo, una de ellas es la de Almacenes de Datos. Actualmente el departamento tiene desplegado el almacén de datos operacional Tecnología, que recopila los datos relacionados con la explotación de las estaciones de trabajo y el control de inventario de software y hardware. El cual se nutre de la información recogida por el sistema informático distribuido T-arenal, que monitorea periódicamente las computadoras de cada docente, generando grandes volúmenes de datos que son procesados por dicha solución.

El mercado de datos Tecnología cuenta con una capa de visualización que dificulta el proceso de análisis de la información por parte de los especialistas del área de tecnología, debido a que la información no se muestra de manera clara y sencilla. No se puede realizar el estudio de los principales indicadores del negocio de forma rápida, que permita conocer el estado actual de la institución, evaluando de esta manera la evolución del negocio. Por todo esto se hace necesario hacer uso de la inteligencia de negocio, que permita mejorar el proceso de toma de decisiones de los directivos de laboratorio de la universidad.

Por todo lo anteriormente planteado, surge como **problema de la investigación**: ¿Cómo contribuir a la toma de decisiones en el mercado de datos Tecnología de la Sala Situacional de la UCI?

La presente investigación tiene como **objeto de estudio** las soluciones de inteligencia de negocio, enmarcado en el **campo de acción** Capa de visualización del mercado de datos Tecnología.

Para darle solución al problema de la investigación, se define como **objetivo general** desarrollar una capa de visualización para el mercado de datos Tecnología de la Sala Situacional UCI que contribuya a la toma de decisiones.

En correspondencia con el objetivo general propuesto se definieron los siguientes **objetivos específicos**:

- ✓ Fundamentar la selección de las metodologías, herramientas y tecnologías a utilizar en el desarrollo de la capa de visualización para el mercado de datos Tecnología de la Sala Situacional UCI.
- ✓ Diseñar los componentes de la capa de visualización para el mercado de datos Tecnología de la Sala Situacional UCI.
- ✓ Implementar la capa de visualización para el mercado de datos Tecnología de la Sala Situacional UCI.
- ✓ Probar la capa de visualización para el mercado de datos Tecnología de la Sala Situacional UCI.

Para dar cumplimiento a los objetivos específicos planteados, se proponen las siguientes **tareas de la investigación**:

- ✓ Selección de las herramientas, tecnologías y métodos para el desarrollo de la capa de visualización.
- ✓ Refinamiento de los requisitos de información.
- ✓ Definición de las vistas lógicas.
- ✓ Definición de las vistas de análisis.
- ✓ Implementación de los cubos multidimensionales.
- ✓ Implementación de las vistas de análisis.
- ✓ Implementación de los gráficos dinámicos.
- ✓ Implementación de los cuadros de mando digitales.
- ✓ Aplicación de los casos de prueba.

El presente trabajo se encuentra estructurado en 3 capítulos:

Capítulo1: Fundamentos teóricos de la inteligencia de negocio. Este capítulo constituye la base teórica de la investigación y aborda los resultados del estudio del estado del arte. Describe los conceptos

principales que se van a tratar durante su desarrollo, así como, la metodología utilizada e incluye el análisis de las herramientas para el desarrollo de la aplicación.

Capítulo 2: Análisis y diseño de la capa de visualización para el mercado de datos Tecnología. En este capítulo se realiza el diseño de las vistas lógicas y del subsistema de visualización. Teniendo como resultado la arquitectura de información y las vistas de análisis.

Capítulo 3: Implementación y prueba. En este capítulo se implementa las vistas lógicas, los cubos multidimensionales y el subsistema de visualización mediante los gráficos dinámicos y cuadros de mando digitales. Luego se realiza el proceso de pruebas de la capa de visualización mediante la herramienta casos de prueba basados en casos de uso. Finalmente se emite la carta de aceptación firmada por el cliente.

CAPÍTULO 1: Fundamentos teóricos de la inteligencia de negocio

Introducción

En el presente capítulo se abordan los conceptos fundamentales relacionados con los Almacenes de Datos Operacionales¹ y la Inteligencia de negocio², así como, la metodología de desarrollo que se va a utilizar y las principales herramientas para el desarrollo de la aplicación. Mediante las descripciones de estos aspectos teóricos se pretende facilitar el entendimiento de la investigación.

1.1 Almacenes de datos operacionales

Un problema frecuente de las empresas en la actualidad es que poseen grandes volúmenes de datos dispersos en múltiples fuentes de información, esto provoca demoras en la obtención de los resultados y obstaculiza el acceso a la información. Con la intención de mejorar estos problemas, se crea un sistema que permita integrar física y lógicamente los datos fuentes, funcionando como un repositorio de información, de este modo surgen los almacenes de datos.

Un ODS es una colección de datos actuales, integrados y volátiles que están orientados a un tema. Puede ser usado como un área transitoria para la reorganización física de los datos operacionales extraídos, proporcionar informes operacionales y apoyar la toma de decisiones (1).

Es un repositorio de datos de la organización, cuyo objetivo es proveer a los niveles tácticos de la empresa, información integrada para el análisis operacional. Dicho repositorio se alimenta de los sistemas operacionales de la organización (2).

Posee las siguientes características (2):

- ✓ Orientado a un tema: es diseñado y organizado alrededor de los temas principales de la organización.
- ✓ Integrado: los datos son transformados en un todo consistente y unificado, dando una visión corporativa unificada.
- ✓ Volátil: los datos son actualizados en cada carga.
- ✓ Evaluado actual: los datos típicamente son actuales.

Entre las principales ventajas que brinda un ODS se encuentran (3):

- ✓ Ahorro de costes por la reutilización de recursos.
- ✓ Disponibilidad de un dato único, evitando múltiples copias del mismo dato en sistemas dispersos.
- ✓ Posibilidad de cruzar la información de diferentes aplicaciones.

¹ Del inglés Operational Data Store, por sus siglas ODS.

² Del inglés Business Intelligence, por sus siglas BI.

Por los elementos anteriormente analizados se entiende por ODS, que es un contenedor de datos operacionales que ayudan al soporte de decisiones y a la operación. Su función es integrar los datos con un mayor nivel de detalle.

1.2 Inteligencia de negocio

Con el transcurso del tiempo la información se ha convertido en uno de los recursos más importantes de cualquier organización. Para elevar el entorno competitivo dentro de una empresa se necesita tomar certeras decisiones, una manera de lograr esto, es por medio de la inteligencia de negocio, la cual se utiliza para comprender, mejorar el rendimiento y reducir los costes e identificar nuevas oportunidades de negocio.

La inteligencia de negocio se puede definir como el proceso de analizar los bienes o datos acumulados en la empresa y extraer cierta inteligencia o conocimiento de ellos. Es una herramienta que permite a través de reportes, consultas y el análisis de datos brindar a los usuarios la información valiosa de forma más sintetizada, con el objetivo de facilitar el proceso de toma de decisiones.

Según el Instituto de Almacenes de Datos la definición de inteligencia de negocios es la siguiente:

“Son los procesos, tecnologías y herramientas que se necesitan para convertir los datos en información, la información en conocimiento y el conocimiento en planes que impulsan acciones rentables para el negocio. La inteligencia de negocios abarca el almacenamiento de datos, herramientas analíticas, contenido y gestión del conocimiento” (4).

Las características que tienen en común este conjunto de técnicas y herramientas son las siguientes (5):

- ✓ Accesibilidad a la información: los datos son la fuente principal. Se garantiza el acceso de los usuarios a los datos con independencia de la procedencia de estos.
- ✓ Apoyo en la toma de decisiones: es ir más allá en la presentación de la información, de manera que los usuarios tengan acceso a herramientas de análisis que les permitan seleccionar y manipular sólo aquellos datos que les interesen.
- ✓ Orientación al usuario final: se busca independencia entre los conocimientos técnicos de los usuarios y su capacidad para utilizar estas herramientas.

La inteligencia de negocio tiene como beneficio una serie de aspectos como los que se explican a continuación (6):

- ✓ Reducción de costos: permite al cliente tener acceso a los datos posibilitando una mejor información y satisfacción para el mismo.
- ✓ Incremento en los ingresos: logra medir el desempeño de la organización en función de sus metas y la industria en donde compite.

- ✓ Mejorar la satisfacción del cliente: brinda a sus usuarios las herramientas para tomar mejores decisiones, proporcionando rápidas respuestas a las preguntas de los mismos.

1.2.1 Ejemplos prácticos de la inteligencia de negocio

Galicia es una empresa conservera de gran prestigio internacional, cuenta con más de 500 empleados y cerca de 100.000.000 € de facturación. Esta empresa no había sido capaz de mejorar la cantidad de producto finalizado que debía almacenar para maximizar sus beneficios. Mediante la utilización de la inteligencia de negocio y tras el análisis minucioso de los datos históricos que guardaba la compañía, resultó posible rediseñar todo el proceso logístico y de almacenamiento productivo hasta el punto de incrementar la rentabilidad económica de la misma en un 10%, independientemente de la producción y la demanda (7).

Un Banco consigue establecer una unión entre sus antiguos sistemas abiertos y heredados con los nuevos datos aportados en las bases de datos departamentales. De este modo es posible identificar quienes son los clientes más rentables de una determinada oficina, y que productos se les puede ofrecer con garantías de obtener ventas cruzadas. Gracias a estas herramientas, el departamento de informática se ve liberado de tener que realizar todas estas tareas y deja que sean los propios usuarios y directores de oficina los que asuman estas funciones (7).

1.3 Metodología de desarrollo

Para el desarrollo de soluciones informáticas se deben realizar acciones encaminadas a lograr el éxito de la solución, para ello se hace necesario seleccionar la metodología que guíe el proceso de desarrollo. Una metodología tiene como principal objetivo aumentar la calidad del software que se produce en todas y cada una de sus fases de desarrollo. A nivel mundial existen dos enfoques para las soluciones de inteligencia de negocio que se corresponden a las metodologías propuesta por Ralph Kimball³ y Bill Inmon⁴.

La principal diferencia que existe entre ambas metodologías está basada en la forma de enfrentar el problema. La visión de Inmon se basa en un enfoque descendente (top-down), propone construir primero el almacén de datos y a partir de este, los mercados de datos; plantea la creación de un repositorio de datos corporativo como fuente de información consolidada, persistente, histórica y de calidad. Al ser construido descendentemente los mercados de datos se nutren del almacén de datos corporativo, convirtiéndose en un complejo empresarial de bases de datos relacionales (8).

³ Ralph Kimball, especialista en el diseño de almacenes de datos y creador del enfoque multidimensional.

⁴ Bill Inmon, creador del término de almacenes de datos y considerado padre de la disciplina.

La propuesta de Kimball se basa en dividir el mundo de inteligencia de negocio entre los hechos y las dimensiones, que conduce a una solución completa en un corto período de tiempo. Es una metodología que se referencia como ascendente (bottom-up), porque plantea que se deben construir primeramente los mercados de datos orientados al tema de su departamento y luego el almacén de datos será la unión de estos (9).

Para guiar el proceso de desarrollo de la capa de visualización del mercado de datos Tecnología, se escoge como metodología la Propuesta de Metodología para el Desarrollo de Almacenes de Datos en DATEC. La cual atraviesa todas las fases de construcción de un almacén de datos, desde el levantamiento de requisitos de información hasta la capa de visualización. Es una metodología mixta que toma como base la propuesta de Ralph Kimball por los siguientes elementos:

- ✓ Identifica las tablas de hechos y dimensiones, lo cual agiliza el proceso de desarrollo y con ello la toma de decisiones.
- ✓ Propone la construcción de mercados de datos departamentales y después el almacén de datos. Esto trae como ventaja que la creación y la puesta en marcha de los mercados de datos se produce en un lapso de tiempo corto, y después se valora si se construye o no el almacén de datos.
- ✓ Existe amplia documentación de la misma, así cualquier duda que exista puede ser atendida rápidamente.
- ✓ Es una metodología madura y reconocida por los usuarios dedicados al tema, además, tiene bien definidas sus etapas, actividades, roles y artefactos.

Esta propuesta toma lo planteado por Leopoldo Zenaido Zepeda en su tesis de doctorado, de incluir los casos de uso para guiar el proceso de desarrollo, y así lograr estar más alineados a las tendencias y normas de trabajo de la universidad. Se agrega además una etapa de prueba que permite comprobar la calidad de los productos que se desarrollan (10).

1.4 Técnicas de inteligencia de negocio

Las técnicas de inteligencia de negocio ayudan a las instituciones empresariales a sacar provecho de la información para incrementar la agilidad corporativa, optimizar el rendimiento de sus procesos y alcanzar objetivos estratégicos. Su principal propósito es facilitar el proceso de la toma de decisiones, resolviendo limitaciones que a menudo demoran las capacidades de respuesta de las instituciones. A continuación se describirán las técnicas que se utilizarán durante el desarrollo de la solución.

1.4.1 Cuadro de mando digital

Un Cuadro de Mando Digital (CMD) es un tablero compuesto por diversas tablas y gráficas, que recoge los principales indicadores del negocio y los presenta de manera clara y sencilla. Es una herramienta que permite realizar el análisis de las diferentes variables de negocio e informa de la evolución de los parámetros fundamentales. Está diseñado para generar un mayor impacto visual al captar rápidamente la información. Constituye una de las vistas más integrales, debido a su característica de dinamismo para mostrar la información, dando a conocer las distintas tecnologías que brindan soporte y agilizan el proceso de toma de decisiones.

“Es una pantalla de visualización de la información, necesaria para lograr uno o más objetivos, consolidada y organizada en una sola pantalla de modo que la información se pueda controlar de un vistazo” (11).

Objetivos de un CMD (6):

- ✓ Información clara y sencilla.
- ✓ Datos relevantes para el negocio.
- ✓ Comparación entre diferentes ejes y dimensiones.
- ✓ Rapidez en la implementación.
- ✓ Útil para la toma de decisiones.
- ✓ Revisión periódica de la información.

Ventajas que brindan los CMD (6):

- ✓ Visión general inmediata de la organización mediante gráficos dinámicos.
- ✓ Detección y corrección de tendencias negativas.
- ✓ Controla las estadísticas vitales de rendimiento.
- ✓ Permite a la empresa enfocarse en las prioridades, eliminando la sobrecarga de información, presentando solo lo relevante.

1.4.2 Modos de almacenamiento de datos

El Procesamiento Analítico en Línea⁵ (OLAP) es una tecnología de software que se puede utilizar por varias herramientas para el análisis de información sobre un almacén de datos. Esta tecnología posee la capacidad de presentar los datos objetos de análisis a través de un modelo multidimensional, llamados cubos OLAP, estos permiten al usuario efectuar un estudio certero desde diferentes puntos de vista. Las herramientas que utilizan estas tecnologías están diseñadas para generar respuestas rápidas a complejas consultas analíticas (12).

Las características básicas de estas herramientas son (13):

⁵ Del inglés Online Analytical Processing, por sus siglas OLAP.

- ✓ Permite recolectar y organizar la información analítica necesaria para los usuarios y disponer de ella en diversos formatos como: tablas, gráficos, reportes, tableros de control, entre otros.
- ✓ Soporta análisis complejos de grandes volúmenes de datos.
- ✓ Complementa las actividades de otras herramientas que requieran procesamiento analítico en línea.
- ✓ Presenta al usuario una visión multidimensional de los datos para cada tema de interés del negocio.

Esta herramienta tiene como ventajas el acceso a las grandes cantidades de datos, permitiendo su análisis desde diferentes perspectivas. Además involucra cálculos complejos entre los datos y brindan rápidas respuestas a los usuarios. Existen tres modos para el proceso analítico en línea de la información: ROLAP, MOLAP y HOLAP.

Procesamiento Analítico en Línea Relacional⁶ (ROLAP): accede a los datos almacenados en un almacén de datos para proporcionar los análisis OLAP. Toda la información es almacenada sobre una base de datos relacional. Se encarga de transformar el esquema relacional del almacén en un esquema multidimensional para el usuario. Principalmente se utiliza para los datos históricos más recientes (14).

Ventajas:

- ✓ Uso total de la seguridad e integridad de la base de datos.
- ✓ Escalable para grande volúmenes de datos.

Desventajas:

- ✓ Consultas más lentas.
- ✓ Los cálculos están limitados a las funciones de la base de datos.

Procesamiento Analítico en Línea Multidimensional⁷ (MOLAP): almacena los datos fuentes del cubo en una estructura multidimensional de alto rendimiento. Utiliza una arquitectura de dos niveles: la base de datos multidimensional, que es la encargada del manejo, acceso y obtención de los datos y el motor analítico (14).

Ventajas:

- ✓ Mejor rendimiento en el tiempo de respuesta.
- ✓ Puede escribir sobre las bases de datos.
- ✓ Posibilita ejecutar cálculos más complicados.

⁶ Del inglés Relational Online Analytical Processing, por sus siglas ROLAP.

⁷ Del inglés Multidimensional Online Analytical Processing, por sus siglas MOLAP.

Desventajas

- ✓ Tamaño limitado para la arquitectura del cubo.
- ✓ No se puede acceder a los datos que no estén en el cubo.

