



UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS

FACULTAD 7

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

Cuadro de Mando Integral para el módulo Consulta Externa del
Sistema de Información Hospitalaria xavia HIS

Autores: Yudith Morell Moya

Daryel Cruz Vega

Tutores: MSc. Yovannys Sánchez Corales

Ing. Arturo Orellana García

La Habana, junio de 2013

“Año 55 de la Revolución”



Tu tiempo es limitado, de modo que no lo malgastes viviendo la vida de alguien distinto. No quedés atrapado en el dogma, que es vivir como otros piensan que deberías vivir. No dejes que los ruidos de las opiniones de los demás acallén tu propia voz interior. Y, lo que es más importante, ten el coraje para hacer lo que te dicen tu corazón y tu intuición.

Steve Jobs

Datos de contacto

MSc. Yovannys Sánchez Corales: Profesor asistente graduado en el año 2005 de Ingeniero en Informática en la CUJAE. Pertenece al Centro de Desarrollo de Software para la Salud. Ha impartido las asignaturas de Inteligencia Artificial, Programación III y Práctica Profesional. Forma parte del proyecto de Atención Primaria de Salud.

Ing. Arturo Orellana García: Graduado de Ingeniero en Ciencias Informáticas en el año 2012 en la UCI. Recién Graduado en Adiestramiento. Se desempeña como desarrollador del sistema alas SIAPS.

Dedicatoria

A mis padres porque me dieron la vida, en especial a mi madre por toda la seguridad y optimismo que siempre me ha transmitido.

A mis abuelos por su sabiduría y paciencia ante todo, en específico a la persona que hizo posible que este sueño se cumpliera, mi abuela (Pucha) que ha sido como una madre para mí, y a la memoria de mi abuelo José porque aunque no esté físicamente, es y será mi faro guía en los caminos de la vida.

Daryel Cruz Vega

A mis padres por su amor y apoyo incondicional, por darme todo lo que han podido y más.

A mi hermana y mi primo Omar Jesús para que un día puedan también ser universitarios.

A mi abuela Fredy por ser como otra madre para mí.

A la memoria de mi bisabuela Argentina a pesar de que ya no esté físicamente por su cariño y dedicación.

Yudith Morell Moya

Agradecimientos

Dicen que el que persevera triunfa y por eso quiero agradecer este triunfo de manera exclusiva a mis padres y abuelos. A mis tíos Miguel y Frank por todo su apoyo y consejos en los momentos difíciles. A mi familia de manera general y a mi novia y compañera de tesis por todo su cariño, dedicación y paciencia.

Agradezco profundamente a Adalennis que siempre me tendió su mano y sin su apoyo no hubiese sido posible realizar este sueño, a mi tutor Yovannys por su constancia y exigencia, a Frank y Arturo por su apoyo, a Dainerys por su ayuda incondicional, a Raiko y Yoana por su empeño y muestras de apoyo.

También a Daylén por su preocupación y ayuda, a mis amigos de forma muy especial a Asiel.

A quienes de una manera u otra aportaron su granito de arena para lograr ser quién soy hoy en día. A todos MUCHAS GRACIAS.

Daryel Cruz Vega

A mis padres por estar allí siempre para mí, por haberme dado una formación de la que hoy estoy orgullosa, por haber sido mis guías todos estos años, los amo.

A mi abuela Fredy por su cariño y preocupación.

A mis tías Yurelkis y Yamila por su cariño y apoyo.

A mi novio y compañero de tesis por su cariño y paciencia.

A mi amiga Adalennis por su amistad durante todos estos años y su ayuda incondicional para la realización de esta tesis como una tutora más.

A mi amiga Liseth por saber escuchar y preocuparse por mis estudios.

A mis amistades Cheila, Daysi, Diana y Yaneslearser por todos los buenos y malos momentos que hemos compartido durante estos cinco años.

A mis tutores Frank, Arturo y en especial a Yovannys por su exigencia y guía para la realización de esta tesis.

A todos los que de una forma u otra me han ayudado a llegar hasta aquí MUCHAS GRACIAS.

Yudith Morell Moya

Resumen

Como aplicación de las Tecnologías de la Información y las Comunicaciones (TIC) se puede mencionar la Inteligencia de Negocios (*Business Intelligence – BI por sus siglas en inglés*), la cual apoya los procesos relacionados con la toma de decisiones tácticas o estratégicas. Como un tipo de herramienta de *BI* se destacan los Cuadros de Mando Integrales (CMI), que priorizan y profundizan el seguimiento de indicadores más que el análisis detallado de la información.

El Sistema de Información Hospitalaria xavia HIS es una aplicación informática que gestiona los procesos clínicos fundamentales en el nivel hospitalario. Uno de sus componentes lo constituye el módulo Consulta Externa. En él se emiten reportes estadísticos a partir de la información que se almacena sobre los pacientes, consultas, referencias y contrarreferencias. Dichos reportes están limitados a la toma de decisiones a un nivel operacional, imposibilitando que la información sea analizada a modo de resumen atendiendo a diferentes indicadores o mediante una relación causa efecto.

Esta investigación tiene como objetivo desarrollar un CMI para el módulo Consulta Externa del Sistema de Información Hospitalaria xavia HIS sustentado en un mercado de datos (*Data Marts – DM por sus siglas en inglés*) que permita la toma de decisiones tácticas y estratégicas. La investigación es validada utilizando el modelo V el cual arrojó un índice de satisfacción grupal de un 0,81.

Palabras claves: Consulta Externa, Cuadro de Mando Integral, Inteligencia de Negocios, Mercado de Datos.

Introducción.....	1
Capítulo 1: Fundamentos teóricos de los cuadros de mando integrales.....	6
1.1 Reseña histórica de los cuadros de mando integrales.....	6
1.2 Inteligencia de negocio.....	7
1.3 Cuadros de mando integrales.....	7
1.4 Almacenes de datos.....	9
1.5 Análisis de soluciones existentes.....	17
1.6 Herramientas y tecnologías utilizadas.....	18
Conclusiones del capítulo.....	21
Capítulo 2: Características y diseño del cuadro de mando integral.....	23
2.1 Caracterización del entorno.....	23
2.2 Reglas del negocio y transformación.....	23
2.3 Necesidades de los usuarios.....	25
2.4 Especificación de requisitos.....	25
2.5 Casos de uso del sistema.....	27
2.6 Arquitectura.....	32
2.7 Diseño del subsistema de almacenamiento.....	34
2.8 Diseño del subsistema de integración.....	38
2.9 Diseño del subsistema de visualización.....	42
2.10 Esquema de seguridad.....	46
Conclusiones del capítulo.....	48
Capítulo 3: Implementación y validación del cuadro de mando integral.....	49
3.1 Implementación del subsistema de almacenamiento.....	49
3.2 Implementación del subsistema de integración.....	49
3.3 Implementación del subsistema de visualización.....	54
3.4 Pruebas.....	61
3.5 Diagrama de despliegue de la solución.....	69
Conclusiones del capítulo.....	70
Conclusiones.....	71
Recomendaciones.....	72
Referencias Bibliográficas.....	73
Glosario de Términos.....	76
Anexos.....	77

Introducción

En las últimas décadas el mundo se ha visto inmerso en un gran desarrollo tecnológico, lo cual ha dado paso al surgimiento de las TIC. Sus ventajas pueden ser aprovechadas en diferentes sectores de la sociedad para mejorar la calidad en los servicios (Cabero, 2005). Específicamente en el ámbito empresarial las organizaciones están favorecidas con el perfeccionamiento de sus estrategias de trabajo. La mejora continua de procesos ha sido necesaria para poder sostenerse en un mercado donde crece la demanda de la calidad de los productos.

Como aplicación de las TIC se puede mencionar *BI*, la cual apoya los procesos relacionados con la toma de decisiones tácticas y estratégicas. Existen varias formas de aplicar *BI* tales como: los Sistemas de Información para Ejecutivos (*Executive Informations System – EIS por sus siglas en inglés*) (Nader, 2003), los Sistemas de Soporte a las Decisiones (*Decision Support System - DSS por sus siglas en inglés*), y los Cuadros de Mando Integrales (Dávila, 1999).

Los *EIS* se refieren a cualquier programa que muestre y monitoree información ejecutiva de las diferentes áreas del negocio. En el momento que se desee conocer en detalle algún dato en un *EIS*, el *DSS* debe permitir profundizar la información, o sea, conocer el detalle de la misma y de las partes que la componen. Los CMI se diferencian del resto de los sistemas de *BI*, en que priorizan y profundizan en el seguimiento de indicadores más que en el análisis detallado de la información. Todas estas herramientas tienen como base un almacén de datos (*DW – Data Warehouse por sus siglas en inglés*) encargado de gestionar, integrar y centralizar la información de una entidad para su posterior análisis. Los mismos se componen por *DM* que son la implementación de un *DW* pero restringidos a un departamento específico de una entidad.

Generalmente la utilización de los CMI ha tenido éxito en empresas tales como: *Mobil Corporation*, *FMC Corporation* y *AT&T*; pertenecientes al sector privado (Gascho & Salterio, 2000). En la actualidad diversas entidades no lucrativas (fundaciones, hospitales públicos), muestran cierto interés en emplear un CMI (Eunice & Ripoll, 2003), en los siguientes estudios (Sanchez, Natalini, et al., 2010), (Sánchez & Días, 2008), (Flores, Uris, et al., 2009) se aprecian algunas de estas entidades. En estas organizaciones el desarrollo del CMI no ha tenido un desenvolvimiento como en las empresas del sector privado. Las organizaciones que optan por utilizar esta herramienta de análisis, cuentan con información precisa para facilitar la toma de decisiones.

Como aplicaciones informáticas para facilitar la toma de decisiones a un nivel operacional en el sector sanitario se destacan los Sistemas de Información para la Salud (SIS) (Tamborero, 2008). Dichos sistemas permiten elevar la calidad de la atención médica mediante informes y el registro de las operaciones diarias basado en las historias clínicas de los pacientes.

El Sistema de Información Hospitalaria xavia HIS es desarrollado por el Centro de Informática Médica (CESIM), perteneciente a la Universidad de las Ciencias Informáticas (UCI) en Cuba. Se especializa en el procesamiento de datos médicos y el almacenamiento de la información resultante. Además, permite la gestión de grandes volúmenes de datos, generados por los procesos de atención al paciente. Incluye además, un conjunto de módulos que implementan los principales procesos hospitalarios, siendo el módulo Consulta Externa uno de sus principales.

En este módulo se emiten reportes estadísticos a partir de la información que se almacena sobre los pacientes, consultas, referencias y contrarreferencias. Dichos reportes no permiten la configuración y adecuación de los mismos a las necesidades de los directivos médicos. Además están limitados a la toma de decisiones a un nivel operacional, imposibilitando que la información sea analizada a modo de resumen atendiendo a diferentes indicadores, tales como: cantidad, porciento y promedio.

Las privaciones anteriores traen consigo que en ocasiones los directivos médicos inviertan mucho tiempo en obtener un resumen del comportamiento de las consultas externas, lo que dificulta la realización de diferentes tipos de análisis basados en una relación causa efecto. Como consecuencia no se permite establecer comparaciones en el tiempo ni con otros indicadores, tales como la cantidad real y la planificada de pacientes atendidos.

La información que gestiona el módulo Consulta Externa constituye una motivación para determinar indicadores que permitan monitorear y dar seguimiento a las consultas externas, así como favorecer la toma de decisiones tácticas y estratégicas vinculada con la dirección médica, dando la posibilidad de ver la evolución de la entidad en esta área, el comportamiento de sus objetivos y mejorar los procesos de la misma.

Por lo anteriormente planteado se define como **problema científico**: ¿cómo contribuir a la toma de decisiones en el módulo Consulta Externa del Sistema de Información Hospitalaria xavia HIS?

Se plantea como **objeto de estudio** los Cuadros de Mando Integrales para la toma de decisiones, enmarcado en el **campo de acción** Cuadros de Mando Integrales para la toma de decisiones a partir de la información recogida en el módulo Consulta Externa del Sistema de Información Hospitalaria xavia HIS.

Se define como **objetivo general** de la investigación: desarrollar un Cuadro de Mando Integral para el módulo Consulta Externa del Sistema de Información Hospitalaria xavia HIS sustentado en un mercado de datos que permita la toma de decisiones tácticas y estratégicas.

El objetivo general se ha desglosado en los siguientes **objetivos específicos**:

- ✓ Realizar el marco teórico relacionado con los Cuadros de Mando Integrales para la salud enfocados a las consultas externas.
- ✓ Desarrollar el Cuadro de Mando Integral correspondiente al módulo Consulta Externa del Sistema de Información Hospitalaria xavia HIS.
- ✓ Validar el Cuadro de Mando Integral correspondiente al módulo Consulta Externa del Sistema de Información Hospitalaria xavia HIS.

Se propone la realización de las siguientes **tareas de la investigación** para dar cumplimiento a los objetivos planteados:

1. Análisis de la fuente de datos existente en el Sistema de Información Hospitalaria xavia HIS.
2. Análisis de los trabajos relacionados con las técnicas de *BI* aplicadas en la salud.
3. Obtención de los artefactos correspondientes a las fases de Requisitos, Arquitectura y Prueba, mediante la propuesta de metodología para el desarrollo de almacenes de datos.
4. Diseño del mercado de datos.
5. Perfilado e integración de los datos.
6. Diseño de los reportes y la estructura del Cuadro de Mando Integral.
7. Validación del correcto funcionamiento del Cuadro de Mando Integral.

Los **métodos de investigación** utilizados se exponen a continuación:

Métodos Teóricos

- ✓ Históricos – Lógicos: para elaborar la fundamentación teórica de la investigación, ya que permiten analizar la trayectoria completa de los CMI y realizar un estudio histórico acerca de los mismos.
- ✓ Análisis – Síntesis: para facilitar el estudio de los CMI se analizarán las diferentes bibliografías consultadas y se sintetizarán las partes previamente analizadas para descubrir sus características generales y las relaciones esenciales entre ellas.
- ✓ Inducción – Deducción: para arribar a proposiciones generales e inferir casos particulares por un razonamiento lógico partiendo del conocimiento general de los CMI.

Métodos Empíricos

- ✓ Observación: para la corroboración empírica de las deficiencias detectadas en cuanto a la gestión de la información en las consultas externas, y para la implantación y validación empírica del sistema elaborado.
- ✓ La medición: con el objetivo de obtener información numérica acerca de una propiedad o cualidad de los CMI, donde se compararan magnitudes medibles y conocidas. Se utiliza el método descriptivo para organizar y clasificar los indicadores cuantitativos obtenidos en las investigaciones empíricas a través de gráficos.
- ✓ Entrevista: se basa en las entrevistas a las analistas del módulo Consulta Externa para validar cuáles de los datos que se tenían en la base de datos fuente eran necesarios controlar y cuáles se podían descartar, además, para la obtención de conocimiento manejando términos, diagnósticos y tratamiento para los procesos de las consultas externas que se analiza.

El presente trabajo se encuentra estructurado en tres capítulos:

Capítulo 1: Fundamentos teóricos de los cuadros de mando integrales

Se expone y documenta el estado del arte relacionado con los CMI. Se definen los conceptos fundamentales relacionados con el tema para un mejor entendimiento, así como las herramientas, tecnologías y metodologías que van a permitir una adecuada solución al problema.

Capítulo 2: Características y diseño del cuadro de mando integral

Luego de realizar un estudio inicial del negocio son identificados los elementos del análisis. Se especifica la arquitectura del sistema propuesto, permitiendo el diseño de los subsistemas que lo componen. La

solución está conformada por un mercado de datos y una herramienta de consulta o análisis de tipo CMI. Se describe el perfilado de datos que asegura la calidad de los mismos y se elabora una política de respaldo y recuperación para la seguridad del sistema. Para la correcta realización del CMI se seleccionan los indicadores correspondientes a las perspectivas propuestas en el capítulo 1.

Capítulo 3: Implementación y validación del cuadro de mando integral

Una vez realizado el diseño multidimensional guiado por la metodología utilizada, se procede a realizar la implementación física del sistema, que implica la implementación de cada uno de los subsistemas que conforman el sistema propuesto. Finalmente se exponen las pruebas realizadas al mismo para su validación, así como los resultados obtenidos en cada una de ellas luego de su aplicación para garantizar el buen funcionamiento del producto final.

Capítulo 1: Fundamentos teóricos de los cuadros de mando integrales.

En el presente capítulo se expone y documenta el estado del arte relacionado con los CMI, así como las tendencias actuales. Se definen los conceptos fundamentales relacionados con el tema para un mejor entendimiento. Además el estudio de las herramientas, tecnologías y metodologías permiten dar soporte a la solución del problema en cuestión.

1.1 Reseña histórica de los cuadros de mando integrales

El concepto de Cuadro de Mando (CM) se deriva del término francés *tableau de bord*, que significa tablero de mando, o cuadro de instrumentos. Su origen se remonta a mediados del siglo XX, aunque es alrededor de 1948 cuando surge la noción de CM empresarial en los Estados Unidos. El desarrollo del CM se basa en cinco ideas esenciales (Santos & Fidalgo, 2006):

- ✓ Herramienta de ayuda durante el proceso de toma de decisiones.
- ✓ Diseño sencillo y eficaz.
- ✓ Aglutina indicadores financieros y no financieros.
- ✓ Flexible frente a los cambios y progresos del entorno.
- ✓ Genera motivación a todos los niveles de responsabilidad.

En Estados Unidos, en la década de los sesenta, la empresa *General Electric* desarrolló un tablero de control para llevar a cabo el monitoreo de los procesos de la misma. Contaba con ocho áreas claves de resultados, entre ellas estaban: rentabilidad, cuota de mercado, formación o responsabilidad pública. *General Electric* detallaba indicadores que resultaran útiles para dar seguimiento y controlar la obtención de objetivos tanto a corto como a largo plazo.

En esta misma década también se extiende a Francia el uso de *tableau de bord*. El tablero de mando combinaba varios ratios o indicadores para el control financiero de la empresa. Con el devenir de los años esta herramienta ha avanzado y combina no sólo ratios financieros, sino también indicadores no financieros los cuales posibilitan controlar los diferentes procesos del negocio (Dávila, 1999).

Actualmente, en el ámbito de la gestión empresarial, se manifiesta una fuerte inclinación a la adaptación de los sistemas de información. Los mismos van surgiendo cada vez con necesidades más específicas relacionadas con la implementación de la estrategia empresarial.

Capítulo 1: Fundamentos teóricos de los cuadros de mando integrales

1.2 Inteligencia de negocio

El término *BI* es la integración por un lado del almacenamiento y por el otro del procesamiento de grandes cantidades de datos, cuyo objetivo principal es transformarlos en conocimiento y permitir la toma de decisiones en tiempo real, a través del análisis y la exploración. Además, se enfoca en la búsqueda de información que no solo se encargue de responder a preguntas de lo que está sucediendo o ya sucedió, sino también, permita la construcción de modelos, a través de los cuales se podrán predecir eventos futuros. Algunos de los principales beneficios se muestran a continuación (Bernabeu, 2007):

- ✓ Proporciona herramientas de análisis para establecer comparaciones y tomar decisiones.
- ✓ Permite a los usuarios no depender de reportes o informes programados, porque los mismos serán generados de manera dinámica.
- ✓ Permitirá predecir el comportamiento futuro con un alto porcentaje de certeza, basado en el entendimiento del pasado.
- ✓ El usuario podrá consultar y analizar los datos de manera sencilla.

1.3 Cuadros de mando integrales

En la década de los noventa, los CMI se definían como: *...un conjunto de indicadores que proporcionan a la alta dirección una visión comprensiva del negocio...* y desde su primera formulación ha evolucionado hasta llegar a ser *...una herramienta de gestión estratégica que trata de articular la correcta implantación de la estrategia a través de indicadores financieros y no financieros...* (Dávila, 1999). Sin embargo en la actualidad constituyen Sistemas de Medición Estratégicos basado en cuatro perspectivas para guiar la implementación de la estrategia en sus factores críticos de éxito y su adaptación al entorno (Kaplan & Norton, 2005). Las perspectivas son las que a continuación se describen:

- ✓ **Financiera:** Incorpora la visión de los accionistas y mide la creación de valor de la empresa.
- ✓ **Cliente:** Muestra cuál es la posición de la empresa en los segmentos del mercado donde a esta le interesa competir.
- ✓ **Interna:** Recoge indicadores de procesos internos que son críticos para el posicionamiento en el mercado y para llevar la estrategia a buen puerto.
- ✓ **Formación y crecimiento:** Crea modelo de negocio apropiado que aprecia la importancia de invertir, innovar y mejorar. Valora el nivel profesional de los involucrados en el negocio, así como las investigaciones e innovaciones de los mismos.

Capítulo 1: Fundamentos teóricos de los cuadros de mando integrales

Como se puede apreciar estas perspectivas están enfocadas al sector financiero, pero pueden ser modificadas en cuanto al número y naturaleza de las que se consideren necesarias en cada caso dependiendo de las necesidades propias de las diferentes entidades.

De acuerdo con lo planteado por (Santos & Fidalgo, 2006), estas perspectivas pueden ser adaptadas al sector de la salud. En tal sentido la principal transformación que se debe realizar es en cuanto a la perspectiva clásica del cliente, sustituyéndose por los diferentes grupos de interés. Para esta investigación la perspectiva clásica del cliente está constituida por los pacientes. En relación a la perspectiva de los procesos internos se introducen las subperspectivas consultas, referencias y contrarreferencias que responden a los procesos significativos que se llevan a cabo en el módulo de Consulta Externa.

1.3.1 Principales características

Entre las características que poseen los CMI se mencionan (Dávila, 1999):

- ✓ Combinan indicadores financieros y no financieros: los indicadores financieros son instrumentos limitados porque sólo explican lo que ha pasado, por lo que se combinan con indicadores no financieros que adelanten lo que luego reflejan los indicadores financieros.
- ✓ Recogen la idea de utilizar indicadores que realicen el seguimiento de la estrategia de una empresa.
- ✓ Ofrecen un método estructurado de selección de indicadores y esto le concede más versatilidad dentro de la gestión de la empresa.
- ✓ Las relaciones causa-efecto son el motor del modelo de negocio. Para dirigir de forma proactiva hay que actuar sobre las causas y no sobre las consecuencias.

1.3.2 Ventajas y desventajas

La aplicación de los CMI trae aparejado ventajas y desventajas en la capa clínica del sector de la salud, las mismas se describen a continuación (Santos & Fidalgo, 2006):

- ✓ El director de la unidad de salud o las direcciones municipales y provinciales pueden tener un seguimiento y control de sus acciones y una visión de la repercusión de su contribución debido al análisis de la información.
- ✓ Permite planificar y establecer objetivos, así como metas a cumplir en un período de tiempo determinado.

Capítulo 1: Fundamentos teóricos de los cuadros de mando integrales

- ✓ Comunicar y vincular los objetivos e indicadores estratégicos, lo que permite comunicar en términos prácticos los objetivos y vincularlos entre sí mediante la relación causa-efecto.
- ✓ Sistematizar la revisión de los objetivos identificando nuevas iniciativas.
- ✓ El éxito en el diseño e implantación del CMI depende de factores importantes como: la definición de las perspectivas, la selección de indicadores, el proceso de creación y la filosofía que trata de comunicarse a través del CMI.

Las ventajas antes mencionadas, deben adecuarse a las características de la institución, de lo contrario se transforman en limitaciones que pueden atentar contra el buen funcionamiento del CMI. A continuación se ponen de manifiesto las debilidades de dicha herramienta:

- ✓ Visión parcial de la organización: en la definición de la estrategia en conjunto con los objetivos específicos se deben tener en cuenta las dimensiones del negocio, de lo contrario el CMI perdería su cualidad integradora.
- ✓ Enfoque generalista de la estrategia: la estrategia no debe ser muy genérica, pues se perdería el enfoque hacia las unidades de negocio de la organización.
- ✓ Exceso o defecto de las mediciones seleccionadas: la cantidad de medidas no deben estar en exceso ni en defecto o podría provocar una inutilidad de la herramienta en el seguimiento diario de los procesos de las consultas externas.

Los CMI constituyen una herramienta de análisis que utilizan las bases de datos multidimensionales conocidas como almacén de datos, que incluye mecanismos propios para extraer, transformar y cargar los datos existentes en los sistemas operacionales o fuentes de datos, siendo nombrados estos mecanismos como proceso de extracción, transformación y carga (*Extraction, Transformation Loading - ETL por sus siglas en inglés*).

1.4 Almacenes de datos

Uno de los conceptos más famosos sobre *DW* es el de W. H. Inmon, quien lo define de la siguiente manera (Inmon, 2002):

“Un Data Warehouse es una colección de datos orientada a temas, integrada, variante en el tiempo y no volátil para el soporte del proceso de toma de decisiones de la gerencia”.

El almacén de datos maneja un gran cúmulo de datos, debido a que fortalece en su estructura la información recogida durante años, procedente de diversas fuentes, en un solo lugar centralizado. Es por esta razón que el depósito puede ser soportado y mantenido sobre diversos medios de almacenamiento.

Además, el *DW* presenta la información sumariada y adicionada desde múltiples versiones, y maneja información histórica. Organiza y almacena los datos que se necesitan para el procesamiento analítico e informático, con la intención de responder a preguntas de negocios y para que los usuarios finales puedan tomar decisiones sobre los datos a pesar de grandes cúmulos de información.

Un *DW* está compuesto por dimensiones, hechos y medidas. Las dimensiones representan aspectos de interés para que los usuarios puedan filtrar y manipular la información almacenada en las tablas de hechos. Los hechos están compuestos por medidas numéricas y las llaves asociadas a las dimensiones con que se relacionan. Los hechos son datos temporales, que son filtrados, agrupados y examinados mediante los contextos delimitados en las tablas de dimensiones (Bernabeu, 2007). Las medidas son los datos numéricos de interés para el cliente que son almacenadas para su rápida respuesta a las peticiones de los usuarios (Martínez, 2007). Las mismas tienen la propiedad de ser aditivas, es decir, que se puede realizar agrupaciones y sumas según todas las dimensiones. Hay otras medidas que son no aditivas según ninguna dimensión, y pueden agregarse calculando valores medios (Zorrilla, 2007).

1.4.1 Topologías para los almacenes de datos

Para el modelado de los almacenes de datos existen tres topologías (Bernabeu, 2007), las cuales se relacionan a continuación:

Esquema en Estrella

El esquema contiene una tabla de hechos central y varias tablas de dimensiones asociadas a esta, a través de sus respectivas llaves. El modelo no debe estar normalizado, es decir, no puede mostrarse en tercera forma normal (3ra FN), trayendo aparejada como principal ventaja obviar uniones (*JOIN*) entre las tablas cuando se realizan consultas. Genera, además, un cierto grado de redundancia, pero el ahorro de espacio no es significativo en este tipo de modelos.