Procesamiento Analítico en Línea Híbrido⁸ (HOLAP): combina atributos de MOLAP y ROLAP, o sea, los datos son almacenados en una estructura multidimensional y la base de datos fuentes, en una base de datos relacional. Es usado para cubos que necesitan rápidas respuestas (14).

Ventajas:

- ✓ Mejor tiempo de respuesta para información resumida.

Desventajas:

- ✓ Elevado costo de mantenimiento para las empresas.

Luego de realizarse el análisis de los diferentes modos de almacenamiento, el tipo de procesamiento a utilizar para el desarrollo es el ROLAP, porque el sistema gestor de base datos PostgreSQL es un gestor de objetos relacionales y el enfoque multidimensional se crea mediante la implementación de los cubos multidimensionales. Además, esta arquitectura accede directamente a los datos almacenados en el almacén de datos y soporta técnicas de optimización de accesos para mejorar el tiempo de respuesta de las consultas.

1.5 Herramientas de inteligencia de negocio

1.5.1 IBM Cognos Business Intelligence

Es un software que se basa en una plataforma abierta empresarial capaz de entregar información consistente y oportuna, el mismo cumple con las exigencias de la inteligencia de negocio y gestión de rendimiento. Posee una arquitectura orientada a servicios y una infraestructura escalable. IBM Cognos Business Intelligence combina funcionalidades que permiten crear cubos OLAP, paneles de instrumentos, tarjetas de puntuación con todos los orígenes de datos y generar informes, lo que convierte a este software en una ayuda para mejorar el proceso de la toma de decisiones. Estas son las principales herramientas que proporciona la suite (15):

- ✓ IBM Cognos Query Studio: es una herramienta que realiza consultas sencillas al usuario. Permite acceder a la misma estructura de datos que utilizan otras herramientas. Aplica filtros, ordenaciones, operaciones de agrupación de datos e incluso crear gráficas.
- ✓ IBM Cognos Analysis Studio: permite la navegación por estructuras multidimensionales como cubos OLAP. Su objetivo general se centra en que el analista pueda navegar por los datos

⁸ Del inglés Hybrid Online Analytical Processing, por sus siglas HOLAP.

cargados en estructuras dimensionales. realiza análisis complejos y comparativos de datos para descubrir tendencias, riesgos y oportunidades.

- ✓ IBM Cognos Event Studio: crea notificaciones y alertas cuando se detectan eventos importantes para el negocio.

Entre las desventajas que presenta la herramienta se encuentran (15):

- ✓ Alta complejidad: es difícil de implementar, de trabajar y mantener.
- ✓ Alto costo: las soluciones resultan difíciles de pagar, debido a los fuertes requerimientos de las tecnologías que necesitan los productos de esta herramienta.

1.5.2 MicroStrategy

Basa su software en el uso de la tecnología de base de datos relacional que permite crear informes y análisis de los datos almacenados en una base de datos relacional. Presenta mejoras de rendimiento y la capacidad para realizar análisis interactivos a través de cuadros de mando y nuevas aplicaciones. Con esta herramienta se puede realizar un completo monitoreo y manejo de las aplicaciones de inteligencia del negocio, manejar fácilmente el ciclo de vida de los proyectos así como reducir los costos de implementación. Entre las características de este software se encuentran (16):

- ✓ Manejo y mantenimiento centralizado.
- ✓ Manejo integral de objetos de esquema y de la aplicación.
- ✓ Permite a los usuarios diseñar nuevos informes, mediante una interfaz y una serie de herramientas muy intuitivas.

Entre las desventajas que presenta esta herramienta se encuentran (16):

- ✓ La implantación de la herramienta es costosa, tanto económicamente como en tiempo.
- ✓ La facilidad de implementación a veces limita las posibilidades que este tipo de herramientas debe ofrecer, porque le faltan funcionalidades que permitan mejorar el proceso de diseñar reportes.

1.5.3 Oracle Business Intelligence Enterprise

Es una plataforma de negocios completa con una amplia oferta en capacidades de análisis y reporte, facilitando a las empresas el análisis profundo de las métricas de desempeño. La herramienta entrega reportes, tablas comparativas y análisis OLAP, dándole al usuario acceso a su almacén de datos, para hacer análisis avanzados y útiles para directivos, consultores y analistas. A continuación se presentan algunas ventajas (17):

- ✓ Gran velocidad de acceso a la información.

- ✓ Seguridad, integra módulos de seguridad de acceso a la información, adaptables a cada perfil de usuario e incluso a cada usuario mismo, con una interfaz muy intuitiva para su mantenimiento.

Posee una serie de desventajas como son (17):

- ✓ Orientada únicamente a grandes empresas debido a su precio.
- ✓ El tiempo de desarrollo puede tardar años para proyectos simples.
- ✓ En el aspecto técnico, las aplicaciones de inteligencia de negocio desarrolladas con esta herramienta son muy estáticas y su evolución al análisis de nuevas aéreas de negocio es complejo.

1.5.4 Pentaho BI Suite

Para el desarrollo de la solución se utilizará la plataforma Pentaho BI Suite, porque es capaz de innovar mucho más rápido que los proveedores comerciales, y ofrece una alternativa de código abierto que supera a las soluciones de inteligencia de negocio propietarias en muchas áreas como arquitectura, soporte de estándares, funcionalidad y simplicidad de implantación. Está basada en tecnología Java al igual que su ambiente de implementación, lo que la hace una herramienta flexible y adaptable a varios ambientes. Esta plataforma se enfoca principalmente en la consultoría y prestación de múltiples servicios de tecnología de información a través de módulos de reportes, análisis OLAP, cuadros de mando digitales, extracción e integración de datos, administración y seguridad, brindándole a las organizaciones las mejores soluciones para sus necesidades. Entre los beneficios de esta herramienta se destaca la superioridad para crear y visualizar las vistas de análisis y cuadros de mando digitales, presentando una capacidad interactiva muy amigable haciendo muy intuitiva su utilización. Pentaho cuenta con una documentación completa lo que hace que su usabilidad sea superior a cualquier otra herramienta de inteligencia de negocio. La suite de Pentaho permite elegir entre el uso completo de las herramientas o la utilización específica de ciertas herramientas para resolver los problemas del negocio (18).

A continuación se describen las herramientas utilizadas de la suite de Pentaho:

Pentaho BI Server 3.8.0

Funciona como un sistema basado en administración web de informes, el servidor de integración de aplicaciones y un motor de flujo de trabajo ligero. Presenta una interfaz de usuario donde se encuentran disponibles todos los informes, vistas OLAP y cuadros de mando digitales. Cuenta con una consola de administración que permite gestionar y supervisar la aplicación y los usuarios (19).

Ventajas:

- ✓ Aplicación extensible, adaptable y configurable.
- ✓ La gestión de la configuración, tanto de la instalación inicial como del mantenimiento está muy bien resuelta.
- ✓ Se integra con la mayoría de entornos y se puede comunicar con otras aplicaciones vía servicios web.
- ✓ Integra todos los recursos informacionales en una única plataforma de explotación.
- ✓ Proporciona mucha libertad al usuario y los desarrolladores para crear contenidos nuevos.
- ✓ Explotación de recursos como servicios web.

Mondrian OLAP Server 3.2.1

Es un servidor OLAP escrito en Java y gestiona la comunicación entre una aplicación OLAP y la base de datos con los datos fuente. Es decir, actúa como “JDBC para OLAP” ejecutando consultas escritas en lenguaje MDX, leyendo datos en una base de datos relacional y presentando los resultados en formato multidimensional (20).

Sus principales ventajas son:

- ✓ Agilizar la consulta de grandes cantidades de datos.
- ✓ Alta velocidad de respuesta.
- ✓ Permite realizar consultas al mercado de datos.
- ✓ Es un motor ROLAP con caché.

Apache Tomcat 6.0.29

Es una implementación de software de código abierto de Java Servlet y tecnologías Java Server Pages. Funciona en cualquier sistema operativo que disponga de una máquina virtual de Java. Tomcat puede utilizarse como un contenedor solitario o como complemento para un servidor web existente. Es una herramienta gratis, fácil de instalar y se puede ejecutar en máquinas con pocos recursos (21).

Pentaho Schema Workbench 3.2.1

Es un entorno visual que permite el desarrollo y prueba de cubos OLAP. El motor de Mondrian procesa las solicitudes de MDX⁹ con el esquema ROLAP. Estos son archivos del modelo de los esquemas XML de metadatos que se crean en una estructura específica. Los modelos XML se consideran como estructuras en forma de cubos que utilizan hechos y tablas de dimensiones existentes (22).

⁹ MDX es la abreviatura de MultiDimensional eXpressions (Expresiones Multidimensionales).

Ofrece las siguientes funcionalidades:

- ✓ Editor de esquemas integrados con un origen de datos subyacente para su validación.
- ✓ Prueba de consultas MDX contra el esquema y la base de datos.
- ✓ Examina la estructura subyacente de bases de datos.
- ✓ Consulta la base de datos que sirve de origen para el esquema de Mondrian.
- ✓ Publica directamente el esquema en el servidor de Pentaho.

1.6 Herramientas de desarrollo de la solución

Visual Paradigm 8.0

Es una herramienta CASE¹⁰ (Ingeniería de Sistemas Asistida por Computadora) que emplea UML¹¹ como lenguaje de modelado, y ha sido concebida para soportar el ciclo de vida del proceso de desarrollo del software mediante la representación de todo tipo de diagramas. Se diseñó para distintos usuarios interesados en la construcción de sistemas de software de forma fiable, utilizando un enfoque orientado a objetos. Sus principales características son (23):

- ✓ Licencia gratuita y comercial.
- ✓ Capacidades de ingeniería directa e inversa.
- ✓ Disponibilidad de múltiples versiones para cada necesidad.
- ✓ Uso de lenguaje estándar común.

Se decidió utilizar Visual Paradigm 8.0 para UML porque es una herramienta CASE profesional que soporta el ciclo completo de vida del desarrollo de software. Modela todos los tipos de diagramas de clases y permite realizar la ingeniería inversa. Esta herramienta presenta licencia gratuita y comercial, además la universidad cuenta con la licencia para su uso.

PostgreSQL 9.1

Se utiliza como sistema gestor de base de datos PostgreSQL 9.1 por ser un sistema estable y de alto rendimiento. Esta versión posee el motor de base de datos más potente de código abierto, ofreciendo diversas características que permiten retirar obstáculos para el despliegue de nuevas aplicaciones. Una de las mejoras de esta versión es la ejecución de búsquedas más potentes implementando un método de indexado. Entre las principales características que avalan la decisión de tomar PostgreSQL como SGBD¹² figuran las siguientes (24):

¹⁰ Del inglés Computer Aided Software Engineering, por sus siglas CASE.

¹¹ Del inglés Unified Modeling Language, por sus siglas UML.

¹² Sistema Gestor de Base de Datos.

- ✓ Replicación sincrónica para clústeres: permite replicar los datos que son agregados mediante la implementación de clústeres de servidores PostgreSQL, limitando la posibilidad de pérdidas de datos.
- ✓ Regionalización por columna para bases de datos multilingües: es la posibilidad que tienen los usuarios para configurar el lenguaje de los textos por cada columna. Incluye soporte para alrededor de 50 juegos de caracteres internacionales diferentes, incluyendo idiomas arábigos y asiáticos.
- ✓ Tablas sin log para incrementar rendimiento: proporcionan una manera de mejorar el rendimiento y el tiempo de respuesta es mucho mayor, manteniendo los datos manejados dentro de PostgreSQL y reduciendo la carga de entradas y salidas.
- ✓ Brinda confiabilidad e integridad a los datos: el PostgreSQL 9.1 puede ser ejecutado en la mayoría de los sistemas operativos más utilizados en el mundo incluyendo, Linux, varias versiones de UNIX y Windows.
- ✓ Es un producto sin costos de licencia: se convierte en una alternativa extremadamente atractiva para las empresas.
- ✓ Posee numerosas interfaces nativas de lenguajes como: C ++, Java, Net, Perl, Python.
- ✓ Excelente documentación.

PgAdminIII 1.14.1

PgAdminIII es una plataforma de desarrollo de PostgreSQL que se puede utilizar en diferentes sistemas operativos. Está diseñado para responder a las necesidades de los usuarios, tanto para la escritura de consultas SQL¹³ (Lenguaje de Consulta Estructurado) como para la creación de complejas bases de datos. Facilita la administración y soporta todas las características de PostgreSQL en su interfaz gráfica. Dentro de las principales características se encuentran (25):

- ✓ Se puede ejecutar en sistemas operativos como Microsoft Windows Server 2000, Linux, FreeBSD, Mac OSX, Solaris, etc.
- ✓ Cuenta con una interfaz multilingüe porque posee más de 12 traducciones compatibles.
- ✓ Acceso a todos los objetos de PostgreSQL, se muestran en su definición de SQL y una lista fácil de usar de los bienes.
- ✓ Licencia gratuita disponible para todos los usuarios.

¹³ Del inglés Structured Query Language, por sus siglas SQL.

1.7 Indicadores y métricas

Para la explotación adecuada del mercado de datos se hace necesario elaborar estadísticas certeras y precisas que posibiliten hacer uso óptimo de los datos que se manejan. Esto está estrechamente entrelazado con el concepto de indicador. El indicador permite proporcionar información relevante acerca de algún aspecto significativo de la realidad.

A continuación se enuncian algunos conceptos relacionados con el negocio (26):

Indicador: es una magnitud asociada a una característica, que permite a través de su medición en períodos sucesivos y por comparación con el valor estándar establecido, evaluar periódicamente dicha característica y verificar el cumplimiento de los objetivos establecidos para el servicio.

Métrica: es una medida numérica directa que representa un conjunto de datos de negocios en la relación a una o más dimensiones.

Indicadores clave de rendimiento (KPI¹⁴): es un indicador que está vinculado a un objetivo. Generalmente se muestra como una tasa o porcentaje y está diseñado para permitir que un usuario de negocios sepa instantáneamente si están dentro o fuera de su plan sin que tenga que buscar información adicional.

Métricas e indicadores clave de rendimiento: son la base para construir un cuadro de mando de gran despliegue visual. Son las herramientas más eficaces para alertar a los usuarios en cuanto a donde se encuentran parados en relación a los objetivos.

Conclusiones del capítulo

Luego de realizar un estudio de los fundamentos teóricos de la presente investigación se concluye que:

- ✓ Se decidió utilizar la Propuesta de metodología para el desarrollo de Soluciones de Almacenes de Datos e Inteligencia de Negocio en DATEC, para guiar el proceso de desarrollo de la solución.
- ✓ Se decidió utilizar como herramienta CASE el Visual Paradigm en su versión 8.0, que posibilita la generación de los diagramas necesarios para el modelado.
- ✓ Se seleccionó como modelo de almacenamiento de datos ROLAP, permitiendo a los usuarios realizar el análisis multidimensional de los datos a través de este motor.
- ✓ Para el desarrollo de la capa de visualización se decidió utilizar las herramientas Pentaho BI Server en su versión 3.8.0, Mondrian OLAP Server en su versión 3.2.1, Apache Tomcat en su versión 6.0.29 y Pentaho Schema Workbench en su versión 3.2.1.

¹⁴ Del inglés Key Performance Indicator, por sus siglas KPI.

CAPÍTULO 2: Análisis y diseño

Introducción

En el desarrollo del presente capítulo se realizará la descripción de la capa de visualización del mercado de datos Tecnología. Se definirán las reglas del negocio, los requisitos de información, los requisitos funcionales y no funcionales. Se diseñarán las vistas lógicas, el subsistema de visualización y se especificarán las Áreas de Análisis (AA) identificadas para agrupar las vistas de análisis agrupadas en los Libros de Trabajo (LT).

2.1 Arquitectura de la solución

Para enriquecer con nuevos indicadores el mercado de datos Tecnología fue necesario hacer uso de un conjunto de herramientas de inteligencia de negocio, que facilitaron la extracción, depuración, análisis y almacenamiento de los datos recogidos por las observaciones sistemáticas del T-arenal sobre los distintos componentes de hardware de las computadoras de la universidad. Una vez consolidada la información anteriormente dispersa en este mercado, se procede a su explotación a través de las herramientas antes mencionadas, y posteriormente se muestran al usuario mediante los gráficos dinámicos y cuadros de mando digitales, los cuales proporcionan una visión amplia y sencilla del negocio, haciendo más sugerente e intuitivo el proceso de toma de decisiones para los directivos de la universidad.

La imagen que se muestra a continuación refleja la arquitectura de la solución de inteligencia de negocio que se propone (Figura 1).