A pesar que es el esquema más fácil de interpretar y optimiza los tiempos de respuesta ante las consultas de los usuarios, es el menos robusto para la carga y es el más lento de construir.

Esquema Copo de Nieve

El esquema contiene una tabla de hechos central, que se relaciona con una o varias tablas de dimensiones y estas a su vez pueden estar relacionadas o no con una o más tablas de dimensiones. Las tablas dimensiones están normalizadas, por lo que el diseño se torna menos complejo y es muy útil en las tablas de dimensiones de muchas tuplas. Sin embargo presenta dos grandes inconvenientes, al poseer múltiples dimensiones y cada una de ellas con jerarquías, el número de dimensiones puede llegar a ser inmanejable. Además, al existir múltiples uniones el desempeño puede verse reducido.

Esquema Constelación

El esquema está compuesto por una serie de esquemas en estrella formado por una tabla de hechos principal y una o más tablas de hechos auxiliares, todas ellas relacionadas con sus respectivas dimensiones. Las diferencias que presenta con el esquema en estrella son sus principales ventajas. Entre ellas se destacan:

- ✓ Permite varias tabla de hechos, por lo cual se pueden analizar más aspectos del negocio con un menor esfuerzo adicional de diseño.
- ✓ Permite la reutilización de dimensiones, ya que una misma dimensión puede relacionarse con varias tablas de hechos.

1.4.2 Dimensiones lentamente cambiantes

Las dimensiones en las cuales sus datos tienden a cambiar y modificarse a través del tiempo, ya sea de forma esporádica o periódica, o implique a una sola tupla o la tabla completa, suelen llamarse dimensiones lentamente cambiantes (*Slowly Changing Dimensions - SCD por sus siglas en inglés*). Cuando ocurre este tipo de cambios, se puede optar por realizar el registro del historial de cambios o reemplazar los valores que sean necesarios.

Inicialmente Ralph Kimball planteó tres estrategias a seguir cuando se tratan las *SCD*: tipo 1, tipo 2 y tipo 3 (Kimball & Ross, 2002). A través de los años, se ha profundizado en el estudio de las definiciones iniciales y han surgido los tipos 0, 4 y 6. Los tipos de *SCD* se describen a continuación:

- ✓ **Tipo 0 (no tiene en cuenta la gestión histórica):** no se realiza ningún esfuerzo para lidiar con los problemas del cambio de la dimensión. De esta manera se sobrescribe parte de la información, mientras que otra queda intacta.

- ✓ **Tipo 1 (sobrescribir):** se utiliza cuando la información histórica no es importante. Sobrescribe los datos antiguos con nuevos y es usado por lo general con el propósito de corregir errores de datos en las dimensiones. Presenta como desventaja principal que no persiste ningún registro histórico en la dimensión.
- ✓ **Tipo 2 (añadir fila):** al ocurrir algún cambio en la dimensión se crea una nueva entrada en la tabla. Al nuevo registro es asignada una nueva llave subrogada, las cuales son de tipo serial y no representan información significativa del negocio, valor que será usado para futuras entradas, mientras que las antiguas utilizarán el valor anterior. De esta manera se gestionan versiones que pueden incluir fechas para indicar los períodos de validez, así como numeradores de registros o indicadores de registros activos o no.
- ✓ **Tipo 3 (añadir columna):** esta estrategia requiere que se agregue una nueva columna a la tabla de dimensión por cada campo cuyos valores deben incluir un historial de cambios. De este modo en la nueva columna se coloca el valor antiguo antes de sobrescribir el valor actual con el nuevo.

Presenta como principal desventaja que solo permite guardar un historial limitado de los datos, dependiendo del número de columnas que se cree.
- ✓ **Tipo 4 (tabla de historia separada):** se almacenan en una tabla adicional los detalles de cambios históricos realizados a la tabla de dimensión. La tabla que contempla la información histórica indicará el tipo de operación que se ha realizado, sobre qué campo se realizó el cambio y la fecha del mismo. Esta tabla tiene como objetivo mantener un detalle de los cambios realizados.
- ✓ **Tipo 6 (híbrido):** el método combina los tipos anteriores 1, 2 y 3; y se le denomina tipo 6 debido a la suma de los tres tipos que integra ($1+2+3=6$). Esta estrategia utiliza el Tipo 1 (sobrescribir) junto con el Tipo 2 (añadir filas) y el Tipo 3 (añadir columnas), añadiendo además, una pareja adicional de columnas para indicar el rango de fechas al cual aplica cada fila en particular.

1.4.3 Estrategia de indexado

Al analizar los datos de interés en una consulta particular, los Sistemas Gestores de Base de Datos (SGBD) tienen siempre la opción de leer todos los registros de una tabla, siendo el método de acceso preferido cuando se trata de tablas pequeñas. Sin embargo, un *DM* contiene tablas con grandes volúmenes de datos, sobre las que se realizan constantemente consultas de gran complejidad que solicitan la información que cumple determinados criterios de búsqueda. El hecho de que muchas de estas

consultas incluyan operaciones de *JOIN* entre tablas, resulta extremadamente costoso atendiendo a los tiempos de respuesta, sobrecarga de disco, concurrencia, entre otros. Con el objetivo de aumentar la velocidad de este tipo de operaciones varios SGBD ofrecen técnicas especializadas como los índices.

Un índice en el contexto de bases de datos es una estructura física de datos que permite un tipo de acceso alternativo al secuencial. Los índices son creados a partir de una o varias columnas de una tabla y evitan el escaneo completo de la misma. Las columnas de las tablas cuyo indexado es necesario, se determinan al conocer los datos y los tipos de consultas que se van a realizar.

Existen varios tipos de índice como son: los *Gis*, *Gin*, *Hash*, *B-Tree* (Cristiá, Morgado, et al., 2010). Los índices por lo general se construyen en forma de árbol balanceado (*B-Tree*), los cuales constituyen estructuras jerárquicas que permiten una búsqueda rápida para obtener la dirección de una fila específica con un determinado valor en una tabla. Un índice *B-Tree* está basado en un árbol balanceado que se procesa jerárquicamente para encontrar el valor de interés.

1.4.4 Mercado de datos

Los *DM*, son bases de datos departamentales, que representan las diferentes áreas sensibles del negocio pero con las mismas características de un *DW*. Existen diferentes autores que brindan sus definiciones, pero una de las más abarcadoras es la propuesta por Ralph Kimball: “...un conjunto flexible de datos, idealmente basado en el dato más atómico posible (*granular*) para ser extraído de las fuentes operacionales y presentado en un modelo simétrico (*dimensional*), que es más resistente cuando se enfrentan con las más inesperadas consultas de los usuarios...”.

Para la realización del presente trabajo se implementa un *DM* puesto que su desarrollo va dirigido a un área del negocio en específico, en este caso, el área de consulta externa. El *DM* posibilitará guardar la información referente a las consultas externas obtenida de la base de datos del xavia HIS, para su posterior análisis a través de los procesos de *ETL*.

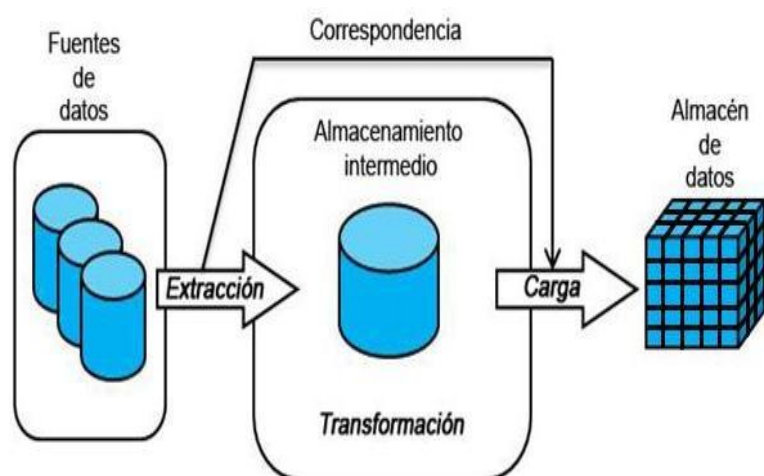
1.4.5 Procesos de extracción, transformación y carga

Para realizar el proceso de extracción de los datos de diversas fuentes que luego serán manipulados, integrados, transformados, y posteriormente cargados los resultados obtenidos en el *DM*, es necesario contar con algún sistema que se encargue de ello. Los procesos de *ETL* son los que cumplirán con esta meta. Los mismos se muestran en la figura 1.

A continuación se explican brevemente los procesos de *ETL* (Almenara, 2005).

Extracción: es el primer paso, apoyándose en las necesidades y requisitos del usuario para la obtención de la información de las diferentes fuentes hacia el *DM*. Los datos pueden encontrarse en formatos desiguales, teniendo en cuenta que pueden provenir de diversas fuentes. Este proceso es el encargado de dejar listo los datos para la transformación.

Transformación: luego de extraída la información se convierte aquellos datos inconsistentes en un conjunto de datos homogéneos y congruentes. Este proceso también se encarga de realizar el proceso de limpieza de datos.



Carga: una vez realizados los procesos de extracción y transformación y si ya se ha asegurado la calidad de los datos, se puede proceder a cargar la estructura de datos del *DM*.

Los procesos de *ETL* permiten dejar los datos listos para ser analizados por diferentes perspectivas para mejorar la toma de decisiones.

Figura 1: Procesos de *ETL*

Fuente: Elaboración propia

1.4.6 Procesamiento analítico en línea

El procesamiento analítico en línea (*System Online Analytical Processing – OLAP por sus siglas en inglés*) (Codd & Salley, 1993) define a una tecnología que se basa en el análisis multidimensional de los datos y que permite al usuario tener una visión rápida e interactiva de los mismos. Optimiza el uso de los *DW* para el análisis de datos en línea, por lo que hace más rápidas las respuestas a consultas analíticas complejas e iterativas, utilizadas generalmente para sistemas de ayuda para la toma de decisiones. *OLAP* permite que los usuarios tengan un mejor entendimiento de la información almacenada en sus bases de datos, lo cual es posible a través de un modelo de datos intuitivo y natural que proporciona. Esto permite que las empresas reconozcan mejor el valor de sus propios datos.

Los cubos *OLAP* constituyen una de las estructuras de datos utilizadas para representar la información de un *DM*. A través de los cubos *OLAP* los usuarios finales reciben una entrega optimizada de la información,

ya que se tienen algunos valores calculados en los datos antes de la solicitud (Nader, 2002). El uso de esta tecnología para visualizar la información relacionada con el módulo Consulta Externa apoya la toma de decisiones tácticas y estratégicas con mayor agilidad.

1.4.7 Metodologías para el desarrollo de almacenes de datos

Las metodologías de desarrollo constituyen la guía del proceso de desarrollo de software que indican qué hacer y cómo actuar ante un determinado problema de desarrollo. Existen numerosas metodologías para el diseño y construcción de un *DW*, de manera general predominan dos enfoques: el de Kimball (Kimball & Ross, 2002) y el de Inmon (Inmon, 2002). Desde el punto de vista arquitectónico, la mayor diferencia entre los dos autores radica en el sentido de construcción del *DW*. Kimball por su parte plantea un enfoque ascendente (*bottom-up*), mientras que Inmon defiende el enfoque descendente (*top-down*).

El enfoque *top-down* es utilizado cuando la tecnología y los problemas existentes en el negocio se conocen de antemano. Se trata de un método sistémico que reduce los problemas de integración, pero es costoso debido a la gran cantidad de datos y su poca flexibilidad. Por su parte el enfoque *bottom-up* se basa en experimentos y prototipos, cuya idea es implementar pequeños *DM* para evaluar las ventajas del nuevo sistema a medida que se avanza.

Existen metodologías que son adaptaciones de los enfoques anteriores como la propuesta por el Centro de Tecnologías de Gestión de Datos (DATEC) de la UCI. La misma constituye una adaptación de la metodología Ciclo de vida Kimball y tiene bien definidas sus fases, dentro de las cuales se encuentra la fase de prueba. Existen otras alternativas como Hefesto (Bernabeu, 2007) que se adaptan también a uno de los dos enfoques principales.

Luego de realizar un estudio de las principales metodologías utilizadas para el desarrollo de *DW*, y valoradas sus características, se escoge la Propuesta de Metodología para el Desarrollo de Almacenes de Datos diseñada en DATEC (González, 2012). La misma forma parte de una tesis de maestría y está validada por la implementación de los productos informáticos relacionados con *BI* que se elaboran en este centro.

Dicha metodología sigue el enfoque ascendente para la construcción de *DW*, la cual define claramente las actividades y tareas que deben realizarse para el adecuado desarrollo de un *DW*. Está en correspondencia con el modelo de calidad del Modelo de Madurez de las Capacidades Integradas (*Capability Maturity Model Integration – CMMI por sus siglas en inglés*), organiza sus roles por grupos de

Capítulo 1: Fundamentos teóricos de los cuadros de mando integrales

trabajo lo que permite aumentar la productividad y está creada para que pueda ser implementada con independencia de las herramientas que se pretendan utilizar. Se complementa con lo planteado por el doctor Leopoldo Zepeda y otros autores de incluir los casos de uso para guiar el proceso de desarrollo (Zepeda, 2008). Como principal aporte se agrega una fase de prueba que permite comprobar la calidad de los productos que se desarrollan (González, 2010).