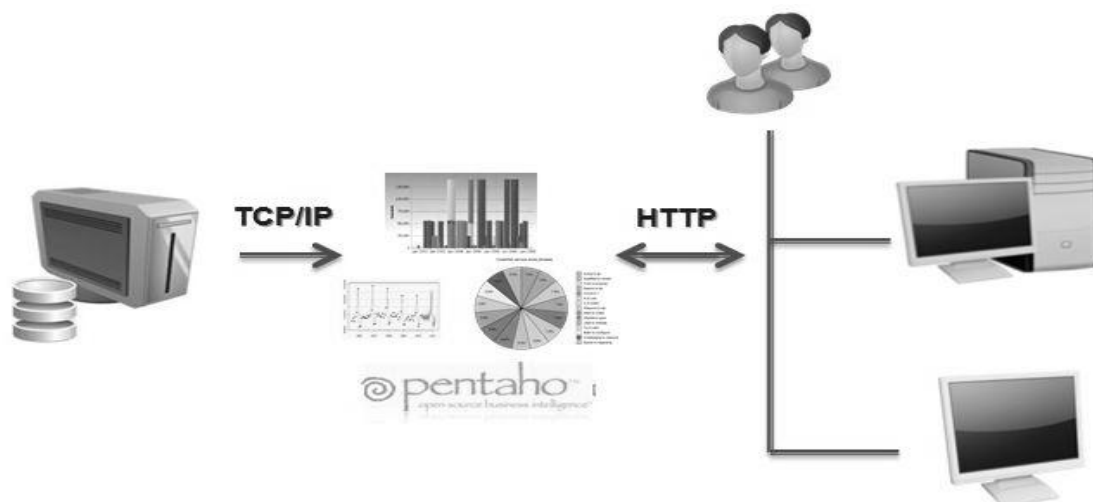


Figura 1: Arquitectura de la solución.

2.2 Especificación de requisitos

2.2.1 Requisitos de información

Los Requisitos de Información (RI) son los encargados de describir la información que se debe almacenar para satisfacer las necesidades del cliente. Durante el proceso de análisis fueron identificados los siguientes requisitos de información:

- RI1. Obtener el total de computadoras reportadas por laboratorio y tiempo.
- RI2. Obtener el tiempo de encendido y tiempo de trabajo de las computadoras por laboratorios y tiempo.
- RI3. Obtener el porcentaje del tiempo de encendido de las computadoras según el tiempo teórico de explotación por laboratorio y tiempo.
- RI4. Obtener el porcentaje del tiempo de encendido de las computadoras según el tiempo real de explotación por laboratorio y tiempo.
- RI5. Obtener el porcentaje del tiempo de trabajo de las computadoras según el tiempo teórico de explotación por laboratorio y tiempo.
- RI6. Obtener el porcentaje del tiempo de trabajo de las computadoras según el tiempo real de explotación por laboratorio y tiempo.
- RI7. Obtener el porcentaje de aprovechamiento de la explotación de las computadoras por laboratorio y tiempo.
- RI8. Obtener el porcentaje de inactividad del CPU por laboratorio, momentos del día y tiempo.
- RI9. Obtener el porcentaje de trabajo de las computadoras por usuario, laboratorio y fecha
- RI10. Obtener el promedio de las observaciones por laboratorio y fecha.
- RI11. Obtener la cantidad total de espacio en disco duro por laboratorio y tiempo.
- RI12. Obtener la cantidad total de espacio en disco duro usado por laboratorio y tiempo.
- RI13. Obtener la cantidad total de espacio en disco duro libre por laboratorio y tiempo.
- RI14. Obtener el promedio de la memoria usada por momentos del día, tipo de memoria, laboratorio y tiempo.
- RI15. Obtener el promedio de paquetes enviados y recibidos en la red por laboratorio, momentos del día y tiempo.
- RI16. Obtener el porcentaje de paquetes enviados y recibidos en la red por laboratorio, momentos del día y tiempo.
- RI17. Obtener la cantidad total del tráfico de paquetes enviados y recibidos en la red por laboratorio, momentos del día y tiempo.

- RI18. Obtener el promedio de memoria RAM utilizada por los procesos en ejecución y categoría de proceso.
- RI19. Obtener el porcentaje de ejecución de los procesos por categoría de proceso.
- RI20. Obtener la cantidad total de computadoras por tipo de CPU, laboratorio y tiempo.
- RI21. Obtener la cantidad total de computadoras por tipo de memoria, laboratorio y tiempo.
- RI22. Obtener la cantidad total de computadoras por tipo de sistema operativo, laboratorio y tiempo.
- RI23. Obtener la cantidad total de computadoras en el dominio por laboratorio y tiempo.

2.2.2 Requisitos funcionales

Los Requisitos Funcionales (RF) son las capacidades o condiciones que el sistema debe cumplir. Es necesaria la identificación de los mismos para satisfacer las necesidades del cliente. A continuación se enumeran las funcionalidades que debe poseer el sistema.

- RF1. Validar usuario y contraseña.
- RF2. Adicionar rol.
- RF3. Eliminar rol.
- RF4. Modificar rol.
- RF5. Adicionar usuario.
- RF6. Eliminar usuario.
- RF7. Modificar usuario.
- RF8. Adicionar reporte.
- RF9. Eliminar reporte
- RF10. Modificar reporte.
- RF11. Personalizar reporte.
- RF12. Exportar reporte a pdf.
- RF13. Exportar reporte a Excel.
- RF14. Imprimir reporte.

2.2.3 Requisitos no funcionales

Los Requisitos No Funcionales (RNF) son propiedades o cualidades que el producto debe tener. Se debe pensar en estas propiedades como las características que hacen al producto atractivo, usable, rápido o confiable. Son importantes para que clientes y usuarios puedan valorar las características no funcionales del producto, pues si se conoce que el mismo cumple con toda la funcionalidad requerida, pueden marcar la diferencia entre un producto bien aceptado y uno con poca aceptación (29).

Usabilidad

RNF1. Agilizar el acceso a las vistas de análisis del mercado de datos mediante la distribución de la información por áreas de análisis.

Fiabilidad

RNF2. El sistema debe estar disponible durante el horario de trabajo.

RNF3. El mantenimiento del sistema se realizará una vez por mes.

Interfaces de usuario

RNF4. Mostrar los mensajes, títulos y demás textos que aparezcan en la interfaz del sistema en idioma español.

RNF5. El sistema debe tener una interfaz amigable y sencilla de utilizar para el usuario.

RNF6. Los reportes se mostrarán de manera sencilla, facilitando la interacción usuario-aplicación.

Interfaces de hardware

RNF7. Para lograr una explotación aceptable del sistema los servidores deben contar con los siguientes requerimientos de hardware:

- ✓ Para las máquinas clientes: mínimo 512 RAM, recomendado 1 GB en adelante.
- ✓ Para los servidores: mínimo 1 GB de RAM, recomendado 2 GB en adelante y 20 GB de disco duro en adelante.

Interfaces de software

RNF8. Las configuraciones de software de las máquinas clientes deben contar al menos con:

- ✓ Firefox 2.0 y Adobe Flash Player.
- ✓ Para el uso de las herramientas de la suite de Pentaho se requiere la instalación de la máquina virtual de java versión 1.6.

Licencia

RNF9. Las herramientas utilizadas deben estar bajo licencia GPL, excepto el Visual Paradigm pues la universidad cuenta con la licencia.

2.3 Reglas del negocio

Las Reglas de Negocio (RN) definen la estructura, funcionamiento y la estrategia de una organización mediante políticas, medidas y restricciones para cumplir con los objetivos. El proceso de especificación implica que hay que identificarlas dentro del negocio y aplicarlas a la solución propuesta. Pueden estar formalmente definidas en manuales, contratos o acuerdos, o bien pueden existir como conocimiento o

experiencia que tienen los miembros de la institución. Durante el análisis se definieron las siguientes RN:

RN. 1 El tiempo de explotación teórico de una computadora se calcula de la siguiente manera: momentos de hora por el número de computadoras del laboratorio.

RN. 2 El tiempo de explotación real de una computadora se calcula de la siguiente manera: momentos de hora por la cantidad de computadoras reportadas.

RN. 3 El plan de encendido de una computadora se calcula de la siguiente manera: tiempo de encendido entre el tiempo de explotación teórica.

RN. 4 El real encendido de una computadora se calcula de la siguiente manera: tiempo de encendido entre el tiempo de explotación real.

RN. 5 El plan de trabajo de una computadora se calcula de la siguiente manera: tiempo de trabajo entre el tiempo de explotación teórica.

RN. 6 El trabajo real de una computadora se calcula de la siguiente manera: tiempo de trabajo entre el tiempo de explotación real.

RN. 7 El porcentaje de aprovechamiento de una computadora se calcula de la siguiente manera: tiempo de trabajo entre el tiempo de encendido por 100.

RN. 8 El consumo de memoria RAM por procesos se calcula de la siguiente manera: uso de memoria total entre el número de procesos.

RN. 9 El promedio de consumo de memoria RAM se calcula de la siguiente manera: consumo de RAM entre la cantidad de observaciones.

RN. 10 La cantidad de Megabytes enviados se calcula de la siguiente manera: bytes enviados por la red entre la cantidad de observaciones.

RN. 11 La cantidad de Megabytes recibidos se calcula de la siguiente manera: bytes recibidos por la red entre la cantidad de observaciones.

RN. 12 El porcentaje de RAM se calcula de la siguiente manera: consumo de RAM entre el consumo total RAM por 100.

RN. 13 El porcentaje de procesos por categorías se calcula de la siguiente manera: cantidad de procesos de una categoría entre la cantidad total de procesos por 100.

2.4 Casos de usos del sistema

Los Casos de Uso (CU) son la interacción del usuario con el sistema para obtener un objetivo específico, se utilizan para representar los requisitos del sistema a desarrollar a través de una serie de iteraciones secuenciales, que tienen lugar cuando el sistema reacciona a un evento efectuado por el

usuario. Durante la fase de análisis los requisitos de información y funcionales identificados se agruparon en tres casos de uso de información y cinco casos de uso funcionales.

2.4.1 Actores del sistema

Tabla 1: Actores del sistema

Actores	Descripción
Analista	Analiza y consulta la información.
Administrador	Administra los usuarios, permisos y reportes

2.4.2 Casos de usos de información

Los requisitos de información fueron agrupados por temas de análisis, en dependencia de las necesidades del cliente, a continuación se describen los Casos de Uso de Información (CUI).

CUI. 1 Obtener información de explotación: visualiza la información relacionada con la explotación de las computadoras.

CUI. 2 Obtener información de procesos: visualiza la información relacionada con la ejecución de procesos.

CUI. 3 Obtener información de inventario: visualiza la información relacionada con el inventario de las computadoras.

A continuación se detalla el CUI **Obtener información de explotación**. Las demás descripciones de los CUI se encuentran en los **Anexos**.

Tabla 2: Descripción textual del CUI Obtener información de explotación.

Objetivo	Obtener la información relacionada con la explotación de las computadoras.
Actores	Analista
Resumen	El caso de uso inicia cuando el analista desea consultar la información relacionada con la explotación de las computadoras en cada laboratorio. Luego selecciona el reporte deseado, el sistema muestra la información contenida en él y las opciones de los posibles cambios que le puede hacer al reporte. El caso de uso finaliza cuando el especialista termina de analizar la información relacionada con la explotación de las computadoras.
Complejidad	Alta
Prioridad	Media

Precondiciones	La información relacionada con la explotación de las computadoras se cargó correctamente en el mercado de datos. Todos los reportes relacionados con la explotación de las computadoras se encuentran disponibles.	
Postcondiciones	Los reportes correspondientes al libro de trabajo L.T Explotación han sido consultados por el especialista.	
Flujo de eventos		
Flujo básico Obtener información de explotación		
Actor	Sistema	
1. El especialista se autentica en el sistema.	2. Muestra la interfaz principal con las áreas de análisis existentes.	
3. El especialista selecciona el área de análisis general A.A.G Sala Situacional UCI.	4. Muestra las áreas de análisis que están contenidas dentro del A.A.G Sala Situacional UCI.	
5. El especialista selecciona el área de análisis A.A Tecnología.	6. Muestra los libros de trabajo que están contenidos dentro del A.A Tecnología.	
7. El especialista selecciona el libro de trabajo L.T Explotación.	8. Muestra los reportes contenidos dentro del L.T Explotación.	
9. El especialista selecciona el reporte deseado.	10. Muestra la información contenida en el reporte seleccionado y brinda la posibilidad al especialista de hacerle cambios al reporte para su análisis. Ir al Caso de Uso: Visualizar reporte. Finaliza el caso de uso.	
Opciones de reportes de explotación		
Perspectivas de análisis	Posibles resultados	
	Medidas	Periodicidad
Variables de entrada relacionadas con el CU Mostrar información de Explotación. Laboratorio Marca de tiempo Tiempo RAM	Variables de salida disponibles en el hecho Explotación. Tiempo de encendido Tiempo de trabajo Inactividad del CPU Cantidad de observaciones Espacio libre Espacio utilizado Espacio total	Rango de tiempo en que se solicitan las variables de salida: Diario

	Consumo de RAM Computadoras encendidas Observaciones Bytes enviados Bytes recibidos	
--	---	--

2.4.3 Casos de usos funcionales

Los requisitos funcionales fueron agrupados en cinco Casos de Uso Funcionales (CUF), los cuales se describen a continuación.

CUF. 1 Autenticar usuario: se realiza la autenticación de los usuarios en el sistema.

CUF. 2 Administrar usuarios: se insertan, modifican o eliminan a los usuarios del sistema.

CUF. 3 Administrar roles: se insertan, modifican o eliminan los roles en el sistema.

CUF. 4 Administrar reporte: se insertan, modifican o eliminan los reportes en el sistema.

CUF. 5 Realizar operaciones sobre reporte: se realiza la visualización de los reportes o funciones deseadas.

A continuación en la tabla se describe el CUF **Administrar roles**. Las demás descripciones de los CUF se pueden encontrar en el artefacto Especificación de casos de uso del expediente de proyecto Tecnología.

Tabla 3: Descripción textual del CUF Administrar roles.

Objetivo	Administrar roles.	
Actores	Administrador.	
Resumen	El caso de uso se inicia cuando el administrador desea insertar, modificar o eliminar algún rol. Luego de seleccionar la opción que desea realizar, el sistema muestra la interfaz correspondiente y el usuario realiza los cambios deseados. El caso de uso finaliza cuando queda creado o eliminado un rol.	
Complejidad	Baja	
Prioridad	Baja	
Precondiciones	El administrador se autenticó en el sistema.	
Postcondiciones	El rol ha sido adicionado, modificado o eliminado.	
Flujo normal de eventos		
Actor	Sistema	
1. El administrador se autentica en la consola de administración.	2. Muestra la consola principal de la consola de administración.	

3. El administrador selecciona la opción Administración del menú Inicio.	4. Muestra la vista para la administración de los roles y usuarios
5. El administrador selecciona la opción Roles.	6. Muestra todos los roles que hay insertados y las opciones de adicionar y eliminar roles.
7. El administrador selecciona la opción que desea realizar: <ul style="list-style-type: none"> ✓ Adicionar rol. Ir a la Sección: Adicionar rol. ✓ Eliminar rol. Ir a la Sección: Eliminar rol. 	
Sección "Insertar rol"	
Actor	Sistema
1. El administrador selecciona la opción Adicionar rol.	2. Muestra una ventana con los datos necesarios para adicionar un rol.
3. El administrador introduce los datos correspondientes y da clic en el botón Aceptar.	4. Guarda los cambios en la base de datos y muestra el rol adicionado. Finaliza el caso de uso.
Sección "Eliminar rol"	
Actor	Sistema
1. El administrador selecciona el rol que desea eliminar y da clic en la opción Eliminar Rol.	2. Muestra una ventana preguntándole si desea eliminar el rol seleccionado.
3. El administrador da clic en el botón Aceptar.	4. Elimina el rol seleccionado. Finaliza el caso de uso.
Sección "Modificar rol"	
Actor	Sistema

1. Selecciona el rol que desea modificar.	2. Muestra la ventana de los datos del rol que se desea modificar
3. Modifica los datos deseados.	4. Guarda los datos en la base de datos y muestra el rol modificado. Finaliza el caso de uso.

2.5 Diagrama de casos de uso

El diagrama de CU muestra la relación entre los actores y los Casos de Uso del Sistema (CUS). Especifican la funcionalidad y el comportamiento de este mediante su interacción con los usuarios u otros sistemas. En la siguiente figura se muestra el diagrama de CUS.