Fases de desarrollo que presenta la propuesta de metodología

- ✓ **Estudio preliminar y planeación:** se realiza un estudio minucioso en la entidad cliente. Esto incluye un diagnóstico integral de la organización, con el fin de determinar qué es lo que se desea construir y qué condiciones existen para el desarrollo y montaje de la misma.
- ✓ **Requisitos:** se realiza el proceso de entrevistas al cliente para determinar los requisitos de información. Se definen, además, los requisitos funcionales y no funcionales. Se hace un levantamiento detallado de las fuentes de datos para validar la disponibilidad de la información.
- ✓ **Arquitectura:** se definen las vistas arquitectónicas de la solución, aspectos como los subsistemas y componentes, la seguridad, la comunicación y la tecnología a utilizar.
- ✓ **Diseño e Implementación:** se define el diseño de las estructuras de almacenamiento de datos, se diseñan los procesos de integración de datos, los cubos *OLAP* para la presentación de la información, así como el diseño gráfico de la aplicación definido por el cliente.
- ✓ **Prueba:** se realizan las pruebas que validan la calidad del producto, comenzando por las pruebas de unidad, las pruebas de integración y sistema, hasta llegar a las pruebas de aceptación con el cliente final.
- ✓ **Despliegue:** consta de dos etapas: despliegue piloto, donde se carga una muestra de los datos en un ambiente controlado, con el fin de mostrarle al cliente final el sistema en funcionamiento. Además, se realiza la capacitación y transferencia tecnológica de la solución a los clientes.
- ✓ **Soporte y Mantenimiento:** comienza cuando la solución está implantada y en explotación y se ejecuta según el contrato firmado y las condiciones de soporte establecidas. Además, se realizan las tareas de manteniendo de la aplicación que garantizan el adecuado funcionamiento y crecimiento del almacén de datos.

- ✓ **Gestión del proyecto:** se ejecuta a lo largo de todo el ciclo de vida del proyecto. Es aquí donde se controla, gestiona y chequea todo el desarrollo, los gastos, las utilidades, los recursos, las adquisiciones, los planes y cronogramas, entre otras actividades relacionadas con la gestión de proyectos.

Se selecciona la propuesta de metodología para conducir el proceso de construcción del *DM* en cuestión. Sus características se ajustan al presente trabajo debido a que se desarrolla previamente un *DM* y luego se aplica la herramienta de análisis de tipo CMI. Además, se facilita la implementación del CMI propuesto al permitir realizar prototipos para ser consultados posteriormente con las personas involucradas. La inclusión de los casos de uso permite representar de forma gráfica el proceso de desarrollo y a su vez guiarlo. La fase de prueba permitirá validar la calidad del sistema propuesto.

1.5 Análisis de soluciones existentes

Los sistemas informáticos para la salud ofrecen a los directivos una adecuada gestión de la información necesaria y oportuna para la toma de decisiones. Luego de un estudio del estado del arte se encontraron algunos trabajos que evidencian el uso de las herramientas utilizadas en *BI*, tales como *EIS* y CMI.

A continuación se muestran los relacionados con el sector sanitario en el ámbito internacional:

Planificación Sanitaria Utilizando Herramientas OLAP (Sanchez, Natalini, et al., 2010). En Argentina se presenta el desarrollo completo de un almacén de datos y herramientas de generación de reportes relacionados con el tratamiento de la tuberculosis. Este sistema es utilizado para monitorear el resultado de los tratamientos, a nivel de distrito para identificar problemas y a nivel provincial o nacional para asegurar el control consistente y de alta calidad en todas las áreas. Los productos involucrados en el desarrollo pertenecen al paquete de soluciones de *BI* de *Microsoft Corporation*®.

Sistema de Soporte a la Decisiones para la detección automatizada de Síndrome Metabólico (Sánchez, Días 2008). En Guanajuato, México se desarrolla una aplicación que muestra cómo a través de un sistema experto y el empleo de herramientas de *DSS*, se puede minimizar el tiempo en el diagnóstico del síndrome metabólico, así como generar un reporte impreso con los resultados, todo esto a partir de un almacén de datos utilizando herramientas *OLAP*. El sistema se desarrolló en Visual Fox versión 9.0, el cual corre en un sistema operativo Windows XP, y puede ser operado en red.

Cuadro de mando integral en el laboratorio clínico: indicadores de perspectiva interna del negocio (Flores & Uris, 2009). En Alicante, España se propone un sistema de indicadores como instrumento de dirección

Capítulo 1: Fundamentos teóricos de los cuadros de mando integrales

para el laboratorio clínico, desde la perspectiva de gestión interna de un CMI. Uno de los gráficos que se muestra contiene la evolución de indicadores de proceso analítico de rutina. Se muestra la evolución mensual de algunos indicadores de proceso analítico que representan el porcentaje de pruebas claves para el funcionamiento de las unidades funcionales (tirotopina, ferritina, sedimento urinario, antígeno y hemograma) que son validadas diariamente en las muestras que proceden de los centros de atención primaria.

En el ámbito nacional a pesar de no existir una gran variedad de trabajos relacionados con *BI*, en el sector de la salud se identificó:

EPIDEN (Sistema informático para el procesamiento transaccional y analítico de la información relacionada con la vigilancia epidemiológica del dengue) (López, Cuevas, et al., 2013). En Santiago de Cuba, Cuba se elaboró un sistema informático de apoyo a la toma de decisiones, donde se utilizaron técnicas de almacenes de datos y *OLAP* para la gestión de la información epidemiológica generada. Entre las principales características de este producto informático figuran: flexibilidad, capacidad de análisis, compatibilidad y gratuidad; también soporta la instrumentación de técnicas de procesamiento transaccional y analítico en línea, enfocado en la explotación de grandes volúmenes de datos del sistema de vigilancia epidemiológica del dengue. El sistema se desarrolló sobre la plataforma de desarrollo Pentaho BI Server, y MySQL como sistema gestor de bases de datos.

A pesar de sus potencialidades, estos sistemas se especializan en la toma de decisiones tácticas y estratégicas de una enfermedad específica. En el caso de *Planificación Sanitaria Utilizando Herramientas OLAP*, monitorea los resultados de los tratamientos de la tuberculosis. Por su parte el *Sistema de Soporte a la Decisiones para la detección automatizada de Síndrome Metabólico* minimiza el tiempo en el diagnóstico del síndrome metabólico. Por último el sistema *EPIDEN* gestiona la información epidemiológica generada. A partir de lo anterior se plantea que estos sistemas no se ajustan a las necesidades de las consultas externas identificadas en el presente trabajo, sin embargo se usaron de guía para comprender el funcionamiento de los sistemas de *BI*.

1.6 Herramientas y tecnologías utilizadas

Se hizo un análisis de las tendencias y tecnologías actuales relacionadas con *BI*, considerando las características del presente trabajo. Una de las tecnologías de *BI* la constituye el conjunto de herramientas de Pentaho las cuales se destacan por su libre distribución y uso. A continuación se describen algunas de estas herramientas:

Pentaho Data Integration

Se selecciona Pentaho Data Integration (*PDI – por sus siglas en inglés*) en su versión 4.2.1 (Medina, Iznaga, et al., 2012) esta es una herramienta líder de las soluciones de código abierto de *BI* que permite la ejecución de los procesos de *ETL* y su versión actual es 4.4.0. Mayormente las herramientas de *ETL* son utilizadas para trabajo con los almacenes de datos, *PDI* también puede ser usado para:

- ✓ Tareas de bases de datos, como consolidar, migrar y sincronizar bases de datos operativas.
- ✓ Migración de datos entre diferentes aplicaciones por cambios de versión o cambio de la aplicación.
- ✓ Sincronización entre diferentes sistemas operacionales.
- ✓ Consolidación de datos: sistemas con grandes volúmenes de datos que son consolidados en sistemas paralelos para mantener datos históricos o para procesos de borrado en los sistemas originales.

Pentaho Schema Workbench

El Pentaho Schema Workbench en su versión 3.2.0 cuenta con una interfaz que permite crear y probar esquemas de cubos *OLAP* visualmente. El motor Mondrian procesa las solicitudes de consultas de expresiones multidimensionales (*Multidimensional Query eXpression – MDX por sus siglas en inglés*) con los esquemas de procesamiento analítico en línea relacionales (*Relational On-Line Analytical Processing – ROLAP por sus siglas en inglés*). Estos archivos son modelos de esquema de lenguaje de marcado extensible (*Extensible Markup Language - XML por sus siglas en inglés*) de metadatos que se crean en una estructura específica utilizada por el motor de Mondrian. Los modelos *XML* se pueden considerar como las estructuras en forma de cubo que utilizan hechos existentes y tablas de dimensiones (Pentaho, 2007). La versión actual del Pentaho Schema Workbench es 3.3.0.

Pentaho BI Server

Se ha seleccionado el uso del Pentaho BI Server en su versión 4.1.0 como plataforma para el desarrollo del *DM* ya que es una aplicación *Java2EE* extensible, adaptable y configurable. Se integra con la mayoría de los entornos y se puede comunicar con otras aplicaciones vía servicios web. Integra todos los recursos de información en una única plataforma y proporciona ventajas al usuario y a los desarrolladores para crear contenidos (Pentaho, 2007). Su versión actual es 4.8.0.

Capítulo 1: Fundamentos teóricos de los cuadros de mando integrales

El Pentaho BI Server incluye el motor Mondrian que permite organizar la información en una estructura dimensional que brinda la posibilidad de moverse por la información a través de sus dimensiones. Mondrian combina la flexibilidad de los motores *ROLAP* con una caché que le proporciona velocidad. Además, puede ser integrado independientemente en cualquier otra plataforma (Stratebi, 2010).

La herramienta que se describe a continuación no es parte del conjunto de soluciones que ofrece la plataforma Pentaho, pero es útil en la implementación de un sistema de *BI*.

DataCleaner

El perfilado de los datos (Bernabeu, 2009) es una de las primeras tareas a realizar en el proceso de calidad de datos y consiste en realizar un análisis inicial sobre los datos de las fuentes, con el propósito de comenzar a conocer su estructura, formato y nivel de calidad.

En el presente trabajo se utiliza DataCleaner en su versión 1.5.3, la misma es una aplicación de código abierto para el perfilado, la validación y comparación de datos. Estas actividades ayudan a administrar y supervisar la calidad de los datos con el fin de garantizar que la información sea útil y aplicable a la situación de negocio. Su versión actual es 2.1.1.

Según su creador Kasper Sorensen, el sistema requiere *Java Runtime Environment 5.0* o una versión superior y drivers de *Java Database Connectivity (JDBC por sus siglas en inglés)* que permite que una aplicación java se conecte con una base de datos. Es una aplicación fácil de usar que genera sofisticados informes y gráficos. Permiten a los usuarios determinar el nivel de calidad de los datos, identificar y analizar la estructura del origen de los mismos para combinar resultados y gráficos, creando vistas fáciles de interpretar.

Visual Paradigm Enterprise Edition

Existen varias herramientas de Ingeniería de Software Asistida por Computadora (*Computer Aided Software Engineering - CASE por sus siglas en inglés*), entre las que se selecciona Visual Paradigm for UML Enterprise Edition en su versión 6.4, la cual es aplicable en todo el ciclo de vida del desarrollo de software. Soporta *Unified Modeling Language (UML por sus siglas en inglés)*, *Systems Modeling Language (SysML por sus siglas en inglés)*, *Business Process Modeling Notation (BPMN por sus siglas en inglés)*, entre otras tecnologías. Permite dibujar todos los tipos de diagramas de clases, generar código desde diagramas y generar documentación. Presenta licencia gratuita y comercial. Es fácil de instalar y actualizar, además de ser compatible entre ediciones. Su versión actual es 10.1.

Capítulo 1: Fundamentos teóricos de los cuadros de mando integrales

Visual Paradigm está compuesta por productos que facilitan a las organizaciones la visualización y diseño de diagramas, así como su integración con lenguajes de programación y SGBD. Es independiente de las plataformas, soportando varios entornos integrados de desarrollo, entre los que se pueden mencionar: Microsoft Visual Studio y Eclipse.

Sistema Gestor de Base de Datos PostgreSQL

Se selecciona PostgreSQL en su versión 8.4 siendo un SGBD objeto–relacional, multiusuario, centralizado y de propósito general. Está liberado bajo la licencia de código abierto (*Berkeley Software Distribution – BSD por sus siglas en inglés*), por lo que es un software preparado para ser modificado siempre y cuando sea para su mejora (Cristiá, Morgado, et al., 2010). Su versión actual es 9.2.4.

PgAdmin III

Se selecciona PgAdmin en su versión 1.10.0 (pgAdmin, 2010). El mismo está desarrollado bajo licencia de código abierto. Es una herramienta que permite gestionar las bases de datos de PostgreSQL. Está diseñada para responder a las necesidades de obtención de información por parte de los usuarios, desde escribir consultas de lenguaje estructurado (*Structured Query Language – SQL por sus siglas en inglés*) simples hasta desarrollar bases de datos complejas. Su versión actual es 1.16.1.

Con la selección de las herramientas a utilizar se garantiza el soporte a cada fase de desarrollo debido a la compatibilidad entre cada una de ellas y el cumplimiento de las políticas de migración a software libre. De esta forma se prevé concluir la presente investigación con resultados satisfactorios, logrando el objetivo propuesto.