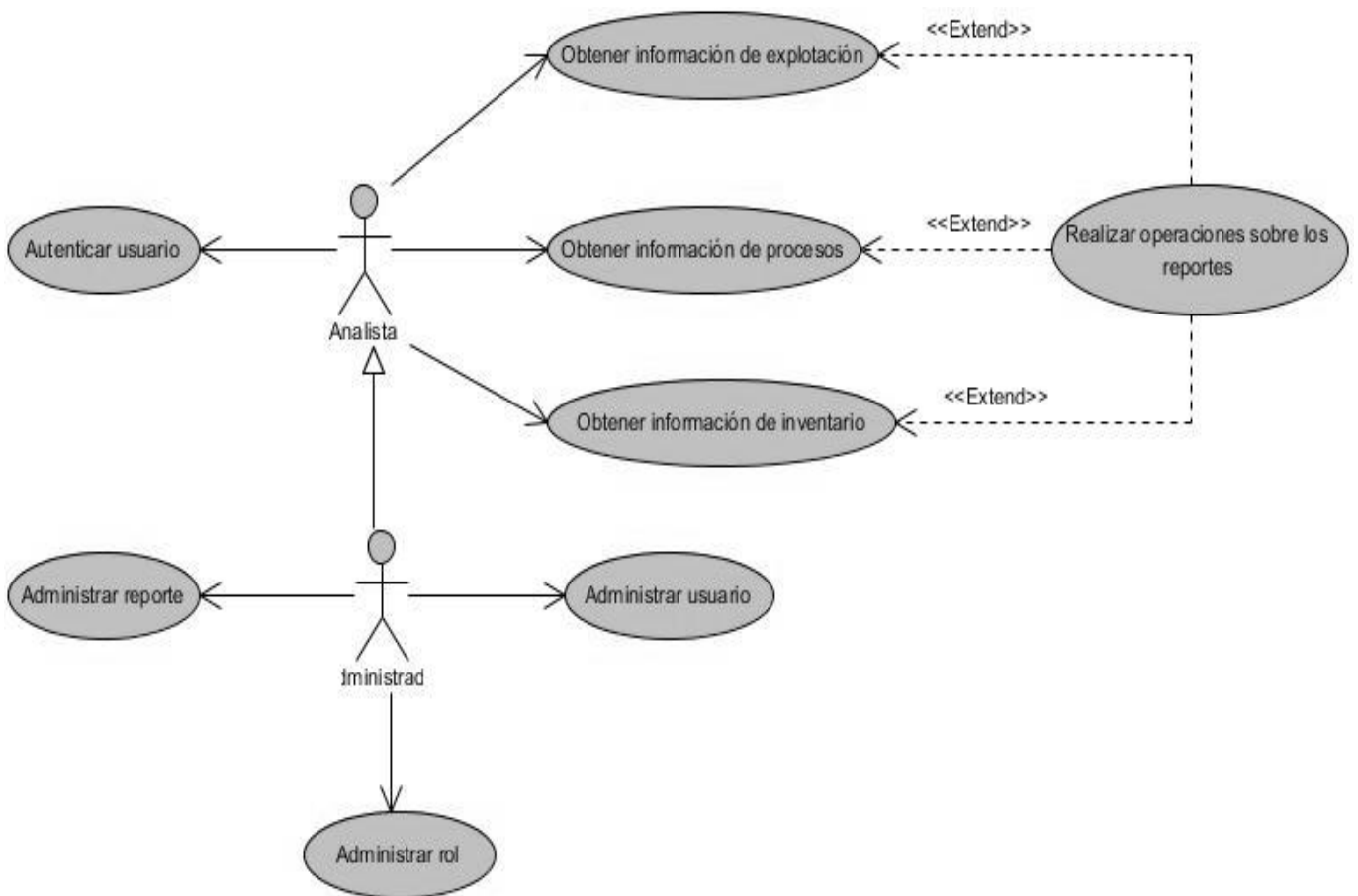


Figura 2: Diagrama de CUS.

2.6 Reglas de navegación

A continuación se describen los elementos que componen las reglas de navegación presentadas en la capa de visualización del mercado de datos Tecnología. Las dimensiones poseen entre sus características principales la definición de jerarquías entre sus atributos para lograr realizar un análisis más detallado de los datos.

dim_temporal_dia: representa una forma de análisis atendiendo los datos por períodos de fechas, es una de las más utilizadas e importantes, debido a que organiza la información de acuerdo a un momento determinado.

dim_so: representa una forma de análisis atendiendo los datos de los sistemas operativos y a sus respectivas versiones.

dim_cpu: representa una forma de análisis atendiendo los datos del tipo de CPU.

dim_usuario: representa una forma de análisis atendiendo los datos del usuario.

dim_marca_tiempo_dia: representa una forma de análisis atendiendo los datos por momentos del día.

dim_proceso: representa una forma de análisis atendiendo los datos por procesos.

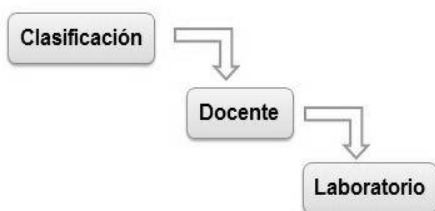
dim_ram: representa una forma de análisis atendiendo los datos de la memoria RAM.

dim_estructura: representa una forma de análisis atendiendo los datos por clasificación, docentes y laboratorios.

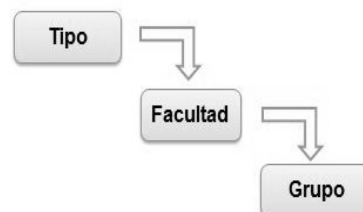
dim_pc: representa una forma de análisis atendiendo los datos de las computadoras.

2.6.1 Agrupaciones de información y jerarquías

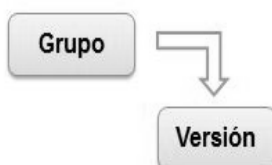
dim_usuario



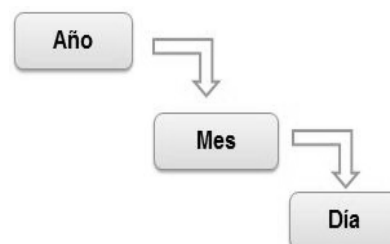
dim_estructura



dim_temporal_dia



dim_so



dim_cpu

Modelo

dim_marca_tiempo

Momento

dim_pc

PC

dim_ram

RAM

dim_proceso

Grupo

2.7 Elementos de la regla de navegación

dim_proceso:

- ✓ Grupo: algunas solicitudes de información se representan a nivel de grupos de procesos para facilitar un correcto análisis de los datos.

dim_temporal_dia:

- ✓ Año: nivel superior de la dimensión temporal_dia, algunas solicitudes de información requieren ser representadas por año.
- ✓ Mes: nivel inferior de año, algunas solicitudes de la información requieren ser representadas por mes.
- ✓ Día: nivel inferior de mes, algunas solicitudes de la información requieren ser representadas por día.

dim_marca_tiempo:

- ✓ Momento: algunas solicitudes de información requieren ser representadas por marcas de tiempo para una mejor organización de la información.

dim_estructura:

- ✓ Clasificación: nivel superior de la dimensión dim_estructura, algunas solicitudes de información requieren ser representadas por diferentes clasificaciones.
- ✓ Docente: nivel inmediato inferior de clasificación, algunas solicitudes de la información requieren ser representadas por docentes.
- ✓ Laboratorio: nivel inmediato inferior de docente, algunas solicitudes de la información requieren ser representadas por laboratorios.

dim_cpu:

- ✓ Modelo: algunas solicitudes de información requieren ser representadas por los modelos de los CPU.

dim_so:

- ✓ Grupo: nivel superior de la dimensión dim_so, algunas solicitudes de información requieren ser representadas por diferentes grupos de sistemas operativos.
- ✓ Versión: nivel inmediato inferior de grupo, algunos reportes de información se muestran por la versión de los sistemas operativos.

dim_ram:

- ✓ RAM: algunas solicitudes de información requieren ser representadas por la memoria RAM.

dim_pc:

- ✓ PC: algunas solicitudes de información requieren ser representadas por las computadoras.

dim_usuario:

- ✓ Tipo: nivel superior de la dimensión dim_usuario, algunas solicitudes de información requieren ser representadas por diferentes usuarios.
- ✓ Facultad: nivel inmediato inferior de tipo, algunos reportes de información se muestran por la facultad.
- ✓ Grupo: nivel inmediato inferior de facultad, algunos reportes de información se muestran por el grupo.

2.8 Diseño de las vistas lógicas

El almacén de datos operacional Tecnología, recoge los datos diarios recopilados por el T-Arenal y permite que su orientación sea totalmente analítica. El sistema cuenta con un total de cinco tablas, que se describen a continuación:

metricas_dinamicas: recoge los datos que varían frecuentemente, como: consumo de ram, uso de cpu, tiempo trabajo, espacio libre en disco, espacio usado en disco, bytes enviados, bytes recibidos y las observaciones.

metricas_estaticas: recoge los datos estáticos del sistema, como: memoria total, cpu total y el espacio en disco total.

pc: recoge los datos relacionados con las computadoras, como: identificador de computadora y el número de ip.

pc_x_procesos: recoge los datos relacionados con los procesos por computadoras, como: frecuencia de proceso, memoria de ejecución por procesos y el nombre del proceso.

usuario_x_procesos: recoge los datos relacionados con los usuarios por procesos, como: nombre de proceso, nombre de usuario, uso de memoria y el uso de cpu.

El proceso de creación de consultas con el objetivo de acceder a un indicador por determinados parámetros se convierte en un proceso engorroso. Para mitigar esta situación los gestores de base de

datos proporcionan métodos para guardar la sentencia SELECT como una vista lógica, la cual se utiliza como una tabla virtual donde se puede consultar e incluirla en otras tablas. A continuación se detallan las cinco vistas lógicas creadas que cumplen satisfactoriamente las necesidades del cliente.

v_metricas_estaticas: contiene las dimensiones relacionadas con la fecha, laboratorios, cpu, ram y sistema operativo. Las medidas cantidad de computadoras y cantidad de computadoras en el dominio.

v_metricas_disco_duro: contiene las dimensiones relacionadas con la fecha, laboratorios, las medidas espacio utilizado, espacio libre del disco duro y el número de computadora.

v_metricas_ram_red: contiene las dimensiones relacionadas con la fecha, laboratorios, momentos del día y ram. Las medidas bytes enviados y recibidos por la red, cantidad de observaciones, cantidad de computadoras y el consumo de la memoria ram.

v_metricas_utilizacion: contiene las dimensiones relacionadas con la fecha, laboratorios y momentos del día. Las medidas cantidad de computadoras, tiempo de encendido y tiempo de trabajo.

v_procesos_usuarios: contiene las dimensiones relacionadas con la fecha, laboratorios, procesos y usuarios. Las medidas número de procesos, uso de la memoria y tiempo de trabajo.

2.9 Diseño del subsistema de visualización

En el diseño del subsistema de visualización es necesario crear los cubos OLAP. Estos proveen un mecanismo para buscar datos con rapidez, con el objetivo de agruparlos y proporcionar una vista más detallada, además, el tiempo de respuesta es uniforme, independientemente de la cantidad de datos. En los cubos se definen las dimensiones, las medidas y los niveles jerárquicos de cada dimensión.

2.9.1. Diseño de los cubos OLAP

La solución cuenta con cinco cubos, nueve dimensiones y 19 medidas, distribuidos de la siguiente manera:

El cubo **Métricas estáticas** contiene cinco dimensiones y una medida física.

El cubo **Procesos por usuarios** contiene cuatro dimensiones y cinco medidas físicas.

El cubo **Métricas utilización** contiene tres dimensiones y seis medidas físicas.

El cubo **Métricas disco duro** contiene tres dimensiones y tres medidas físicas.

El cubo **Métricas ram red** contiene cuatro dimensiones y cuatro medidas físicas.

2.9.2. Diseño de la arquitectura de información

Un área de análisis permite agrupar la información referente a los indicadores organizados en los LT según las necesidades de una empresa. La definición de estas constituye un punto importante en la solución propuesta, donde se definieron en función de las necesidades de información del cliente.

Estos datos describen una serie de comportamientos del hardware y software de las computadoras de la universidad.

A continuación se muestra una representación del diseño establecido para la arquitectura de la información del sistema (Figura 3). La misma está compuesta por el Área de Análisis General Sala Situacional de la UCI (AAG.SSUCI) que contiene el Área de Análisis (AA) Tecnología, agrupando la información referente a los indicadores en libros de trabajo.

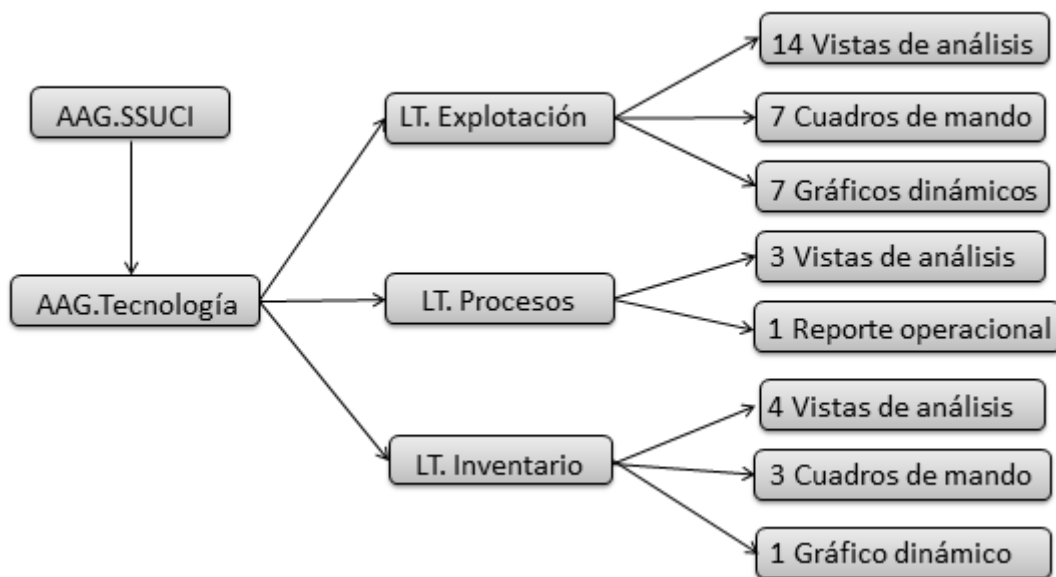


Figura 3: Arquitectura de la información.

2.9.3. Descripción de los libros de trabajo

Los LT son estructuras organizativas que agrupan las vistas de análisis y son generadas dentro de las AA. Pueden ser creados teniendo en cuenta un criterio que permita organizar la información. En la solución propuesta se definieron tres LT, Explotación, Procesos e Inventario, los que se describen a continuación:

LT1. Explotación: contiene 14 vistas de análisis, siete gráficos dinámicos y siete cuadros de mando digitales, permitiendo el análisis correspondiente a la explotación de las computadoras y la visualización de estas vistas de análisis.

LT2. Inventario: contiene cuatro vistas de análisis, un gráfico dinámico y tres cuadros de mandos digitales, permitiendo realizar el análisis correspondiente del hardware y software de las computadoras.

LT3. Procesos: contiene tres vistas de análisis, que permiten realizar el análisis correspondiente a la ejecución de los procesos según las categorías y usuarios, además cuenta con un reporte operacional con el objetivo de llegar a un mayor nivel de especificación de la información.

2.10 Esquema de seguridad

Los sistemas informáticos deben estar provistos de mecanismos que aseguren la integridad de los datos y el sistema en general. Usualmente se establecen roles y permisos que se le aplican a los distintos usuarios que se definen en una aplicación. La Plataforma Pentaho BI contiene su propia seguridad, compuesta por cuatro áreas esenciales (27):

Seguridad de acceso a datos de objetos: incluye usuarios, contraseñas, autorizaciones permitidas, recursos web y protección a datos.

Autenticación: incluye el procesamiento de información interactiva de inicio de sesión (por ejemplo: nombre de usuario y contraseña), comparándola con la información recuperada del almacén de datos de seguridad.

Autorización de recursos web (URL): ofrece protección a las URL para responder a cada usuario si pueden o no acceder a una determinada página. El administrador de recursos web es el encargado de brindar a cada usuario autenticado un permiso de seguridad, delimitando las páginas a las que tiene acceso y a las que no.

Autorización a objetos del dominio: en el sistema los únicos objetos del dominio protegidos por la plataforma, son los objetos del repositorio otorgados al usuario autenticado.

Tabla 4: Seguridad en la aplicación

Roles	Permisos		Descripción
	Lectura	Escritura	
Analista	X		Tiene acceso de solo lectura al AA. Tecnología. Puede visualizar todas las vistas de análisis de esta área.
Administrador	X	X	Tiene acceso de lectura y escritura para gestionar los permisos, roles y usuarios, pero solo tiene permiso de lectura en el AA. Tecnología.

Conclusiones del capítulo

Finalizado el análisis y diseño de la capa de visualización del mercado de datos Tecnología, se concluye que:

- ✓ Se identificaron 23 requisitos de información, 14 requisitos funcionales y nueve requisitos no funcionales del sistema, agrupados en tres CUI y cinco CUF, permitiendo identificar las necesidades de información del cliente agrupadas en temas de análisis.

- ✓ Se identificaron 13 reglas del negocio, que se corresponden con las restricciones de implementación de la solución propuesta.
- ✓ Durante el diseño se definieron cinco vistas lógicas, que permitieron diseñar cinco cubos multidimensionales.
- ✓ Se realizó el diseño de la arquitectura de información, donde se identificó en el AA Tecnología tres libros de trabajos que contienen 21 vistas de análisis, ocho gráficos dinámicos, 10 cuadros de mandos digitales y un reporte operacional.

CAPÍTULO 3: Implementación y prueba

Introducción

En este capítulo se describen los procesos asociados a la fase de implementación y prueba de la capa de visualización del mercado de datos Tecnología. Se abordarán aspectos referentes al transcurso de la solución para lograr su certificación. Se especificarán detalles acerca de la creación de las vistas de análisis, así como, los gráficos dinámicos y cuadros de mando digitales. Se realizarán las pruebas pertinentes como parte de una actividad más para el aseguramiento de la calidad del producto, comprobando el correcto funcionamiento de la aplicación.