Conclusiones del capítulo

Una vez estudiados los fundamentos teóricos del objeto de estudio de la investigación se concluye que el capítulo constituye la base sobre la cual se construirá la solución del presente problema. Además, se realizó un estudio de todos los temas teóricos que soportan la investigación científica, relacionada con el objeto de estudio del presente trabajo, demostrándose que los requerimientos y necesidades descritas no están contemplados en los sistemas existentes.

El sistema propuesto constituye una solución viable a la problemática planteada, teniendo en cuenta que se permite el almacenamiento de la información concerniente a los procesos de las consultas externas, en función del apoyo a la toma de decisiones por parte de los directivos del sector de la salud. Además, la

Capítulo 1: Fundamentos teóricos de los cuadros de mando integrales

propuesta de metodología seleccionada permite guiar el proceso de construcción del sistema a través de cada fase del ciclo de vida.

Capítulo 2: Características y diseño del cuadro de mando integral

Luego de realizar un estudio preliminar son identificadas las características del sistema. Además, se especifica la arquitectura del mismo, permitiendo el diseño de los subsistemas que lo componen. La solución está conformada por un mercado de datos y una herramienta de consulta o análisis de tipo CMI. Se describe el perfilado de datos que asegura la calidad de los mismos y se elabora una política de respaldo y recuperación para la seguridad del sistema. Para la correcta realización del CMI se seleccionan los indicadores correspondientes a las perspectivas propuestas en el capítulo anterior.

2.1 Caracterización del entorno

En el área de consulta externa se llevan a cabo un conjunto de procesos. Uno de ellos es el de atención al paciente que comienza cuando la persona previamente citada llega a la consulta. Independientemente del tipo de consulta el médico siempre interroga y examina para llegar a un diagnóstico y prescripción de un tratamiento. Con el interrogatorio se conoce el motivo de la consulta, los hábitos psicobiológicos que presenta el paciente y los antecedentes personales y familiares que pueden influir en el comportamiento de la patología presentada. La realización del examen funcional y el físico son otras de las actividades que complementan la atención del paciente.

Una vez que emite el diagnóstico presuntivo o definitivo concluye si el paciente necesita tratamiento. Antes de culminar la consulta el médico decide la conducta que debe seguir el paciente y crea los modelos correspondientes para garantizar la continuidad de la atención en el hospital o en otro centro asistencial.

Teniendo en cuenta la información que se gestiona se realiza la identificación de los temas de análisis. Esto significa agrupar todos los elementos relativos a un mismo objeto o evento. En el presente trabajo los temas de análisis identificados son: pacientes, consultas, referencias y contrarreferencias.

2.2 Reglas del negocio y transformación

Las reglas del negocio y transformación describen políticas, normas, operaciones, definiciones y restricciones especiales que se deben cumplir en algunas actividades de los procesos diarios de la entidad. Se definieron algunas reglas durante la etapa inicial de la investigación. Otro conjunto de reglas son identificadas producto de un análisis de perfilado de datos el cual es descrito en la sección 2.9.1.

Capítulo 2. Características y diseño del cuadro de mando integral

Dichas reglas constituyen la entrada fundamental para los procesos de diseño del almacén. Las reglas del negocio y transformación identificadas se describen a continuación:

RN1-Los identificadores de las medidas no pueden estar repetidos.

RN2-Los identificadores de los indicadores no pueden ser nulos.

RN3-Los identificadores de las dimensiones no pueden ser nulos.

RN4-Los valores que indiquen cantidad tienen que ser mayores o igual que cero.

RN5-Los servicios brindados en el área de consulta externa se definen tal y como se encuentra almacenado en el campo *nombre* de la tabla *servicio* correspondiente al esquema *comun* de la base de datos fuente.

RN6-La especialidad se define tal y como se encuentra almacenada en el campo *nombre* de la tabla *especialidad* correspondiente al esquema *comun* de la base de datos fuente.

RN7-El sexo estará definido de la siguiente manera: F para femenino y M para masculino.

RN8-El tipo de consulta se define: triaje general, primera, laboratorio, charla, episódica, sucesiva, interconsulta, pre-empleo, terminación, tratamiento fisioterapéutico y estudio radiológico.

RN9-El diagnóstico se define tal y como se encuentra almacenado en el campo *descripcion_enfermedad* de la tabla *diagnostico_medico_enfermedad* correspondiente al esquema *hc_local* de la base de datos fuente.

RN11-El médico se define según el valor contenido en el campo *nombre* que se encuentra almacenado en la tabla *usuario* correspondiente al esquema *comun* de la base de datos fuente y con minúscula excepto la letra inicial que irá con mayúscula.

RN12-El hospital se define según el valor contenido en el campo *nombre* que se encuentra almacenado en la tabla *entidad* correspondiente al esquema *comun* de la base de datos fuente.

RN14-El tipo de población se define de la siguiente manera: Población PDVSA, Población no PDVSA y Población Coromoto.

RN15-El tipo de paciente se define de la siguiente manera: beneficiario, titular, pasante, comunidad, jubilado y otros.

RN16-El tipo de referencia se define de la siguiente manera: internas y externas.

RN17-El tiempo se define día/mes/año.

2.3 Necesidades de los usuarios

Para realizar un análisis satisfactorio del negocio es preciso conocer las carencias de los clientes, así como la función de los mismos en el módulo Consulta Externa. Para ello se identificaron las necesidades de información que muestran los directivos del área de consulta externa, las cuales son el primer peldaño para el diseño del *DM*. Las mismas son la información sobre los pacientes, consultas, referencias y contrarreferencias que los directivos requieren para ver el desempeño de la entidad.

2.4 Especificación de requisitos

En la fase de requisitos se definen los requerimientos necesarios para hacer factible la solución propuesta. Dentro de los mismos se encuentran tanto los requisitos de información, como los funcionales y no funcionales, que de manera íntegra, satisfacen las necesidades del cliente. De esta correcta especificación depende la adecuada implementación.

Requisitos de información

Los requisitos de información dan respuesta a las necesidades del cliente. Los mismos son implementados posteriormente por los reportes y gráficos generados al consultar el CMI con sus indicadores claves de rendimiento, especificados en la sección 2.10.4. A continuación se muestran los requisitos de información más significativos, el resto se encuentra en el artefacto de especificación de requerimientos de software. El orden en que aparecen es condicionado por la consecutividad con que se describen algunos de ellos a lo largo del documento.

RI1. Obtener la cantidad de pacientes atendidos de un mes y año por servicio, según tipo de paciente.

RI2. Obtener la cantidad de pacientes atendidos de un mes y año por especialidad según el servicio.

RI3. Obtener la cantidad de pacientes atendidos de un mes y año según la especialidad, tipo de paciente, y el servicio.

RI4. Obtener la cantidad de pacientes atendidos en un mes y año por tipo de población según servicio, especialidad y sexo.

RI5. Obtener la cantidad de pacientes atendidos de un mes y año por especialidad, según el diagnóstico del paciente.

Capítulo 2. Características y diseño del cuadro de mando integral

RI13. Obtener la cantidad de pacientes atendidos de un mes y año por sexo, según especialidad.

RI18. Obtener el porcentaje de pacientes atendidos de un mes y año según tipo de paciente.

RI19. Obtener el promedio de pacientes atendidos de un mes y año según tipo de paciente.

Requisitos funcionales

Los requisitos funcionales se encargan de identificar y definir las metas o funciones que no deben faltar para la adecuada implementación del sistema. A continuación se muestran los requisitos funcionales utilizados en el presente trabajo:

RF1. Autenticar usuario.	RF8. Visualizar usuario.	RF15. Realizar transformación de los datos.
RF2. Adicionar rol.	RF9. Modificar usuario.	RF16. Cargar los datos al mercado de datos Consulta Externa.
RF3. Eliminar rol.	RF10. Adicionar reporte.	
RF4. Visualizar rol.	RF11. Visualizar reporte.	RF17. Personalizar reporte.
RF5. Modificar rol.	RF12. Eliminar reporte	RF18. Exportar a otro formato.
RF6. Adicionar usuario.	RF13. Modificar reporte.	
RF7. Eliminar usuario.	RF14. Extraer datos de la fuente.	

Requisitos no funcionales

Los requisitos no funcionales son requisitos que imponen restricciones en el diseño o la implementación. Se relacionan con atributos de calidad sistémica que conllevan al buen funcionamiento del sistema como un todo. A continuación se relacionan algunos de ellos:

RNF1. Cumplir con las pautas de diseño de las interfaces.

El sistema debe tener una interfaz gráfica uniforme incluyendo pantallas, menús y opciones. Las pautas de diseño se realizarán siguiendo la arquitectura de información definida.

RNF2. Mostrar los mensajes, títulos y demás textos que aparezcan en la interfaz del sistema en idioma español.

Los títulos de los componentes visuales, los mensajes para interactuar con los usuarios y los mensajes de error, deben ser en idioma español y tener una apariencia uniforme en todo el sistema. Los mensajes de error deberán ser explícitos para dar a conocer la gravedad del error.

RNF7. Asegurar la disponibilidad del sistema.

El sistema debe estar disponible durante el horario de trabajo. En caso de fallo, la recuperación del servicio no deberá ser de un período de tiempo mayor de cuarenta y ocho horas.

RNF9. Garantizar la persistencia de la información.

Para asegurar la persistencia de la información se ejecutará un respaldo total de los datos del mercado de datos con una frecuencia semestral.

2.5 Casos de uso del sistema

Los casos de uso constituyen la interacción entre los actores del sistema y sus respectivas actividades. Los requisitos de información y los requisitos funcionales son agrupados en casos de uso de información y casos de uso funcionales respectivamente. Luego los casos de uso del sistema están compuestos por casos de uso funcionales y de información. En las próximas secciones se describen cada uno de los tipos de casos de uso. Los actores que intervienen en dichos casos de uso se especifican en la tabla 2.

Tabla 2: Actores del sistema

Actor	Objetivo
Analista	Responsable de la visualización de los reportes del mercado de datos Consulta Externa.
Administrador	Responsable de gestionar los usuarios y roles del sistema, así como los reportes a visualizar.
Administrador ETL	Responsable de realizar los procesos de extracción, transformación y carga de los datos del sistema fuente.

Capítulo 2. Características y diseño del cuadro de mando integral

El diagrama de casos de uso del sistema es una representación gráfica de todos los actores del sistema y el flujo de eventos que los mismos guían, es decir, la relación entre los actores y los casos de uso. Para un mejor modelado del diagrama se utilizan los patrones de casos de uso tales como: “múltiples actores” mediante la generalización/especialización entre el Analista y el Administrador, “concordancia adición” mostrado en el caso de uso Realizar operaciones sobre reportes, y el patrón “CRUD completo” (*Create Read Update Delete*) evidenciado en los casos de uso Gestionar rol, Gestionar usuario y Gestionar reportes (Övergaard & Palmkvist, 2004). La figura 2 muestra el diagrama correspondiente.

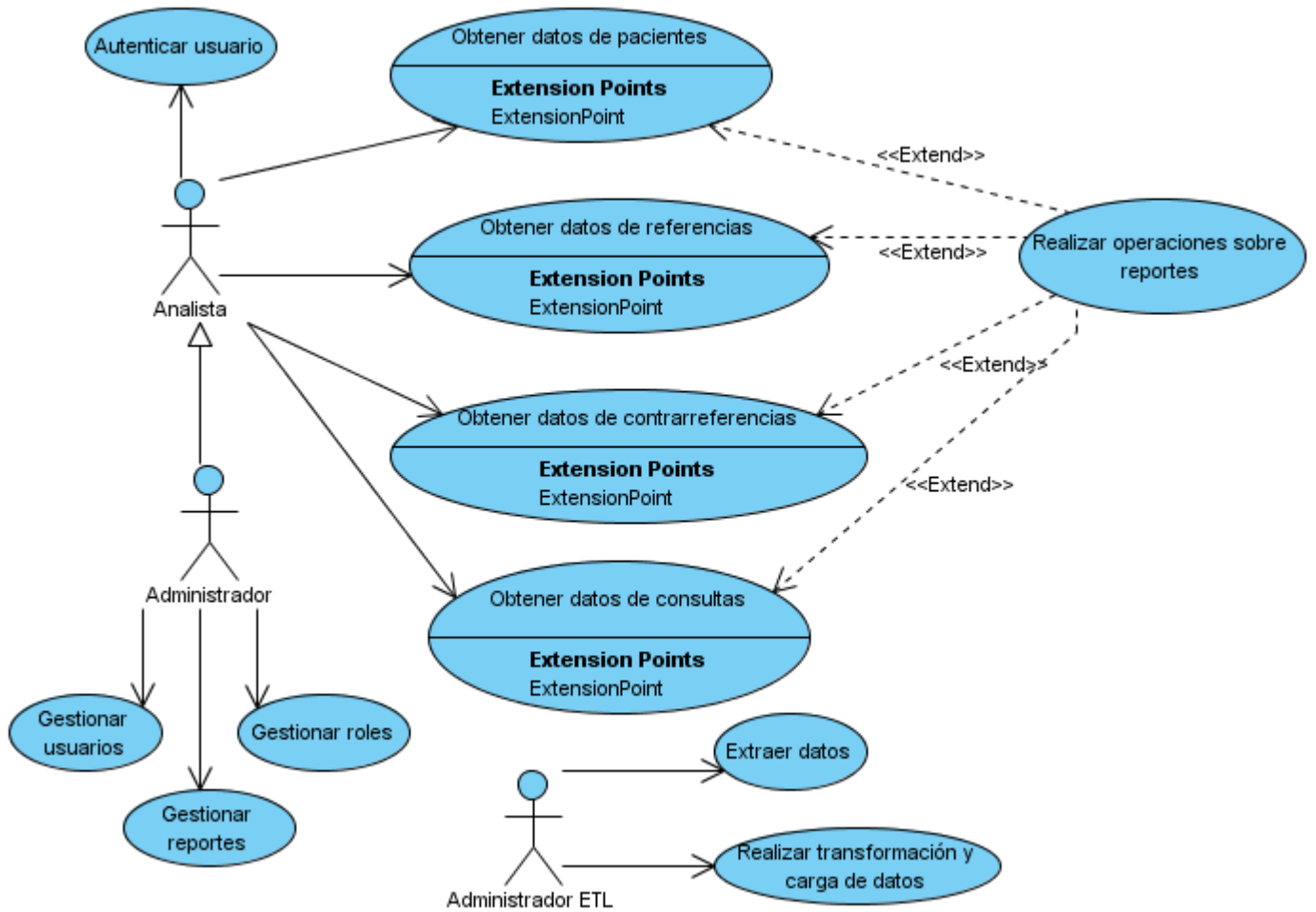


Figura 2: Diagrama de CUS
Fuente: Elaboración propia

2.5.1 Casos de uso de información

Los casos de uso de información son la agrupación de los requisitos de información por temas de análisis para tener la información organizada. A continuación se relacionan los casos de uso de información (CUI) definidos:

CUI1. Obtener datos de pacientes: visualiza los reportes de las medidas seleccionados de pacientes.