3.1 Implementación de los cubos OLAP

Las aplicaciones OLAP están orientadas a bases de datos multidimensionales que permiten procesar grandes volúmenes de información, en campos bien definidos, y con un acceso inmediato a los datos para su consulta y posterior análisis (28). Los cubos OLAP representan una tecnología que mejora significativamente el proceso de análisis de la información. Para la implementación de las vistas de análisis de la capa de visualización se crearon cinco vistas lógicas, a partir de las cuales se modelaron los cubos multidimensionales en la implantación del subsistema de visualización. Cada cubo comprende sus respectivas medidas, a las que se le definieron el tipo de datos, la vista origen y la visibilidad, se definieron también las dimensiones asociadas al cubo, a las que se les especificaron el nombre que se mostrará en cada vista de análisis y su correspondiente jerarquía.

A continuación se muestran los cinco cubos que componen la aplicación (Figura 4).

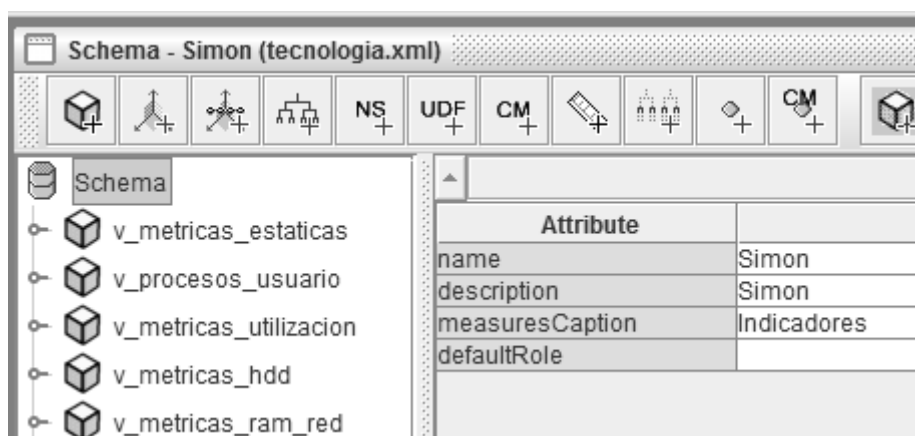


Figura 4: Representación de los cinco cubos del mercado de datos Tecnología.

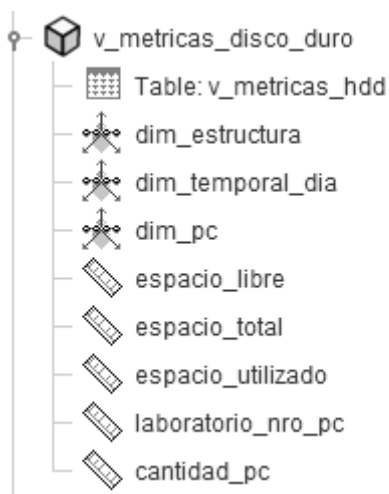


Figura 5: Representación de los elementos que componen el cubo métricas de disco duro. Se visualiza las dimensiones y medidas asociadas.

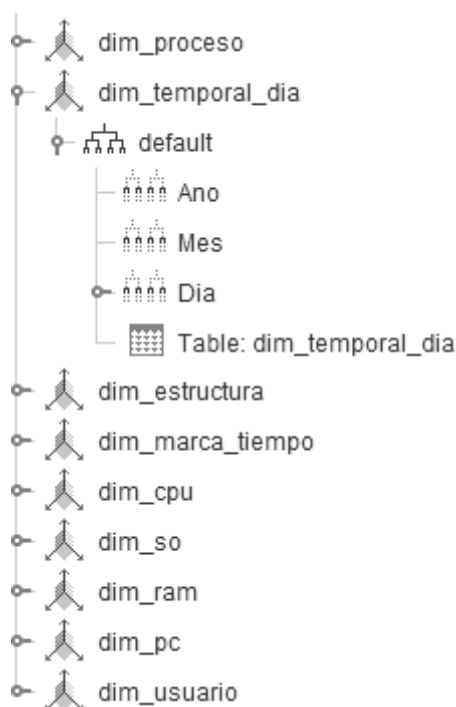


Figura 6: Representación de las dimensiones y la jerarquía que la compone.

3.2 Arquitectura de información

La capa de visualización hace uso de los cuadros de mando digitales para facilitar el proceso de toma de decisiones referente al área de Tecnología. Gracias a la capacidad de integrar datos en una sola gráfica y al despliegue visual de estos tableros, los directivos de la universidad podrán identificar las tendencias y correlaciones de la información. A continuación se detallan los elementos que componen

la estructura de navegación de la información (Figura 7). El desarrollo de la capa de visualización contiene un AA, conformada por tres LT que contienen 21 vistas de análisis, con sus respectivos cuadros de mando digitales y gráficos dinámicos.

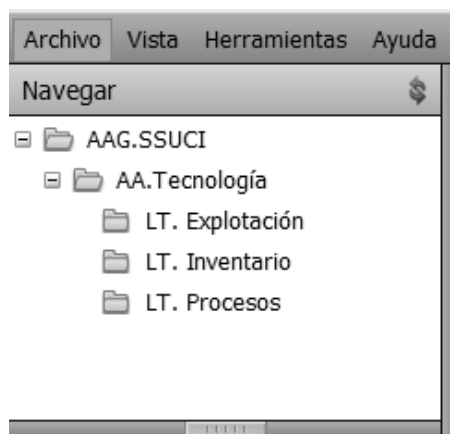


Figura 7: Arquitectura de información: el AA contiene los LT, en los que se ubican las vistas de análisis, cuadros de mandos digitales y gráficos dinámicos.

3.3 Implementación de las vistas de análisis

Las vistas de análisis son la unidad básica utilizada para mostrar la información almacenada en el almacén de datos operacional, y representan la información que el cliente desea que se muestre como finalidad del producto. Estos son implementados a través de consultas MDX, que son en los sistemas OLAP el equivalente a las consultas SQL en las bases de datos relacionales.

LT Explotación: contiene 14 vistas de análisis, siete cuadros de mandos digitales y siete gráficos dinámicos, permitiendo realizar un análisis de los datos en función de los indicadores que a continuación se muestran:

- ✓ Aprovechamiento.
- ✓ Cantidad de observaciones.
- ✓ Cantidad de computadoras reportadas.
- ✓ Total de tiempo encendido y tiempo de trabajo.
- ✓ Disco duro.
- ✓ Inactividad del CPU.
- ✓ Por ciento de trabajo según el tiempo real de explotación.
- ✓ Por ciento de trabajo según el tiempo teórico de explotación.
- ✓ Por ciento de encendido según el tiempo real de explotación.
- ✓ Por ciento de encendido según el tiempo teórico de explotación.
- ✓ Por ciento de paquetes enviados y recibidos por la red.

- ✓ Promedio de paquetes enviados y recibidos por la red.
- ✓ Cantidad total del tráfico de red enviado y recibido por la red.
- ✓ Uso de la memoria RAM.

A continuación se muestra la vista de análisis asociada al indicador espacio en disco duro (Figura 8). Las vistas restantes se encuentran en los Anexos.

		Laboratorio	
Fecha	Indicadores	+ Docencia	+ Proyecto
+ Enero	Espacio libre (gb)	4.233	702.095
	Espacio usado (gb)	1.883	1.351.045
	Espacio Total	6.116	2.053.140
+ Febrero	Espacio libre (gb)		903.130
	Espacio usado (gb)		1.357.220
	Espacio Total		2.260.350

Figura 8: Vista de análisis correspondiente al espacio libre y espacio usado del disco duro.

LT Proceso: contiene tres vistas de análisis y un reporte operacional, permitiendo realizar un análisis de los datos en función de los indicadores que a continuación se muestran:

- ✓ Porcentaje de ejecución de los procesos por categorías.
- ✓ Ejecución de los procesos por usuarios.
- ✓ Promedio de memoria RAM utilizada por los procesos.

A continuación se muestra la vista de análisis asociada al indicador ejecución de procesos por categorías (Figura 9):

		Laboratorio			
		+ Docencia		+ Proyecto	
		Indicadores		Indicadores	
Fecha	Proceso	● Procesos	● Porcentaje	● Procesos	● Porcentaje
+ Enero	Chat	1	12,5 %	475	20,6 %
	Desarrollo	2	25,0 %	675	29,3 %
	Docencia			25	1,1 %
	Internet	2	25,0 %	409	17,7 %
	Juego			68	2,9 %
	Multimedia	2	25,0 %	432	18,7 %
	Oficina	1	12,5 %	223	9,7 %

Figura 9: Vista de análisis correspondiente a la ejecución de procesos por categorías.

LT Inventario: contiene cuatro vistas de análisis, tres cuadros de mandos digitales y un gráfico dinámico, permitiendo realizar un análisis de los datos en función de los indicadores que a continuación se muestran:

- ✓ Cantidad de computadoras por tipo de CPU.
- ✓ Cantidad de computadoras por tipo de RAM.
- ✓ Cantidad de computadoras por sistemas operativos.
- ✓ Cantidad de computadoras en el dominio.

A continuación se muestra la vista de análisis asociada al indicador cantidad de computadoras por tipo de memoria RAM (Figura 10):

		Fecha	
RAM	Laboratorio	+ Enero	+ Febrero
512 MB	- Proyecto	11	58
	+ Docente 4	11	58
1 GB	- Docencia	5	
	+ Docente 1	3	
	+ Docente 2	2	
	+ Proyecto	667	526
2 GB	- Proyecto	125	106
	+ Docente 4	125	106

Slicer: [Medida=Cantidad de PC]

Figura 10: Vista de análisis correspondiente a la cantidad de computadoras por tipo de RAM.

Los gráficos dinámicos que componen la capa de visualización graficarán interactivamente los indicadores previamente definidos durante la fase de análisis y diseño de la aplicación. Contendrán filtros ubicados en la parte superior que refinarán la perspectiva de búsqueda del usuario final. Por otra parte, los cuadros de mando digitalizarán la información en distintas gráficas, permitiendo la navegación en la misma ventana de forma que se puede profundizar más en la información mostrada. Esto permite al usuario tener una visión más general e inmediata de los principales indicadores a medir, facilitando tomar decisiones certeras.

3.4.1 Tipos de gráficos

En el desarrollo de la solución se utilizarán los siguientes tipos de gráficos:

- ✓ Gráficos de barras: los gráficos de barras ilustran comparaciones entre elementos individuales.

- ✓ Gráficos de pastel: los gráficos circulares muestran el tamaño proporcional de los elementos que conforman un dato en función de la suma de los elementos. Siempre mostrará una única serie de datos, y es útil cuando se desea destacar un elemento significativo.
- ✓ Gráficos de líneas: los gráficos de líneas muestran una serie como un conjunto de puntos conectados mediante una sola línea. Los mismos se usan para representar los datos que tienen lugar durante un período continuado de tiempo.

A continuación se describen varios ejemplos de la solución propuesta:

Cuadro de mando digital **Aprovechamiento**: muestra en cuatro gráficos de barras el aprovechamiento de las computadoras en la última semana, por docente, tipo de laboratorio y laboratorio respectivamente. Luego en un gráfico de pastel muestra el aprovechamiento del laboratorio seleccionado por momentos del día. En caso de no existir datos muestra un mensaje “No hay datos disponibles”.

Cuadro de mando digital **Consumo promedio de RAM**: muestra en tres gráficos de barras el consumo de RAM en la última semana, por docente y tipo de laboratorio respectivamente. Luego en un gráfico de pastel muestra el consumo de RAM por el tipo de laboratorio, marca de tiempo, tipo de memoria RAM y el consumo por procesos. En caso de no existir datos muestra un mensaje “No hay datos disponibles”.

Las figuras 11 y 12 muestran los ejemplos descritos.

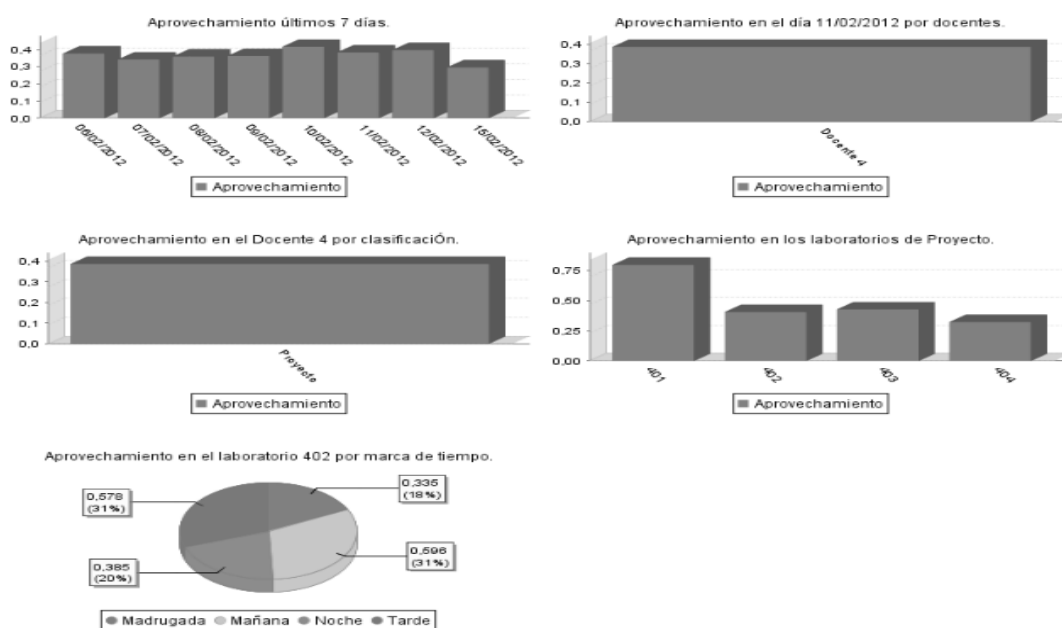


Figura 11: Cuadro de mando digital correspondiente al aprovechamiento de las computadoras.

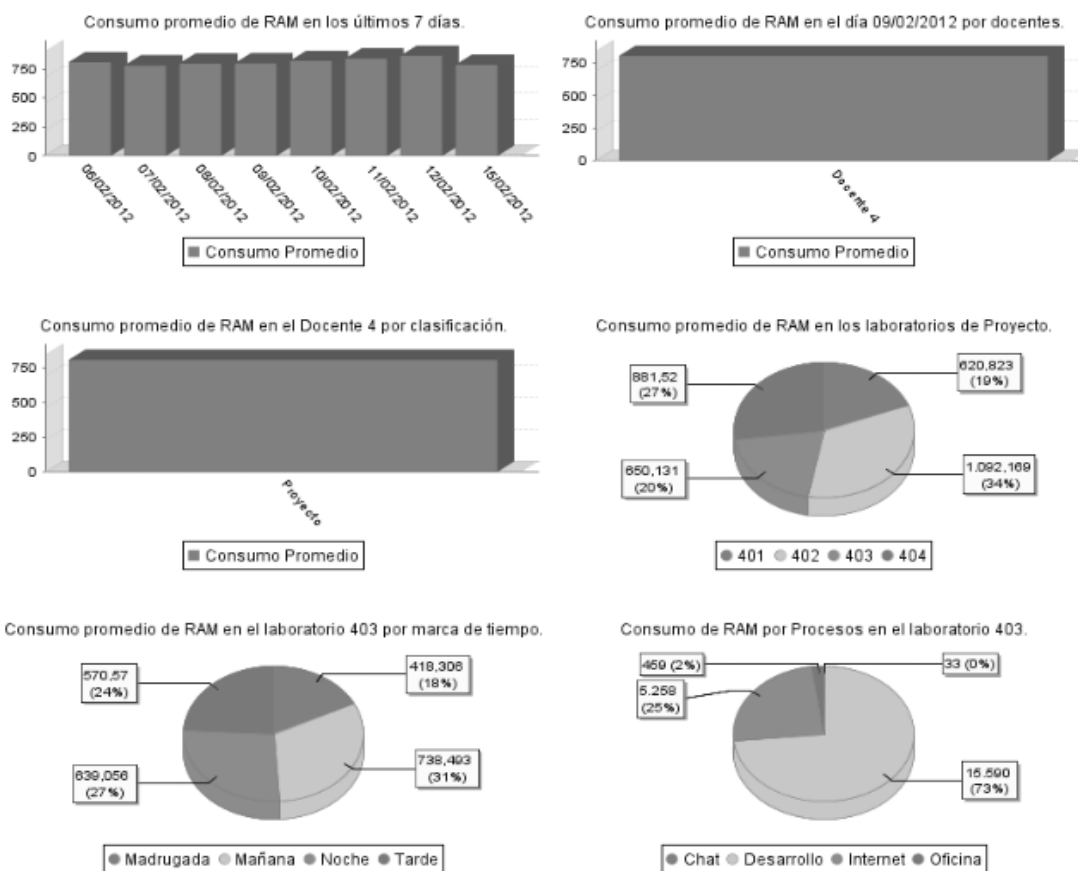


Figura 12: Cuadro de mando digital correspondiente al consumo de RAM de las computadoras.

Gráfico **Tráfico de red**: permite escoger en una lista desplegable la clasificación de los docentes, el docente y el laboratorio, y muestra en un gráfico de línea los paquetes enviados y paquetes recibidos por la red del laboratorio seleccionado durante las horas del día. En caso de no marcar algún criterio de selección se emite el mensaje “No hay información disponible”.

Gráfico **Observaciones**: permite escoger en una lista desplegable la clasificación de los docentes, el docente y el laboratorio, y muestra en un gráfico de barras el promedio de observaciones del laboratorio seleccionado durante la última semana. En caso de no marcar algún criterio de selección se emite el mensaje “No hay información disponible”.

Las figuras 13 y 14 muestran los ejemplos descritos anteriormente.

Clasificación: Proyecto Docente: Docente 4 Laboratorio: 402

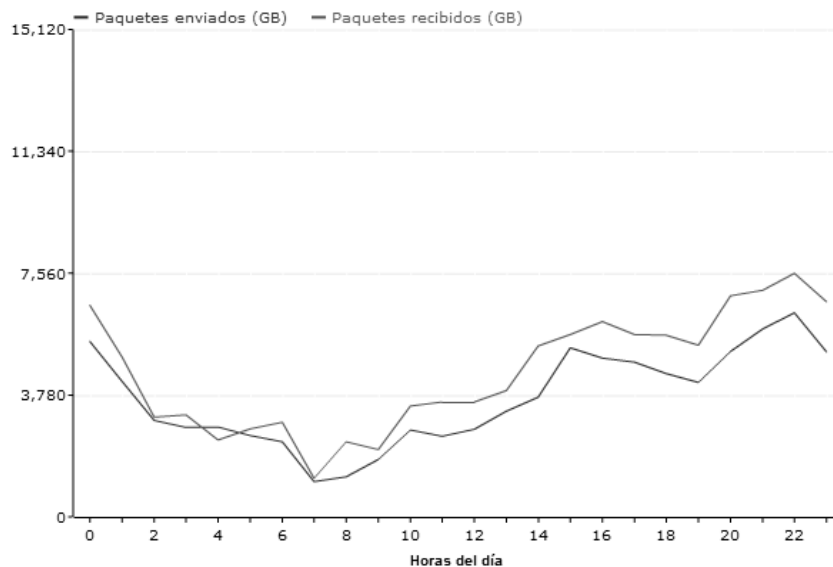


Figura 13: Gráfico dinámico correspondiente a la cantidad de paquetes recibidos y enviados por la red.

Clasificación: Proyecto Docente: Docente 4 Laboratorio: 401

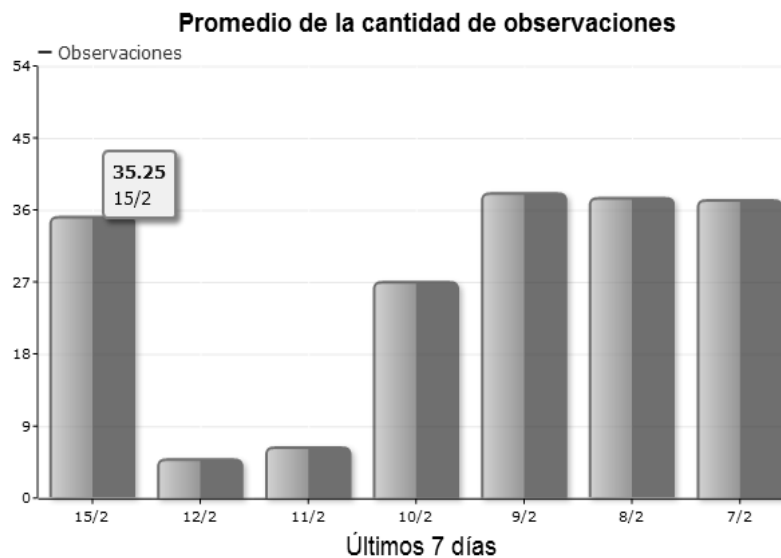


Figura 14: Gráfico dinámico correspondiente al promedio de las observaciones por laboratorios.

3.4 Pruebas

Para mejorar la calidad del producto de software es necesario realizar un conjunto de evaluaciones durante todo el proceso de desarrollo que implique al cliente y al desarrollador. El objetivo de diseñar las pruebas parte de un pensamiento enfocado a la calidad del proceso desde el inicio, permitiendo comprobar en varios momentos del ciclo de desarrollo la conformidad con los requisitos de software y la detección temprana y corrección de errores, de ahí que las pruebas deben ser planificadas teniendo en cuenta quien las va a realizar, en qué momento y como elemento más importante, qué es lo que se va a revisar. Siguiendo en todo momento la idea principal, que consiste en probar cada detalle, y probarlo varias veces bajo diferentes condiciones para de esta forma detectar la mayor cantidad de errores.

En la presente investigación para evaluar la capa de visualización del mercado de datos Tecnología se utilizó el Modelo V, que fue definido por CALISOFT¹⁵ y lo pone en práctica DATEC, con el fin de crear un estándar para comprobar que el producto cumpla con las especificaciones del negocio.

El modelo V constituye una evolución del modelo en cascada, con la diferencia que este considera que las actividades de prueba se ejecuten paralelamente con las actividades de análisis y diseño. Representa las etapas del ciclo de desarrollo de las pruebas con un nivel de abstracción elevado. Este modelo hace corresponder con cada etapa de desarrollo una prueba pertinente en su nodo paralelo de la columna de aseguramiento de la calidad. Cada una de las pruebas se encarga de prevenir futuros desperfectos en los requisitos, diseño e implementación del sistema.

¹⁵ Centro de Calidad para Aplicaciones Tecnológicas.

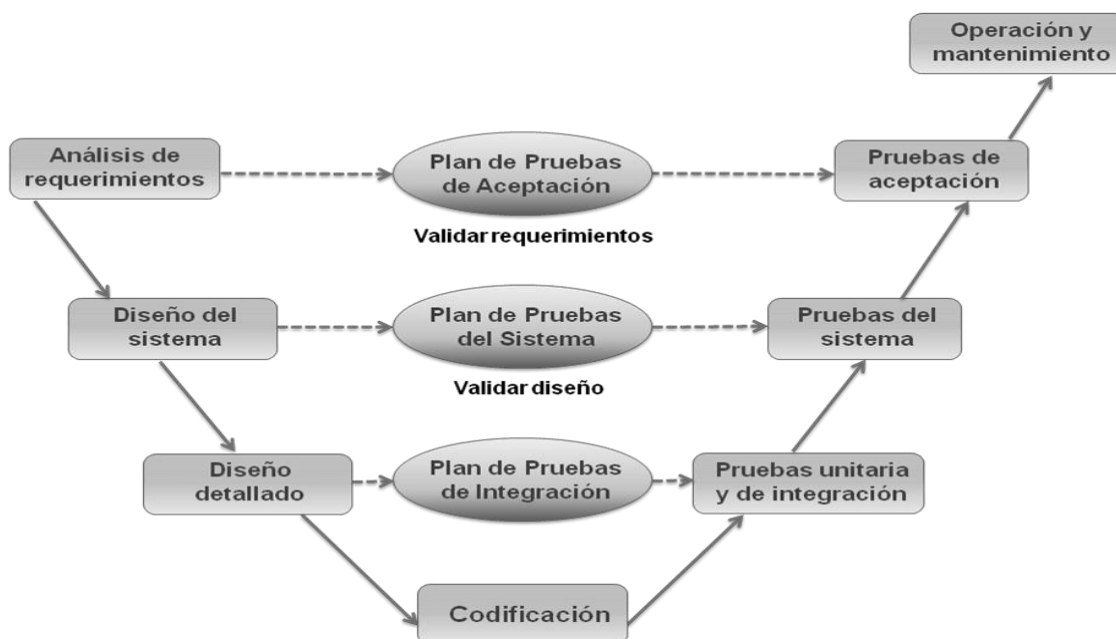


Figura 15: Modelo V. Permite realizar pruebas a lo largo del ciclo de desarrollo del software.

3.5 Herramientas para la realización de las pruebas

Casos de prueba: se utilizan para identificar posibles fallos de implementación, se basan en los CUI definidos en el DCUS, con el propósito de validar las necesidades del cliente. En el mercado de datos Tecnología se aplicaron tres casos de pruebas correspondientes a los CU. La siguiente tabla describe el caso de prueba asociado al CUI Obtener información de inventario. Las demás descripciones se encuentran en los Anexos.

Tabla 5: Diseño del caso de prueba inventario

Escenario	Descripción	Perfiles de análisis	Indicadores a medir	Respuesta del sistema	Flujo central
EC 1.1: Cantidad computadoras dominio	Muestra las cantidades de computadoras en el dominio	Tiempo Laboratorio	Cantidad de computadoras	El sistema muestra todas las variables disponibles para los análisis, ubicadas en las filas y las columnas que pueden ser visualizadas	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT.

				para cada vista de análisis.	Inventario. Selecciona la vistas de análisis cantidad de computadoras en el dominio
EC 1.2: Cantidad computadoras por tipo de SO	Muestra las cantidades de computadoras por tipo de SO	Tiempo Laboratorio	Cantidad de computadoras	El sistema muestra todas las variables disponibles para los análisis, ubicadas en las filas y las columnas que pueden ser visualizadas para cada vista de análisis	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Inventario. Selecciona la vista de análisis cantidad computadoras por tipo de SO
EC 1.3: Cantidad computadoras por tipo de CPU	Muestra las cantidades de computadoras por tipo de CPU	Tiempo Laboratorio	Cantidad de computadoras	El sistema muestra todas las variables disponibles para los análisis, ubicadas en las filas y las columnas que pueden ser visualizadas para cada vista de	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Inventario. Selecciona la vista de

				análisis.	análisis. cantidad de computadoras por tipo de CPU
EC 1.4: Cantidad computadoras por tipo de RAM	Muestra las cantidades de computadoras por tipo de RAM	Tiempo Laboratorio	Cantidad de computadoras	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las columnas que pueden ser visualizadas para cada vista de análisis.	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Inventario. Selecciona la vista de análisis cantidad de computadoras por tipo de RAM

3.6 Resultado de las pruebas

Pruebas unitarias

Las pruebas unitarias se enfocan en un programa o componente que desempeña una función específica, verificando que funcione tal y como lo define la especificación del programa. Los programadores siempre prueban el código durante el desarrollo, por lo que las pruebas unitarias son realizadas solamente después de que el programador considera que el componente se encuentra libre de errores.

Una vez concluida la etapa de implementación se realizaron las pruebas unitarias a la capa de visualización, donde se detectaron dos no conformidades, una no conformidad de complejidad alta y una no conformidad de complejidad baja, las cuales fueron resueltas satisfactoriamente (Figura 16).

NC1: Errores ortográficos, de acentuación en las vistas de análisis.

NC2: Se tienen datos fijos en los reportes.

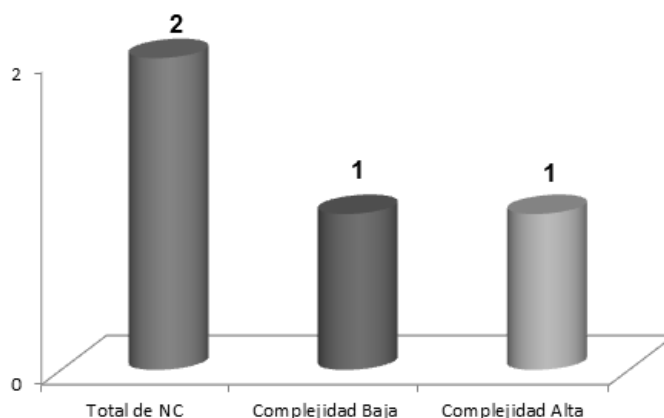


Figura 16: Resultados de las pruebas unitarias y de integración.

Pruebas de integración

Su objetivo es identificar los errores introducidos por la combinación de programas o componentes probados unitariamente, además, verificar que las especificaciones de diseño sean alcanzadas. No son verdaderamente pruebas de sistema debido a que los componentes no se encuentran implementados en el ambiente operativo.

Al sistema se le realizaron pruebas de integración donde no se detectaron ninguna no conformidad, al tener un correcto funcionamiento el subsistema de almacenamiento con el subsistema de visualización.

Pruebas del sistema

Son usualmente conducidas para asegurar que todos los módulos trabajan como sistema, sin error. Es similar a la prueba de integración pero con un alcance mucho más amplio. Las pruebas del sistema examinan qué tan bien el sistema cumple con los requerimientos de la organización y su utilidad, seguridad y desempeño. También se realizan estas pruebas a la documentación del sistema.

Las pruebas del sistema fueron aplicadas por el grupo de calidad interna del departamento de Almacenes de Datos, a través de los casos de prueba guiados por casos de uso. Los tres casos de prueba basados en CU permitieron identificar un total de cuatro NC, dos de complejidad alta, una de complejidad media y una de complejidad baja (Figura 17).

NC1: Uso excesivo de mayúsculas.

NC2: No coinciden las variables de entrada especificadas en los casos de prueba con las mostradas en la aplicación.

NC3: Los gráficos correspondientes a Métricas RAM y Métricas CPU no contienen las leyendas.

NC4: Errores ortográficos de acentuación en la vista de análisis Procesos por usuarios del LT Procesos.

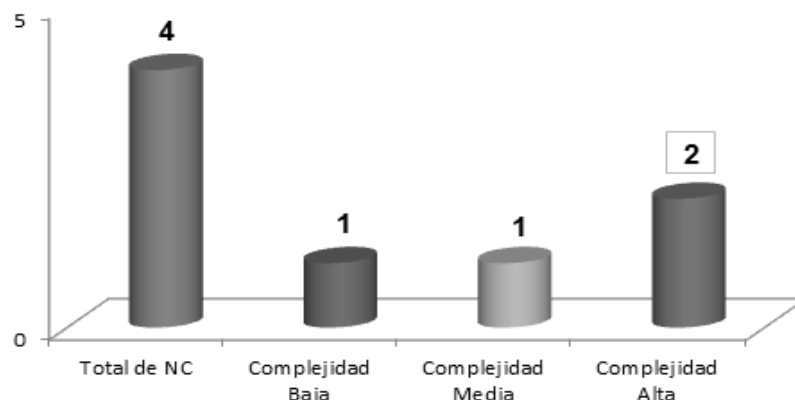


Figura 17: Resultados de la aplicación de los casos de prueba basados en CU al sistema.

Pruebas de aceptación

Las pruebas de aceptación son realizadas con el objetivo de determinar el grado de satisfacción del cliente con el sistema. También se confirma que el sistema está terminado y cumple con los requisitos definidos. Para verificar el estado de cumplimiento de la aplicación se realizó una entrevista con el cliente, donde se le mostró la aplicación y quedó conforme con la misma, firmando una carta de aceptación (ver Anexo figura 27). La siguiente figura muestra el resultado final de las pruebas aplicadas según la complejidad.

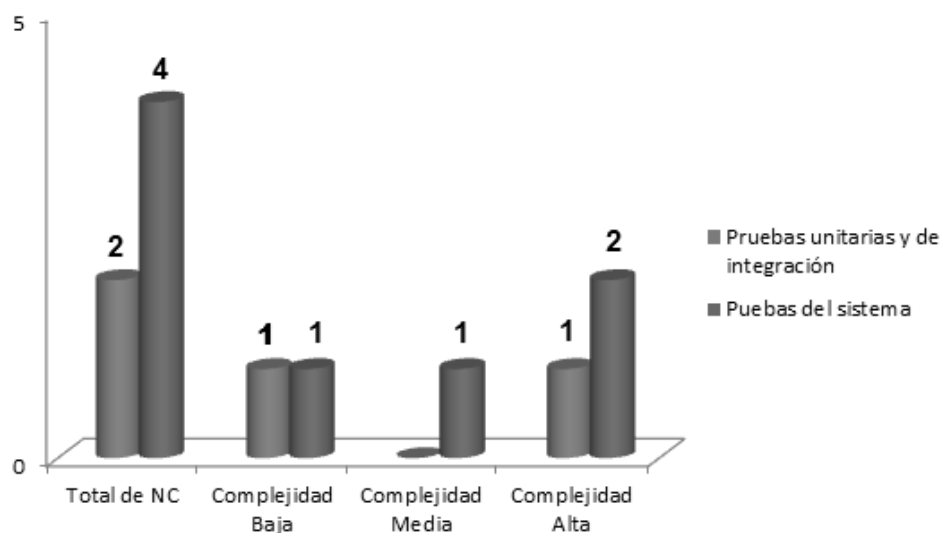


Figura 18: Resultado general de las pruebas aplicadas a la solución.

Conclusiones del capítulo

En la realización de este capítulo se abordaron los elementos referentes al desarrollo de las etapas de implementación y prueba de la capa de visualización para el mercado de datos Tecnología donde:

- ✓ Durante la implementación del subsistema de visualización se realizaron seis cubos OLAP, tres libros de trabajos contenidos en un área de análisis, 21 vistas de análisis, 10 cuadros de mando digitales, ocho gráficos dinámicos y un reporte operacional, permitiendo resolver las necesidades de información del cliente.
- ✓ Mediante las pruebas unitarias se detectaron dos NC, una no conformidad de complejidad baja y una no conformidad de complejidad alta.
- ✓ Se diseñaron y aplicaron tres casos de prueba para validar la solución, permitiendo identificar cuatro no conformidades, dos de complejidad baja, una de complejidad media y una de complejidad baja, las cuales fueron corregidas satisfactoriamente.
- ✓ Se comprobaron las funcionalidades del sistema con el especialista del área, donde se obtuvo la carta de aceptación que fue firmada por el cliente.