CUI2. Obtener datos de referencias: visualiza los reportes de las medidas seleccionados de referencia.

CUI3. Obtener datos de contrarreferencias: visualiza los reportes de las medidas seleccionados de contrarreferencias.

CUI4. Obtener datos de consultas: visualiza los reportes de las medidas seleccionados de consulta.

A continuación se muestra en la tabla 3 la descripción textual del CUI Obtener datos de pacientes. El resto de las descripciones textuales se encuentran en el artefacto de especificación de casos de uso dentro del Expediente de Proyecto.

Tabla 3: Descripción textual CUI1. Obtener datos de pacientes

Caso de Uso:	Obtener datos de pacientes.	
Tipo:	Caso de Uso de Información.	
Actores:	Analista	
Resumen:	El CUI inicia cuando el analista selecciona visualizar información de pacientes, luego se realiza la selección de la información a mostrar, con el objetivo que la misma sea visualizada por el usuario. El CUI finaliza cuando se muestran los resultados.	
Precondiciones:	El analista debe estar autenticado.	
Referencias	RI1 - RI19	
Prioridad	Media	
Flujo Normal de Eventos		
Sección “Obtener datos de pacientes”		
Acción del Actor	Respuesta del Negocio	
1. El analista selecciona el Área de	2. Muestra las Áreas de Análisis (A.A)	

Capítulo 2. Características y diseño del cuadro de mando integral

Análisis General (A.A.G) xavia HIS.	que están contenidas dentro del A.A.G xavia HIS.
3. El analista selecciona el A.A Consulta Externa.	4. Muestra los libros de trabajo que están contenidos dentro del A.A Consulta Externa.
5. El analista selecciona el Libro de Trabajo (L.T) Pacientes.	6. Muestra los reportes contendidos dentro del L.T Pacientes.
7. El analista selecciona el reporte deseado.	8. Muestra la información contenida en el reporte seleccionado y brinda la posibilidad al actor de hacerle cambios al reporte para su análisis. Ir al caso de uso funcional Realizar operaciones sobre los reportes. Finaliza el CUI.

Opciones de reportes de Pacientes

Entradas	Posibles resultados	
	Salidas	Periodicidad
Variables de entrada relacionadas con el caso de uso Obtener datos de pacientes. <ul style="list-style-type: none"> ✓ Especialidad ✓ Servicio ✓ Sexo ✓ Diagnóstico ✓ Hospital ✓ Tipo de paciente ✓ Tipo de población 	Variables de salida disponibles en el hecho Pacientes. <ul style="list-style-type: none"> ✓ Cantidad de pacientes ✓ Promedio de pacientes ✓ Porciento de pacientes 	Rango de tiempo en que se solicitan las variables de salida: <ul style="list-style-type: none"> ✓ Semestral

✓ Tiempo		
Pos condiciones	Disponibilidad de opciones (cruce de variables) de reportes relacionados con el CUI Obtener datos de pacientes.	

2.5.2 Casos de uso funcionales

Los casos de uso funcionales están relacionados básicamente con la administración de usuarios, roles y reportes visualizados, la autenticación de usuarios y la realización de las consultas en la base de datos. También se definen requisitos vinculados a los procesos de *ETL*. A continuación se relacionan los casos de uso funcionales (CUF) definidos:

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

CUF 2. Gestionar usuario: se insertan, modifican o eliminan los usuarios que interactúan con el sistema, asignando los roles correspondientes a los usuarios.

CUF 3. Gestionar rol: se registran, modifican o eliminan roles en el sistema.

CUF 4. Gestionar reporte: se eliminan, insertan o modifican los reportes que se visualizan.

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

CUF 6. Extraer datos: se realiza la extracción de los datos necesarios de las diferentes fuentes.

CUF 7. Realizar la transformación y carga de los datos: se realiza la transformación y carga de los datos necesarios para la construcción del mercado de datos Consulta Externa.

A continuación se muestra la descripción textual del CUF2 Gestionar usuario. El resto de las descripciones textuales se encuentran en el artefacto especificación de casos de uso dentro del Expediente de Proyecto.

Tabla 4: Descripción textual CUF2. Autenticar usuario

Caso de Uso:	Autenticar usuario.
Actores:	Analista, Administrador.
Resumen:	El CUF inicia cuando el actor se conecta al sistema para acceder a las funcionalidades del mismo. Para ello debe proporcionar un usuario y

Capítulo 2. Características y diseño del cuadro de mando integral

	contraseña. El caso de uso termina -una vez validado el usuario y contraseña cuando el usuario logra acceder al sistema.
Precondiciones:	El sistema debe estar disponible. El usuario debe existir en la base de datos.
Referencias	RF1
Prioridad	Media.
Flujo Normal de Eventos	
Acción del Actor	Respuesta del Sistema
1. El usuario interactúa con el sistema.	2. El sistema muestra un formulario para introducir los datos.
3. El usuario introduce los datos.	4. El sistema valida los datos.
	5. El sistema le da los permisos para entrar a la aplicación. Finaliza el caso de uso.
Flujos Alternos	
Acción del Actor	Respuesta del Sistema
	5.1 Si los datos no son válidos, muestra un mensaje de error y regresa al punto 2 del Flujo Normal de Eventos . Finaliza el caso de uso.
Poscondiciones	El usuario ha sido autenticado en el sistema.

2.6 Arquitectura

La arquitectura de software según Pressman (Pressman, 2005): “... proporciona una visión global del sistema a construir. Describe la estructura y la organización de los componentes del software, sus propiedades y las conexiones entre ellos. Los componentes del software incluyen módulos de programas y varias representaciones de datos que son manipulados por el programa”. En este trabajo la arquitectura de software se refiere a la forma en que se estructura el sistema propuesto teniendo en cuenta los requisitos y exigencias del usuario como se muestra en la figura 3.

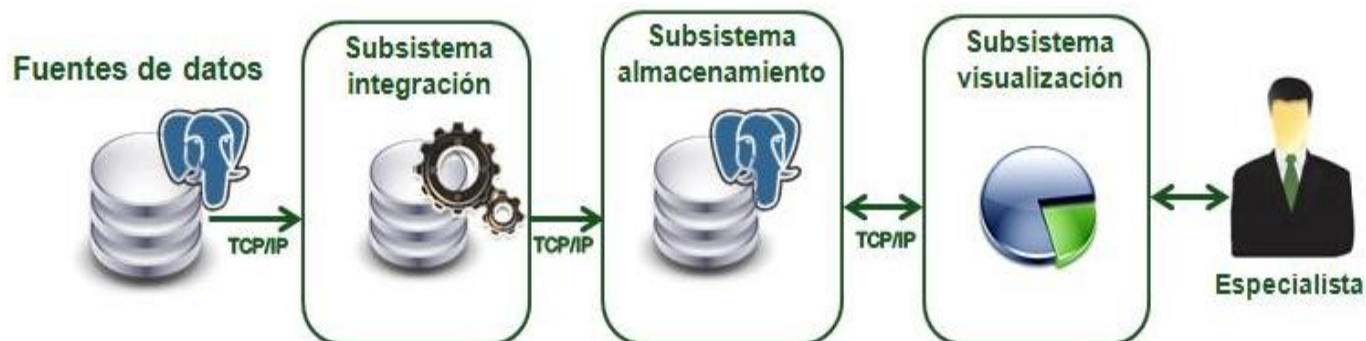


Figura 3: Diseño de la arquitectura
Fuente: Elaboración propia

Para llevar a cabo la construcción del sistema se debe tener presente la arquitectura lógica de este tipo de sistemas la cual queda de la siguiente forma:

- ✓ **Fuentes de datos:** se refiere al origen de los datos.
- ✓ **Subsistema de almacenamiento:** base de datos que contiene las tablas de hechos identificadas en el negocio, las dimensiones seleccionadas para la solución y las relaciones que existen entre estas.
- ✓ **Subsistema de integración:** incluye los procesos que permiten la carga de los datos de la fuente hacia el *DM*.
- ✓ **Subsistema de visualización:** comprende las interfaces orientadas a usuarios, facilitando una búsqueda rápida de la información para la toma de decisiones tácticas y estratégicas.

El subsistema de integración se nutre de la fuente de datos, que es la base de datos utilizada por el sistema xavia HIS. Lleva a cabo los procesos que integran y transforman la información para su almacenamiento. Los encargados de la administración de dichos procesos son el administrador de *ETL* y el administrador de base de datos.

El subsistema de almacenamiento recoge la información manipulada durante los procesos de *ETL*, la cual se almacena en el mercado de datos Consulta Externa soportada por el SGBD PostgreSQL y administrada por los usuarios autorizados mediante la herramienta PgAdminIII.

Finalmente, el subsistema de visualización permite mostrar la información estandarizada y organizada en los reportes y gráficos por áreas de análisis a través de la herramienta Pentaho BI Server utilizando el

editor de cuadro de mando (*Community Dashboard Editor- CDE por sus siglas en inglés*). El administrador y el analista son los autorizados a la visualización de los reportes.

Diseño de la solución

Se realiza el diseño conformado por los tres subsistemas: integración, almacenamiento y visualización. Los subsistemas de almacenamiento e integración están orientados fundamentalmente al mercado de datos, mientras que el subsistema de visualización se enfoca en la herramienta de análisis CMI.

2.7 Diseño del subsistema de almacenamiento

Para realizar el diseño del subsistema de almacenamiento se crea el modelo dimensional el cual contiene las tablas de hechos identificadas en el negocio con sus medidas asociadas. Las mismas pueden estar constituidas por jerarquías o niveles, que permiten detallar la información de forma eficiente. Se debe definir, además, una política de respaldo y recuperación que garantice la integridad de los datos almacenados. Se define la convención de nombres para las tablas de hechos y dimensiones.

2.7.1 Dimensiones

En el diseño del subsistema de almacenamiento se identifican las dimensiones que se detallan a continuación. Los nombres de las tablas dimensiones llevan el sufijo *dim* seguido del guión bajo.

Dimensión médico (*dim_medico*): describe el universo de valores bajo los cuales puede clasificarse la información atendiendo al médico que asiste a un paciente.

Dimensión especialidad (*dim_especialidad*): describe el universo de valores bajo los cuales puede clasificarse la información atendiendo a la especialidad.

Dimensión servicio (*dim_servicio*): describe el universo de valores bajo los cuales puede clasificarse la información atendiendo al servicio.

Dimensión sexo (*dim_sexo*): describe el universo de valores bajo los cuales puede clasificarse la información atendiendo al sexo.

Dimensión hospital (*dim_hospital*): describe el universo de valores bajo los cuales puede clasificarse la información atendiendo al hospital.

Dimensión tipo de población (*dim_tipo_poblacion*): describe el universo de valores bajo los cuales puede clasificarse la información atendiendo al tipo de población.

Dimensión tipo de referencia (*dim_tipo_referencia*): describe el universo de valores bajo los cuales puede clasificarse la información atendiendo al tipo de referencia.

Dimensión tipo de paciente (*dim_estado_pac*): describe el universo de valores bajo los cuales puede clasificarse la información atendiendo al tipo del paciente.

Dimensión diagnóstico (*dim_diagnostico*): describe el universo de valores bajo los cuales puede clasificarse la información atendiendo al diagnóstico.

Dimensión tipo de consulta (*dim_tipo_consulta*): describe el universo de valores bajo los cuales puede clasificarse la información atendiendo al tipo de consulta.

Dimensión tiempo (*dim_tiempo*): esta dimensión es una de las más comunes e importantes en el diseño de mercados de datos, debido a que define una línea de tiempo para enmarcar la información almacenada y la organiza atendiendo al momento en que fue captada.

Se define una jerarquía Fecha Anno-Mes-Día que contiene tres niveles: el nivel año, el nivel mes y el nivel día. Esto posibilita ver la información en detalle.

2.7.2 Tablas de hechos

Para el desarrollo del sistema se identifican cuatro tablas de hechos:

- ✓ **Hecho pacientes** (*hech_paciente*): recoge los datos relacionados con la medida numérica *cant_pacientes*, asociada al valor real de la cantidad de pacientes atendidos durante un mes o año.
- ✓ **Hecho consultas** (*hech_consulta*): recoge los datos relacionados con la medida numérica *cant_consultas*, asociada al valor real de la cantidad de consultas realizadas durante un mes o año.
- ✓ **Hecho referencias** (*hech_referencia*): recoge los datos relacionados con la medida numérica *cant_referencia*, asociada al valor real de la cantidad de referencias emitidas durante un mes o año.
- ✓ **Hecho contrarreferencias** (*hech_crreferencia*): recoge los datos relacionados con la medida numérica *cant_contrarreferencias*, asociada al valor real de la cantidad de contrarreferencias emitidas durante un mes o año.

2.7.3 Matriz bus o matriz dimensional

En la matriz bus o matriz dimensional se representa la información como matrices multidimensionales o cuadros de múltiples entradas denominados cubos. Representa la relación que existe entre los hechos y las dimensiones. La realización de la matriz dimensional permite conocer que de los cuatro hechos definidos para el modelo de datos a diseñar, todos ellos comparten dimensiones, sin embargo, no existen dos o más hechos que se relacionen con exactamente las mismas dimensiones. Esto demuestra la inexistencia de solapamiento de hechos. La tabla 5 muestra la matriz realizada.

Tabla 5: Matriz BUS

Dimensiones	Hechos			
	Pacientes	Referencias	Contrarreferencias	Consultas
Servicio	x	x	x	x
Especialidad	x	x	x	x
Tipo de paciente	x			
Diagnóstico	x			
Sexo	x	x		
Médico		x	x	x
Hospital	x	x		
Tiempo	x	x	x	x
Tipo población	x			
Tipo referencia		x		
Tipo consulta				x

2.7.4 Modelo de datos

Luego de definidas las dimensiones, hechos y medidas se procede a realizar el modelo de datos. A continuación la figura 4 muestra una parte del modelo de datos diseñado para el desarrollo del sistema propuesto. El modelo sigue una topología constelación de hechos, puesto que las tablas de hechos comparten algunas de sus dimensiones.

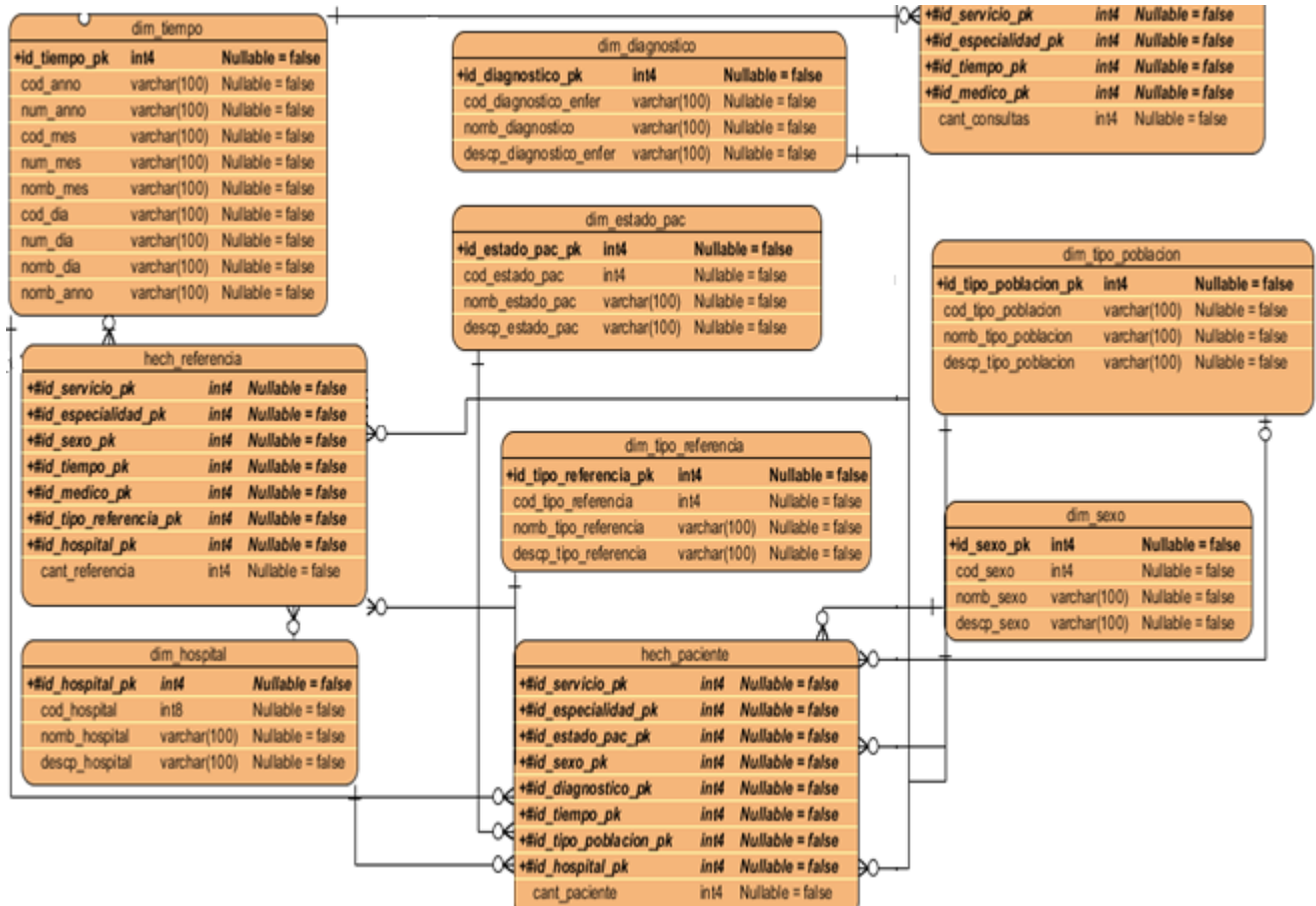


Figura 4: Modelo de datos
Fuente: Elaboración propia

2.7.5 Estructura física de almacenamiento

Se definen cuatro esquemas que conforman la estructura de almacenamiento del mercado de datos Consulta Externa. A continuación se relacionan:

El esquema “nomencladores” y “fecha” contiene las dimensiones que pueden ser compartidas con el resto de los mercados de datos. El esquema “consulta” recoge las tablas de dimensiones y hechos propias del mercado de datos Consulta Externa. El esquema “metadatos” se define con el objetivo de realizar el control de cambios, donde se ha diseñado una estructura basada en el uso de cuatro tablas de metadatos.

2.7.6 Política de respaldo y recuperación

Con el objetivo de garantizar la persistencia de la información, se establece una política de respaldo y recuperación que incluye la utilización de respaldos (*backups*) (Hosting diario, 2013). De los distintos tipos de backups existentes se selecciona la utilización de dos de ellos:

Respaldo completo: Se hace un respaldo completo de todas las carpetas y archivos que recogen la información del sistema propuesto dos veces al año. El respaldo abarca el 100% de los datos, por lo que suele ser el que lleva más tiempo en realizarse.

Respaldo incremental: Una vez que se realicen cambios en el sistema propuesto, se realiza un respaldo de todos los archivos que han sido modificados desde que fue ejecutado el último respaldo completo o incremental, siendo el método más rápido para realizar respaldos.

2.8 Diseño del subsistema de integración

El subsistema de integración incluye inicialmente un análisis de perfilado de los datos, la extracción de los mismos desde sus respectivas fuentes, así como una serie de transformaciones basadas en las reglas del negocio y transformación descritas en el epígrafe 2.3. Finalmente devuelve los datos preprocesados y listos para el almacenamiento, de forma tal, que permita a los desarrolladores construir la capa de presentación de la aplicación.

2.8.1 Perfilado de los datos

Perfilado de datos se refiere a un análisis de los datos iniciales para comprender mejor su contenido, estructura, calidad y dependencias. El mismo permite conocer la calidad de los datos iniciales, así como gestionar y controlar el estado de los mismos. Además, reduce los errores de la información que será cargada en el mercado.

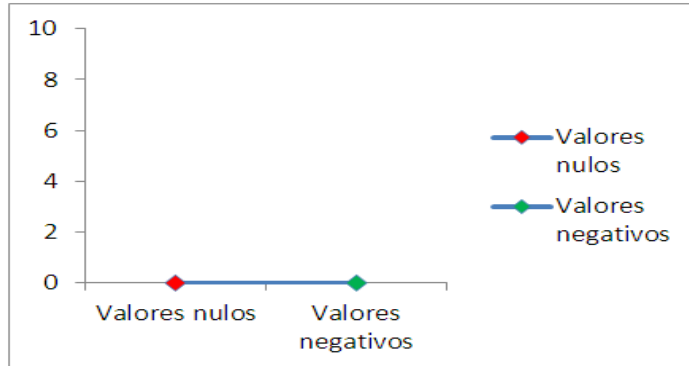


Figura 5: Valores nulos o negativos
Fuente: Elaboración propia

Los datos iniciales son extraídos de varias tablas utilizadas para el funcionamiento del módulo Consulta Externa. De manera general se observa de las 14076 tuplas contenidas en la fuente de datos, la inexistencia de valores nulos o negativos como se muestra en la figura 5, sin embargo existen 11749 valores duplicados como se muestra en la figura 6.

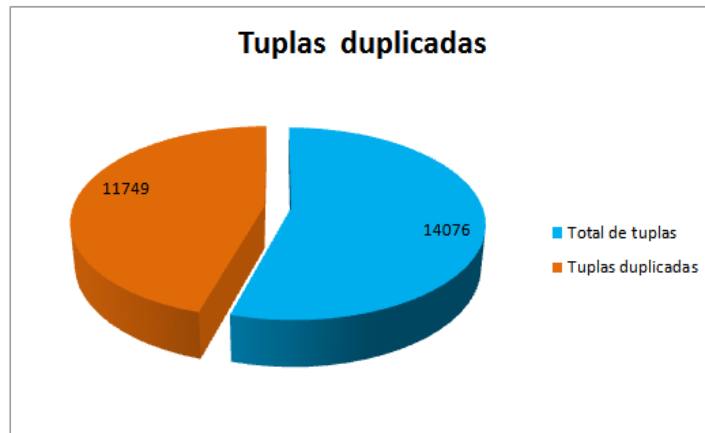


Figura 6: Tuplas duplicadas
Fuente: Elaboración propia

También se identifican los tipos de datos contenidos en los diferentes campos de algunas tablas que conforman los esquemas: *comun*, *publico*, *nomencladores*, *hc* y *hc_local* como se muestra en la figura 7.

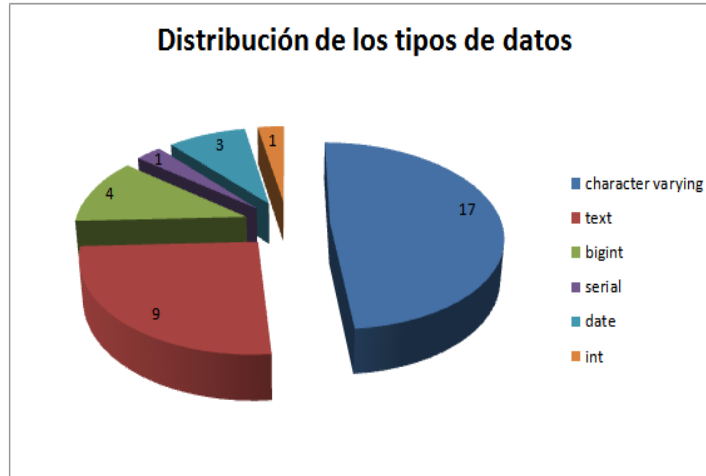


Figura 7: Tipos de datos
Fuente: Elaboración propia

Según lo planteado por Kimball, para garantizar la calidad de los datos los mismos deben ser (Kimball & Caserta, 2004):

- ✓ **Correctos:** los valores y descripciones deben ser correctos y precisos.
- ✓ **Inequívocos:** los valores y descripciones deben ser únicos, es decir, deben tener un solo significado.
- ✓ **Coherentes:** los valores y descripciones de los datos deben tener una nomenclatura invariable.
- ✓ **Completo:** los valores y las descripciones deben ser definidos para cada caso. Además, todos los datos deben ser cargados.

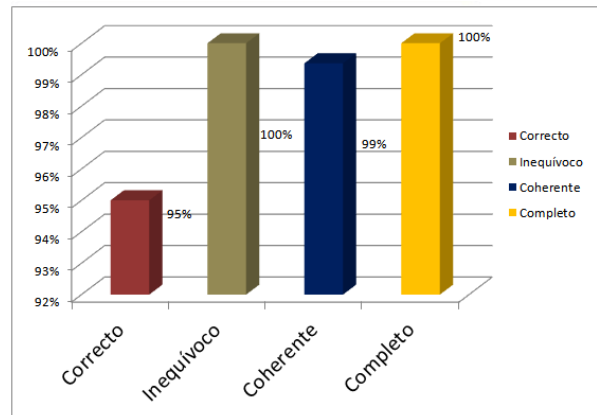


Figura 8: Perfilado de datos
Fuente: Elaboración propia

Aplicando los criterios anteriores se obtuvo como resultado 95% de datos correctos, 100% inequívocos y completos, además de un 99% coherentes, como se muestra en la figura 8.