Conclusiones generales

Al finalizar el trabajo de diploma “Capa de visualización para el mercado de datos Tecnología para la Sala Situacional de la Universidad de las Ciencias Informáticas” se concluye que se ha cumplido con los objetivos específicos planteados. Para ello:

- ✓ La selección de la Propuesta de Metodología para el Desarrollo de Almacenes de Datos en DATEC garantizó el proceso de desarrollo del software. Las herramientas utilizadas permitieron desarrollar satisfactoriamente la aplicación y darle respuestas a las necesidades de información del cliente.
- ✓ Se realizó el diseño del subsistema de visualización, sirviendo como base para la fase de implementación, donde se diseñaron cinco cubos multidimensionales y nueve dimensiones.
- ✓ Se implementaron 21 vistas de análisis, 10 cuadros de mando digitales, ocho gráficos dinámicos y un reporte operacional, apoyando el proceso de toma de decisiones.
- ✓ La aplicación de las pruebas a la capa de visualización para el mercado de datos Tecnología, permitió identificar seis no conformidades que fueron resueltas satisfactoriamente.

Recomendaciones

Con el propósito de mejorar la propuesta realizada en este trabajo, se sugiere:

- ✓ Aplicar técnicas de minerías de datos que permitan detectar patrones de comportamiento sobre la información almacenada.

Referencia bibliográfica

1. **Inmon, Bill.** *Information Management on the five classes of ODS.*
2. Geocities. *Geocities.* [En línea] Lixbeth Rodríguez, 07 de 05 de 2007. [Citado el: 07 de 11 de 2011.] http://www.geocities.ws/lixbeth/Seminario/t3.html#_Toc159925936.
3. Data Warehousing. *Institute.* [En línea] 21 de 03 de 2007. [Citado el: 07 de 11 de 2011.] <http://tdwi.org/pages/education/education-main-page.aspx..>
4. Institute Data Warehousing. [En línea] [Citado el: 07 de 11 de 2011.] <http://tdwi.org/pages/education/education-main-page.aspx..>
5. **Santos, Romina Elizabeth.** *Bases de Datos Multiplataforma como Soporte Para La Inteligencia de Negocios.*
6. **Rodríguez, Anilsey Delfino.** *Cuadros de Mando Digitales, Soporte a la Toma de Decisiones.* La Habana : s.n.
7. Stratebi. *Stratebi.* [En línea] [Citado el: 28 de 10 de 2011.] http://www.stratebi.es/todobi/jun10/Comparativa_OSBI.pd.
8. **Alberto Límia Navarro, Anisley Delfino, Asnioby Hernández López, Doris Medina Mustelier, Iván M. Cárdenas Tandrón, Julio Ernesto Ortiz, Marisleidys Socas, Osniel Calvo, Yanet Peña Vazquez, Yanisbel Gonzales Hernández y otros.** *Metodología para el Desarrollo de Soluciones de Almacenes de Datos e Inteligencia de Negocios en DATEC.* Ciudad de La Habana: s.n., 2010.
9. **Sanz, Miguel Rodríguez.** [En línea] 21 de 12 de 2010. [Citado el: 10 de 11 de 2011.] <http://e-archivo.uc3m.es/handle/10016/9856>.
10. **Jorge Luis Tufiño López.** [En línea] 09 de 2011. [Citado el: 05 de 12 de 2011.] <http://bibdigital.epn.edu.ec/handle/15000/4101>.
11. **Few, Stephen.** *Information Dashboard Design.* 2006.
12. **Sánchez, Leopoldo Zenaido Zepeda.** *Metodología para el Diseño Conceptual de Almacenes de Datos.* Valencia : s.n., 2006.
13. DataCleaner reference documentation. *DataCleaner reference documentation.* [En línea] 01 de 2009. [Citado el: 14 de 11 de 2011.] <http://datacleaner.eobjects.org/resources/docs/documentation.html>.
14. Sinnexus. *Sinnexus.* [En línea] 2007. [Citado el: 18 de 11 de 2011.] http://www.sinnexus.com/business_intelligence/datawarehouse.aspx..
15. IBM Cognos. *IBM Cognos.* [En línea] [Citado el: 18 de 11 de 2011.] <http://www-01.ibm.com>.
16. LWP. *LWP.* [En línea] 30 de Octubre de 2011. [Citado el: 29 de Octubre de 2011.] http://www.lawebdelprogramador.com/foros/SAP/62311-Que_es_Microstrategy.html.

17. **Robert Stackowiak, Joseph Rayman, Rick Greenwald.** *Oracle Data Warehousing and Business Intelligence Solutions*. Indianapolis : Wiley Publishing, Inc., 2007. IN 46256.
18. Pentaho open source bussines intelligence. *Pentaho open source bussines intelligence*. [En línea] [Citado el: 18 de 11 de 2011.] <http://mondrian.pentaho.com>.
19. Portada sobre la plataforma Pentaho Open Source Business Intelligence. *Portada sobre la plataforma Pentaho Open Source Business Intelligence*. [En línea] 2006. [Citado el: 05 de 12 de 2011.] <http://pentaho.almacen-datos.com>.
20. Pentaho open source bussines intelligence. *Pentaho open source bussines intelligence*. [En línea] [Citado el: 03 de 12 de 2011.] <http://mondrian.pentaho.com>.
21. *Professional Apache Tomcat*.
22. **Wood, Sherman.** Pentaho Mondrian Project. *Pentaho Mondrian Project*. [En línea] 2008. [Citado el: 05 de 12 de 2011.] <http://mondrian.pentaho.com/documentation/workbench.php>.
23. Visual_Paradigm. *Visual_Paradigm*. [En línea] [Citado el: 28 de 11 de 2011.] http://www.ecured.cu/index.php/Visual_Paradigm.
24. PostgreSQL Tools. *PostgreSQL Tools*. [En línea] [Citado el: 10 de 12 de 2011.] <http://www.pgadmin.org>.
25. pgAdmin PostgreSQL Tools. *pgAdmin PostgreSQL Tools*. [En línea] [Citado el: 20 de 01 de 2012.] <http://www.pgadmin.org>.
26. Sixtina ConsultingGroup. *Sixtina ConsultingGroup*. [En línea] [Citado el: 22 de 11 de 2011.] <http://www.gestiopolis.com/administracion-estrategia/dashboard-kpi-metricas.htm#>.
27. **Inmon, W. H.** *Building the Data Warehouse*. s.l. : 0-471-08130-2, 2006.
28. OLAP. *OLAP*. [En línea] [Citado el: 05 de 12 de 2011.] <http://www.informatica-hoy.com.ar/Cubo-OLAP-una-base-de-datos-multidimensional.php..>
29. **RUMBAUGH, J.** *El lenguaje Unificado de modelado*. 2000.

Bibliografía

1. **Cabrera María Evelia Casales** Facultad de Ciencias [Online]. Facultad de Ciencias. Enero 2009. [Cited: Noviembre 07, 2011.] <http://hp.fciencias.unam.mx/~alg/bd/dwh.pdf>
2. **DATEC**, METODOLOGÍA PARA EL DESARROLLO DE SOLUCIONES DE ALMACENES DE DATOS E INTELIGENCIA DE NEGOCIO EN DATEC, 2008.
3. **DÍAZ, S. O.** *Inteligencia de Negocios*, 2009.
4. **Espinosa Roberto.** El Rincón del BI [Online]. Mayo 10, 2010. [Cited: Marzo 2011.] <http://churriwifi.wordpress.com/2010/05/10/16-3-construccion-procesos-etl-utilizando-kettle-pentaho-data-integration/>.
5. **EVELIA, C.** [En línea] [Citado el: 12 de 04 de 2012.] <http://hp.fciencias.unam.mx/~alg/bd/dwh.pdf>
6. **Few, Stephen.** *Information Dashboard Design*. 2006.
7. **GALICIA, B. I.** [En línea] [Citado el: 24 de 05 de 2012.] http://www.sinnexus.com/business_intelligence
8. **Huamantumba Rayner.** [Online]. Agosto 31, 2007. [Cited: Noviembre 20, 2011] WWW.RUEDATECNOLOGICA.COM.
9. **Inmon, W. H.** *Building the Data Warehouse*. ISBN: 0-471-08130-24.
10. **Kimball, Ralph y Ross, Margy.** *The Data Warehouse Toolkit*. ISBN: 0-471-20024-7.
11. **MEDINA, M.** Vistas materializadas en almacenes de datos para el monitoreo y control de proyectos, 2011
12. **Ortiz, Marta Cecilia Ortiz.** La inteligencia de negocios aplicada a las organizaciones en Latinoamérica. [En línea] [Citado el: 4 de Noviembre de 2011.] http://www.google.com/url?sa=t&source=web&cd=1&ved=0CBcQFjAA&url=http%3A%2F%2Fww2.epm.com.co%2Fbibliotecaepm%2Fbiblioteca_virtual%2Fdocuments%2FLa_inteligencia_de_negocios_aplicada_a_las_organizaciones.pdf&rct=j&q=La%20inteligencia%20de%20negocios%20es%2.
13. **Pereira, J. E.** Cuadro de Mando Integral, CMI. [En línea] [Citado el: 27 de noviembre de 2011.] http://www.mercadeo.com/41_scorecard.htm.
14. **PENTAHO.** [En línea] [Citado el: 11 de 12 de 2011.] <http://mondrian.pentaho.com/>
15. **pgAdmin PostgreSQL Tools.** *pgAdmin PostgreSQL Tools*. [En línea] [Citado el: 20 de 01 de 2012.] <http://www.pgadmin.org>.
16. **PostgreSQL Tools.** *PostgreSQL Tools*. [En línea] [Citado el: 10 de 12 de 2011.] <http://www.pgadmin.org>.

17. **Rodríguez, Anisley Delfino.** *Cuadros de Mando Digitales, Soporte a la Toma de Decisiones.* Ciudad de La Habana : s.n., 2010.
18. **Santos, Romina Elizabeth Dos.** *Bases de Datos Multiplataforma como Soporte Para La Inteligencia de Negocios.*
19. **Sánchez, Leopoldo Zenaido Zepeda.** *Metodología para el Diseño Conceptual de Almacenes de Datos.* Valencia, España : s.n., 2008.
20. **Sanz, Miguel Rodríguez.** [En línea] 21 de 12 de 2010. [Citado el: 10 de 11 de 2011.] <http://e-archivo.uc3m.es/handle/10016/9856>.
21. **Wood, Sherman.** Pentaho Mondrian Project. *Pentaho Mondrian Project.* [En línea] 2008. [Citado el: 05 de 12 de 2011.] <http://mondrian.pentaho.com/documentation/workbench.php>.

Anexos

```

CREATE OR REPLACE VIEW
staging_tecnologia.t_metricas_estaticas AS
SELECT
    s.dim_temporal_dia_key, s.dim_laboratorio_id,
    s.dim_cpu_id, s.dim_so_id, s.dim_ram_id,
    s.pc_dominio, count(s.id_pc)::integer AS nro_pc
FROM (
    SELECT
        metricas_estaticas.dim_temporal_dia_key, metricas_estaticas.dim_laboratorio_id,
        metricas_estaticas.id_pc, metricas_estaticas.pc_dominio,
        max(metricas_estaticas.dim_cpu_id) AS dim_cpu_id,
        max(metricas_estaticas.dim_so_id) AS dim_so_id,
        max(metricas_estaticas.dim_ram_id) AS dim_ram_id
    FROM staging_tecnologia.metricas_estaticas
    GROUP BY
        metricas_estaticas.dim_temporal_dia_key, metricas_estaticas.dim_laboratorio_id,
        metricas_estaticas.id_pc, metricas_estaticas.pc_dominio
    ) s
WHERE (s.dim_temporal_dia_key IN
(
    SELECT dim_temporal_dia.dim_temporal_dia_key
    FROM dimensiones.dim_temporal_dia
    ORDER BY dim_temporal_dia.dim_temporal_dia_key DESC
    LIMIT 330
)
)
GROUP BY s.dim_temporal_dia_key, s.dim_laboratorio_id, s.dim_so_id, s.dim_cpu_id, s.dim_ram_id, s.pc_dominio
ORDER BY s.dim_temporal_dia_key DESC;

```

Figura 19: Vista lógica métricas estáticas

```

CREATE OR REPLACE VIEW staging_tecnologia.v_metricas_hdd AS
SELECT
    metricas_dinamicas.dim_temporal_dia_key, dim_laboratorio.dim_laboratorio_id,
    metricas_dinamicas.id_pc, dim_laboratorio.laboratorio_nro_pc,
    max(metricas_dinamicas.hora_fin) AS hora_fin,
    max(metricas_dinamicas.espacio_disco_utilizado)::real)::
    integer AS espacio_utilizado, max(metricas_dinamicas.espacio_disco_libre)::real)::integer AS espacio_libre,
    max(metricas_dinamicas.espacio_disco_utilizado)::
    integer + max(metricas_dinamicas.espacio_disco_libre)::integer AS espacio_total
FROM
    dimensiones.dim_laboratorio, staging_tecnologia.metricas_dinamicas
WHERE
    dim_laboratorio.dim_laboratorio_id = metricas_dinamicas.localizacion AND
    (metricas_dinamicas.dim_temporal_dia_key IN (
        SELECT dim_temporal_dia.dim_temporal_dia_key
        FROM dimensiones.dim_temporal_dia
        ORDER BY dim_temporal_dia.dim_temporal_dia_key DESC
        LIMIT 330
    ))
GROUP BY
    metricas_dinamicas.dim_temporal_dia_key, dim_laboratorio.dim_laboratorio_id,
    metricas_dinamicas.id_pc, dim_laboratorio.laboratorio_nro_pc;

```

Figura 20: Vista lógica métricas de disco duro.

```

CREATE OR REPLACE VIEW staging_tecnologia.v_metricas_utilizacion AS
SELECT
  t.dim_temporal_dia_key, t.dim_marca_tiempo_lab_id, t.dim_laboratorio_id,
  count(t.id_pc)::integer AS cantidad_pc, sum(t.t_encendido) AS t_encendido,
  sum(t.t_trabajo) AS t_trabajo, max(t.momento_nro_horas) AS momento_nro_horas,
  max(t.laboratorio_nro_pc) AS laboratorio_nro_pc
FROM (
  SELECT
    metricas_dinamicas.dim_temporal_dia_key,
    dim_laboratorio.dim_laboratorio_id, dim_marca_tiempo_dia.dim_marca_tiempo_lab_id,
    metricas_dinamicas.id_pc, dim_marca_tiempo_dia.momento_nro_horas,
    dim_laboratorio.laboratorio_nro_pc, sum(metricas_dinamicas.tiempo_encendido) / 60::numeric AS t_encendido,
    sum(metricas_dinamicas.tiempo_trabajo) / 60::numeric AS t_trabajo
  FROM
    dimensiones.dim_laboratorio, dimensiones.dim_marca_tiempo_dia, staging_tecnologia.metricas_dinamicas
  WHERE
    dim_laboratorio.dim_laboratorio_id = metricas_dinamicas.localizacion AND
    dim_marca_tiempo_dia.hora_inicio <= metricas_dinamicas.hora_inicio AND
    dim_marca_tiempo_dia.hora_fin >= metricas_dinamicas.hora_fin AND
    (metricas_dinamicas.dim_temporal_dia_key IN (
      SELECT dim_temporal_dia.dim_temporal_dia_key
      FROM dimensiones.dim_temporal_dia
      ORDER BY dim_temporal_dia.dim_temporal_dia_key DESC
      LIMIT 330))
  GROUP BY
    metricas_dinamicas.dim_temporal_dia_key, dim_laboratorio.dim_laboratorio_id,
    dim_marca_tiempo_dia.dim_marca_tiempo_lab_id, metricas_dinamicas.id_pc,
    dim_marca_tiempo_dia.momento_nro_horas, dim_laboratorio.laboratorio_nro_pc
) t
GROUP BY t.dim_temporal_dia_key, t.dim_laboratorio_id, t.dim_marca_tiempo_lab_id;

```

Figura 21: Vista lógica métricas utilización

```

CREATE OR REPLACE VIEW staging_tecnologia.v_procesos_usuario AS
SELECT
  usuario_x_procesos.dim_laboratorio_id, usuario_x_procesos.dim_temporal_dia_key,
  dim_proceso.dim_proceso_id, dim_usuario.dim_usuario_id,
  sum(usuario_x_procesos.tiempo_trabajo) AS tiempo_trabajo,
  sum(usuario_x_procesos.numero_procesos) AS numero_procesos,
  sum(usuario_x_procesos.uso_memoria) AS uso_memoria
FROM
  dimensiones.dim_proceso, metadatos.clasificacion_procesos, staging_tecnologia.usuario_x_procesos,
  dimensiones.dim_usuario
WHERE
  usuario_x_procesos.nombre_proceso = clasificacion_procesos.proceso_nombre AND
  dim_proceso.grupo::text = clasificacion_procesos.proceso_categoria AND
  dim_usuario.usuario::text = usuario_x_procesos.usuario AND
  clasificacion_procesos.proceso_categoria <> 'Sistema'::text AND
  (usuario_x_procesos.dim_temporal_dia_key IN (
    SELECT dim_temporal_dia.dim_temporal_dia_key
    FROM dimensiones.dim_temporal_dia
    ORDER BY dim_temporal_dia.dim_temporal_dia_key DESC
    LIMIT 330
  ))
GROUP BY
  usuario_x_procesos.dim_laboratorio_id, usuario_x_procesos.dim_temporal_dia_key,
  dim_proceso.dim_proceso_id, dim_usuario.dim_usuario_id;

```

Figura 22: Vista lógica procesos por usuarios



Figura 23: Representación del cubo métricas de RAM y red.

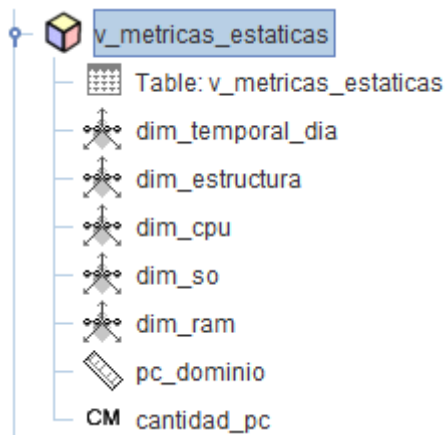


Figura 24: Representación del cubo métricas estáticas.

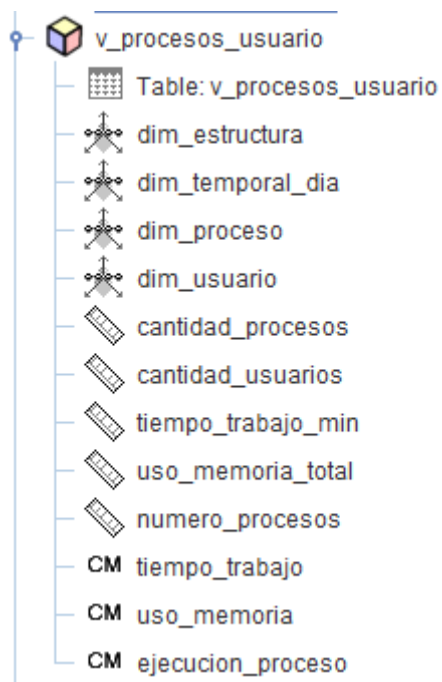


Figura 25: Representación del cubo procesos por usuario.

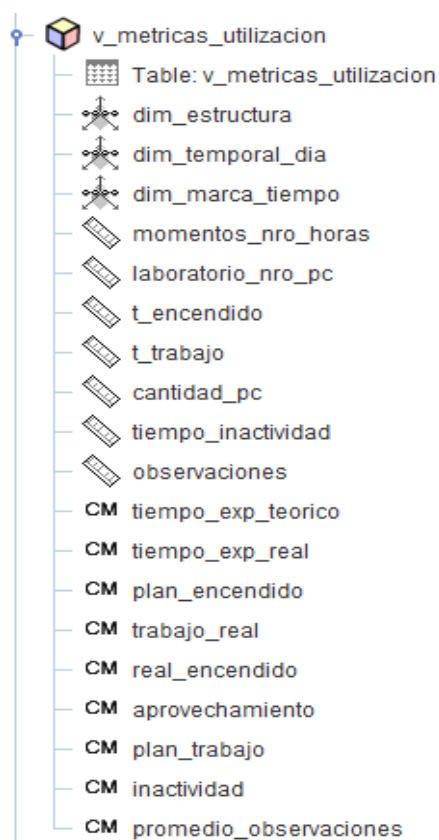


Figura 26: Representación del cubo métricas de utilización.

ACTA DE ACEPTACIÓN

En La Habana, a los 07 días del mes de noviembre del 2012

De una parte, el departamento Dirección de laboratorios, representado en este acto por Omar Mar Cornelio, quien a los fines y efectos derivados del presente documento se denominará como "El cliente", y de otra Parte, el centro de Tecnologías de Gestión de Datos, conocido de forma abreviada como DATEC de la Universidad de las Ciencias Informáticas (UCI), representada en este acto por Asnioby Hernández López, que a los fines y efectos derivados del presente documento se denominará DATEC.

Primero: Que en cumplimiento de los acuerdos, han sido efectuadas las actividades que se describen, **Las partes** DECLARAN:

CONSIDERANDO: Que se han efectuado las actividades siguientes:

1. – Diseño del subsistema de visualización de datos: Capa de visualización de Tecnología.
2. – Proceso de extracción, transformación y carga de los datos.
3. – Implementación de las vistas de análisis OLAP.

CONSIDERANDO: Que las actividades realizadas han sido desarrolladas con la calidad requerida y bajo las condiciones pactadas y aprobadas por **Las Partes**.

CONSIDERANDO: Que las actividades que se han ejecutado cumplen con los requerimientos de **El Cliente**.

CONSIDERANDO: Que DATEC ha entregado la documentación que avala la ejecución de este acto al **El Cliente**.

POR TANTO: **Las Partes** acuerdan formalizar mediante la presente Acta, Aceptadas las actividades que han sido ejecutadas en esta fecha.

Y para que así conste, se extiende la presente Acta en dos (2) ejemplares, rubricados por **Las Partes**.

Por El Cliente

Nombre y Apellidos

Por DATEC

Nombre y Apellidos

Figura 27: Carta de aceptación del cliente

Almacenes de Datos
BATEC
Centro de Tecnologías de Gestión de Datos

UCI Universidad de las Ciencias Informáticas

Acta de Liberación Interna de Productos Software

Fecha de liberación: 29/11/2012

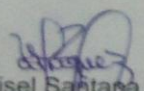
Emitida a favor de: Proyecto Sala Situacional de la Universidad de las Ciencias Informáticas (SSUCI). Departamento de Almacenes de Datos e Inteligencia de negocio, Centro de Tecnologías de Gestión de Datos.

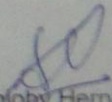
Nombre del Proyecto: Sala Situacional de la Universidad de las Ciencias Informáticas (SSUCI).

Nombre del Producto: Capa de visualización para el mercado de datos Tecnología.

1. Datos del producto

Artefacto	Versión	Estado final	Cantidad Iteraciones	Tipos de pruebas realizadas
Capa de visualización para el mercado de datos Tecnología.	1.0	0	2	Pruebas funcionales.


Ing. Marisel Santana Rodríguez
 Aseguradora de la Calidad
 Dpto. Almacenes de Datos


MsC. Asnoby Hernández López
 Jefe Dpto. Almacenes de Datos

Grupo Aseguramiento de la Calidad Departamento Almacenes de Datos (AD)

Figura 28: Acta de liberación interna del producto de software.

Tabla 6: Descripción textual del CU: Obtener información de procesos

Objetivo	Obtener la información de los procesos.	
Actores	Analista	
Resumen	El caso de uso inicia cuando el analista desea consultar la información relacionada con la ejecución de los proceso. Luego de seleccionar el reporte deseado, el sistema muestra la información contenida en él y las opciones de los posibles cambios que le puede hacer al reporte. El caso de uso finaliza cuando el especialista termina de analizar la información.	
Complejidad	Alta	
Prioridad	Media	
Precondiciones	La información relacionada con la ejecución de los procesos se cargó correctamente en el mercado de datos.	
Postcondiciones	Los reportes correspondientes al libro de trabajo L.T Procesos han sido consultados por el especialista.	
Flujo de eventos		
Flujo básico Obtener información de procesos		
Actor	Sistema	
1. El especialista se autentica en el sistema.	2. Muestra la interfaz principal con las áreas de análisis existentes.	
3. El especialista selecciona el área de análisis general A.A.G Sala Situacional UCI.	4. Muestra las áreas de análisis que están contenidas dentro del A.A.G Sala Situacional UCI.	
5. El especialista selecciona el área de análisis A.A Tecnología.	6. Muestra los libros de trabajo que están contenidos dentro del A.A Tecnología.	
7. El especialista selecciona el libro de trabajo L.T Procesos.	8. Muestra los reportes contenidos dentro del L.T Procesos.	
9. El especialista selecciona el reporte deseado.	10. Muestra la información contenida en el reporte seleccionado y brinda la posibilidad al especialista de hacerle cambios al reporte para su análisis. Ir al Caso de Uso: Visualizar reporte. Finaliza el caso de uso.	

Opciones de reportes de procesos		
Perspectivas de análisis	Posibles resultados	
	Medidas	Periodicidad
Variables de entrada relacionadas con el CU Mostrar información de Procesos. Usuario Tiempo Laboratorio	Variables de salida disponibles en el hecho Procesos. Cantidad de procesos Uso RAM	Rango de tiempo en que se solicitan las variables de salida: Diario

Tabla 7: Descripción textual del CU: Obtener información de inventario

Objetivo	Obtener la información de Inventario.
Actores	Analista
Resumen	El caso de uso inicia cuando el analista desea consultar la información relacionada con el inventario de las computadoras. Luego de seleccionar el reporte deseado, el sistema muestra la información contenida en él y las opciones de los posibles cambios que le puede hacer al reporte. El caso de uso finaliza cuando el especialista termina de analizar la información.
Complejidad	Alta
Prioridad	Media
Precondiciones	La información relacionada con el inventario de las computadoras se cargó correctamente en el mercado de datos.
Postcondiciones	Los reportes correspondientes al libro de trabajo L.T Inventario han sido consultados por el especialista.
Flujo de eventos	
Flujo básico Obtener información de inventario	
Actor	Sistema
1. El especialista se autentica en el sistema.	2. Muestra la interfaz principal con las áreas de análisis existentes.

3. El especialista selecciona el área de análisis general A.A.G Sala Situacional UCI.	4. Muestra las áreas de análisis que están contenidas dentro del A.A.G Sala Situacional UCI.
5. El especialista selecciona el área de análisis A.A Tecnología.	6. Muestra los libros de trabajo que están contenidos dentro del A.A Tecnología.
7. El especialista selecciona el libro de trabajo L.T Inventario.	8. Muestra los reportes contenidos dentro del L.T Inventario.
9. El especialista selecciona el reporte deseado.	10. Muestra la información contenida en el reporte seleccionado y brinda la posibilidad al especialista de hacerle cambios al reporte para su análisis. Ir al Caso de Uso: Visualizar reporte. Finaliza el caso de uso.
Opciones de reportes de inventario	
Perspectivas de análisis	Posibles resultados
	Medidas
	Periodicidad
Variables de entrada relacionadas con el CU Mostrar información de Inventario. Tiempo Laboratorio RAM SO CPU	Variables de salida disponibles en el hecho Inventario. Cantidad de computadoras
	Rango de tiempo en que se solicitan las variables de salida: Diario

Tabla 8: Diseño del caso de prueba explotación

Escenario	Descripción	Perfiles de análisis	Indicadores a medir	Respuesta del sistema	Flujo central
-----------	-------------	----------------------	---------------------	-----------------------	---------------

EC 1.1: Aprovechamiento	Muestra el aprovechamiento de las computadoras	Tiempo Laboratorio Marca tiempo	Aprovechamiento	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las columnas que pueden ser visualizadas para cada vista de análisis.	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Explotación. Selecciona la vista de análisis correspondiente.
EC 1.2: Cantidad de tiempo encendido y tiempo de trabajo	Muestra la cantidad de tiempo encendido y tiempo de trabajo de las computadoras	Tiempo Laboratorio Marca tiempo	Cantidad de tiempo encendido y tiempo de trabajo	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las columnas que pueden ser visualizadas para cada vista de análisis	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Explotación Selecciona la vista de análisis correspondiente.
EC 1.3: Tiempo trabajo según explotación teórica	Muestra el tiempo trabajo según el total de computadoras en el laboratorio	Tiempo Laboratorio Marca tiempo	Tiempo trabajo según explotación teórica	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las columnas que pueden ser visualizadas para cada vista	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Explotación.

				de análisis.	Selecciona la vista de análisis correspondiente
EC 1.4: Tiempo de trabajo según explotación real	Muestra el tiempo de trabajo según el total de las computadoras encendidas	Tiempo Laboratorio Marca tiempo	Tiempo trabajo según explotación real	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las columnas que pueden ser visualizadas para cada vista de análisis.	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Explotación. Selecciona la vista de análisis correspondiente.
EC 1.5: Tiempo encendido según explotación teórica	Muestra el tiempo de encendido según el total de computadoras en el laboratorio	Tiempo Laboratorio Marca tiempo	Tiempo encendido según explotación teórica	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las columnas que pueden ser visualizadas para cada vista de análisis.	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Explotación. Selecciona la vista de análisis correspondiente.
EC 1.6: Tiempo encendido según explotación real	Muestra el tiempo de encendido según la cantidad de computadoras encendidas	Tiempo Laboratorio Marca tiempo	Tiempo encendido según explotación real	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las columnas que pueden ser	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología.

				visualizadas para cada vista de análisis.	Selecciona el LT. Explotación. Selecciona la vista de análisis correspondiente.
EC 1.7: Cantidad de observaciones	Muestra la cantidad de observaciones	Tiempo Laboratorio Marca tiempo	Cantidad de observaciones	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las columnas que pueden ser visualizadas para cada vista de análisis.	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Explotación. Selecciona la vista de análisis correspondiente.
EC 1.8: Inactividad del CPU	Muestra la Inactividad del CPU	Tiempo Laboratorio Marca tiempo	Inactividad del CPU	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las columnas que pueden ser visualizadas para cada vista de análisis.	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Explotación. Selecciona la vista de análisis correspondiente.
EC 1.9: Disco duro	Muestra el espacio libre y usado en el disco duro por computadoras	Tiempo Laboratorio Marca tiempo	Disco duro	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador.

				columnas que pueden ser visualizadas para cada vista de análisis.	Selecciona el AAG. Tecnología. Selecciona el LT. Explotación. Selecciona la vista de análisis correspondiente.
EC 1.10: Promedio de paquetes enviados y recibidos	Muestra el promedio de paquetes enviados y recibidos por la red	Tiempo Laboratorio Marca tiempo	Promedio de paquetes enviados y recibidos	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las columnas que pueden ser visualizadas para cada vista de análisis.	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Explotación. Selecciona la vista de análisis correspondiente.
EC 1.11: Cantidad de paquetes enviados y recibidos	Muestra la cantidad de paquetes enviados y recibidos	Tiempo Laboratorio Marca tiempo	Cantidad de paquetes enviados y recibidos	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las columnas que pueden ser visualizadas para cada vista de análisis.	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Explotación. Selecciona la vista de análisis correspondiente.
EC 1.12: Porcentaje de paquetes enviados y recibidos	Muestra el porcentaje de paquetes enviados y recibidos por la	Tiempo Laboratorio Marca tiempo	Porcentaje de paquetes enviados y recibidos	El sistema muestra todas las variables disponibles para los análisis,	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral

	red			ubicados en las filas y las columnas que pueden ser visualizadas para cada vista de análisis.	izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Explotación. Selecciona la vista de análisis correspondiente.
EC 1.13: Consumo de RAM	Muestra el consumo de RAM por computadoras	Tiempo Laboratorio Marca tiempo	Consumo de RAM	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las columnas que pueden ser visualizadas para cada vista de análisis.	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Explotación. Selecciona la vista de análisis correspondiente.
EC 1.14: Cantidad computadoras reportadas	Muestra la cantidad computadoras reportadas	Tiempo Laboratorio Marca tiempo	Cantidad computadoras reportadas	El sistema muestra todas las variables disponibles para los análisis, ubicados en las filas y las columnas que pueden ser visualizadas para cada vista de análisis.	Se autentica. Entra al sistema. Despliega hacia la derecha el componente ubicado en el lateral izquierdo que contiene el navegador. Selecciona el AAG. Tecnología. Selecciona el LT. Explotación. Selecciona la vista de análisis correspondiente.

Glosario de términos

Área de análisis: se le denomina área de análisis a la agrupación de información según su propósito, aunque el criterio depende de las necesidades de la institución o empresa donde se aplica el sistema. Permite restringir el número de usuarios que acceden a los datos.

BI: Inteligencia del negocio.

Cubo: colección de dimensiones y medidas en un área temática particular.

DATEC: Centro de Tecnologías de Gestión de Datos.

Libro de Trabajo: estructura organizativa que agrupa las vistas de análisis generados dentro de las áreas de análisis. Puede ser creado teniendo en cuenta criterio que permitan organizar la información: Emisor de las vistas de análisis, receptor de las vistas de análisis, contenido, etc.

SSUCI: Sala situacional de la Universidad de las Ciencias Informáticas.

SQL: lenguaje de consulta estructurado o SQL (por sus siglas en inglés Structured Query Language) es un lenguaje declarativo de acceso a bases de datos relacionales que permite especificar diversos tipos de operaciones en ellas.

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