2.8.2 Diseño de los procesos de integración

Luego de conocidas la estructura, contenido y fiabilidad de los datos, se procede a diseñar los procesos de integración que se corresponde con los procesos de *ETL* que también suelen llamarse transformaciones. Lo anterior no asegura que los diseños no puedan modificarse a la hora de la implementación de las transformaciones.

Para cargar la información hacia el mercado de datos Consulta Externa se accede a la base de datos fuente, de donde son extraídos los campos correspondientes a los datos que quedarán registrados en las dimensiones con que se relaciona cada hecho y se procede a realizar las transformaciones adecuadas.

Primeramente se extraen los datos de la fuente y a su vez se ordenan para un mejor análisis, seleccionando los registros y campos relevantes para el mercado de datos Consulta Externa, teniendo en cuenta el modelo de datos realizado. La etapa de transformación y limpieza entrega la información lista para ser cargada en el mercado. Mediante la validación de datos se detectan las incongruencias, tales como datos con el mismo significado y diferente nomenclatura, datos con el mismo nombre y diferente código, entre otras. Además, se detectan las entradas duplicadas que son los datos que se repiten en la fuente y no deben ser cargados hacia el mercado.

Mediante la carga de las tablas de dimensiones y hechos los datos son transmitidos hacia el mercado de datos Consulta Externa. La información cargada es depurada y utilizada para la generación de los reportes que son analizados por especialistas. La figura 9 muestra los pasos anteriormente descritos.



Figura 9: Diseño de la transformación para la carga de las dimensiones
Fuente: Elaboración propia

Para el diseño de las transformaciones de las tablas de hechos primeramente se realiza la extracción de los datos y a su vez el ordenamiento y la validación de los mismos. Luego se efectúa una búsqueda de las llaves primarias de las dimensiones siendo posteriormente agrupadas con el objetivo de contar por todas las llaves. Después se validan las llaves, si no poseen errores son insertadas en el mercado de datos Consulta Externa, de lo contrario se envían a un excel de errores para su posterior análisis. La figura 10 muestra el diseño de las transformaciones para las tablas de hechos.

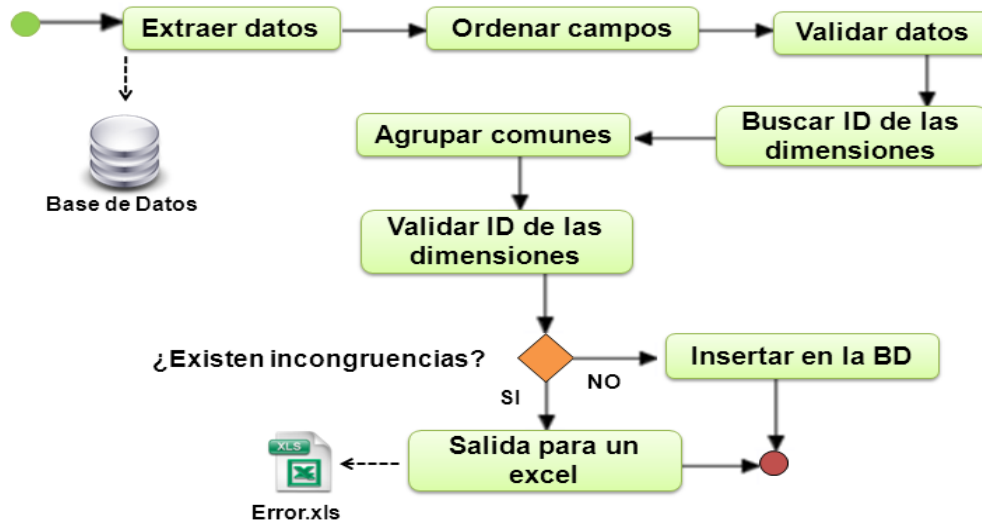


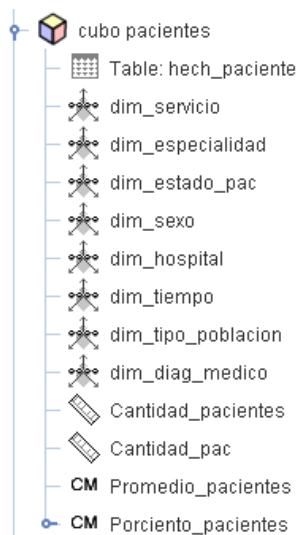
Figura 10: Diseño de la transformación para las tablas de hechos
Fuente: Elaboración propia

2.9 Diseño del subsistema de visualización

En el subsistema de visualización es necesario diseñar los reportes candidatos y los cubos *OLAP*. Luego el CMI tomará dichos reportes como entrada para mostrar la información consultada. Un reporte *OLAP* permite diseñar y simular un esquema de cubos, a través del mapeo de los datos dimensionales y permite crear reportes en línea haciendo uso de dimensiones y medidas.

2.9.1 Diseño de los cubos multidimensionales

El mercado de datos Consulta Externa posee cuatro cubos multidimensionales, correspondientes a las tablas de hechos definidas y once dimensiones. Además, de una medida aditiva: cantidad y dos no aditivas: promedio y porcentaje para cada cubo. La figura 11 muestra como ejemplo el cubo pacientes con las dimensiones y medidas asociadas.



Las dimensiones son utilizables y permiten ser aprovechadas en cubos diferentes distribuidas como sigue:

El cubo **Pacientes** contiene ocho dimensiones utilizables.

El cubo **Consultas** contiene cinco dimensiones utilizables.

El cubo **Referencias** contiene siete dimensiones utilizables.

El cubo **Contrarreferencias** contiene cuatro dimensiones utilizables.

Figura 11: Cubo pacientes

Fuente: Elaboración propia

2.9.2 Diseño de la arquitectura de información

La definición de las Áreas de Análisis (A.A) constituye un paso importante dentro del diseño del subsistema de visualización. La solución propuesta se realiza en función de la información que frecuentemente utiliza el módulo Consulta Externa. En tal sentido se define el A.A Consulta Externa, la cual se encuentra contenida dentro del Área de Análisis General (A.A.G) Sistema de Información Hospitalaria xavia HIS (A.A.G xavia HIS). Dentro de ella se agrupa la información organizada en libros de trabajo, los cuales son directorios que contienen los reportes agrupados por temas de análisis con el objetivo de lograr una mejor organización de los reportes. Incluye, además, los L.T que contienen los reportes y gráficos dinámicos del CMI agrupados también por temas de análisis.

La figura 12 muestra el mapa de navegación de la capa de presentación de la aplicación, que constituye una representación del diseño definido para la arquitectura de información del sistema.



Figura 12: Diseño de la arquitectura de información
Fuente: Elaboración propia

2.9.3 Diseño de reportes candidatos

Los reportes correspondientes a dicha A.A se encuentran distribuidos dentro de los libros de trabajos de la siguiente manera:

L.T Pacientes: libro de trabajo contenido dentro del A.A Consulta Externa. Contiene diecinueve reportes que permiten realizar un análisis general de los datos y se realizan en función de los indicadores cantidad, promedio y porcentaje de pacientes.

L.T Consultas: libro de trabajo contenido dentro del A.A Consulta Externa. Contiene once reportes que permiten realizar un análisis general de los datos y se realizan en función de los indicadores cantidad, promedio y porcentaje de consultas.

L.T Referencias: libro de trabajo contenido dentro del A.A Consulta Externa. Contiene dieciséis reportes que permiten realizar un análisis general de los datos y se realizan en función de los indicadores cantidad, promedio y porcentaje de referencias.

L.T Contrarreferencias: libro de trabajo contenido dentro del A.A Consulta Externa. Contiene seis reportes que permiten realizar un análisis general de los datos y se realizan en función de los indicadores cantidad, promedio y porcentaje de contrarreferencias.

2.9.4 Indicadores del cuadro de mando integral

A partir de las perspectivas definidas en el capítulo anterior en la sección 1.3 se seleccionan los indicadores claves de rendimiento (*Key Performance Indicators - KPI por sus siglas en inglés*) (Buenas Tareas, 2010) para cada una de ellas. Los indicadores son descripciones de un determinado análisis que pueden ser cuantitativos o cualitativos. Permiten, además, las comparaciones en el tiempo, con otros indicadores y la vinculación de los mismos a través de relaciones causa-efecto como se muestra en la figura 13.

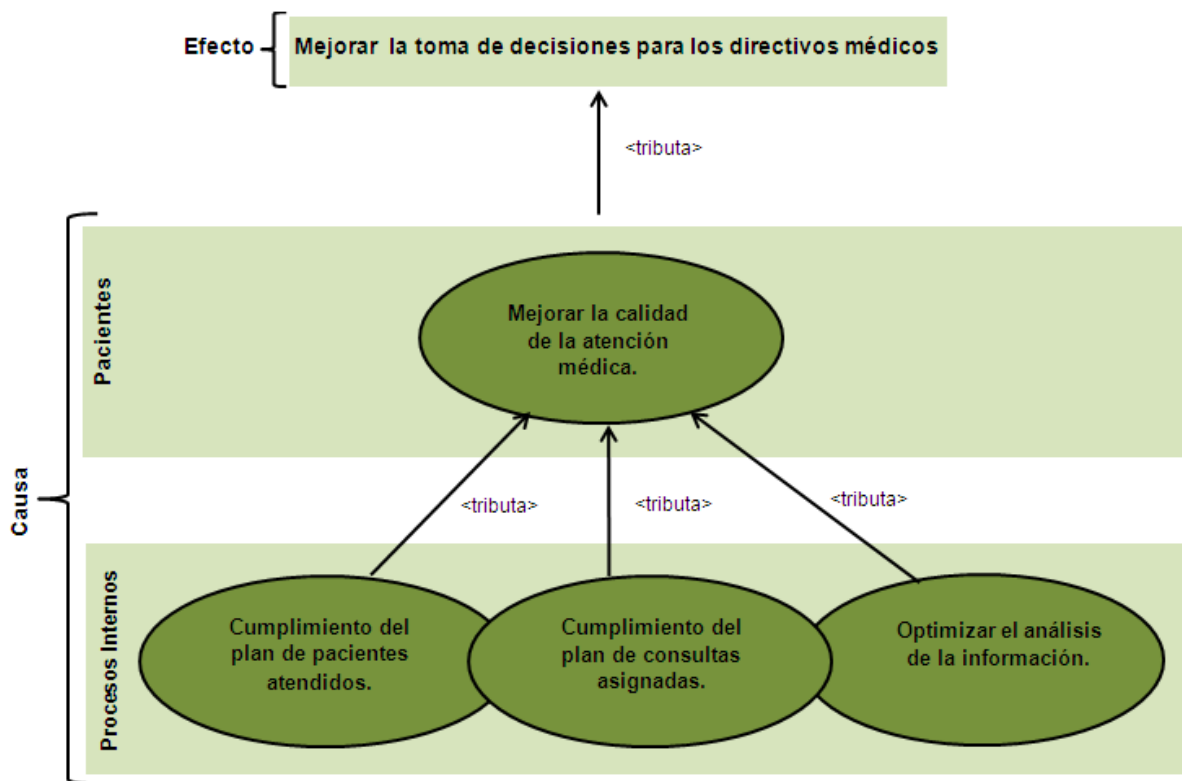


Figura 13: Relación causa efecto entre indicadores
Fuente: Elaboración propia

Con el objetivo de mejorar la toma de decisiones para los directivos médicos se crea una relación causa-efecto entre los indicadores cualitativos y cuantitativos de las perspectivas seleccionadas, permitiendo mantener un seguimiento y control de los objetivos propuestos por las entidades de salud.

2.10 Esquema de seguridad

Para el acceso al sistema propuesto se ha establecido un usuario por cada uno de los roles existentes en el sistema, con el objetivo de garantizar un control de acceso basado en roles (Vogelmann , 2008). De esta manera, cada usuario opera en el sistema según los permisos que se le definan al rol correspondiente. Para un mejor entendimiento del esquema de seguridad llevado a cabo en el sistema se realiza un desglose a continuación por cada uno de los subsistemas identificados.

Seguridad en el subsistema de almacenamiento

Tabla 6: Seguridad en la BD

Roles	Permisos
Administrador BD	Realiza la administración del mercado de datos. Tiene todos los permisos sobre la BD. Otorga permisos a nivel de usuarios.
Administrador ETL	Realiza los procesos de <i>ETL</i> sobre los datos y tiene permisos de lectura y escritura sobre los esquemas pertenecientes al mercado de datos Consulta Externa.

Seguridad en el subsistema de integración

La seguridad en el subsistema de integración se garantiza por parte del sistema operativo que en el caso del mercado de datos Consulta Externa se trata de GNU/Linux. Dicho sistema operativo permite otorgar permisos a los archivos de configuración y ejecución para los usuarios avanzados. De esta manera, se puede restringir o permitir el acceso de un determinado usuario a un archivo para su visualización de contenidos, modificación y ejecución (en caso de un archivo ejecutable). Esta ventaja es utilizada con el objetivo de restringir el acceso por parte de usuarios no autorizados a los archivos que contienen las transformaciones y trabajos que permiten el desarrollo de los procesos de integración en el mercado de datos Consulta Externa.

Seguridad en el subsistema de visualización

Tabla 7: Seguridad en la aplicación

Roles	Permisos
Analista	Tiene acceso de solo lectura al A.A Consulta Externa. Puede visualizar todos los reportes de esta área.
Administrador	Tiene acceso total para gestionar los permisos, roles y usuarios, así como los reportes contenidos en el A.A Consulta Externa.

Para garantizar la integridad de los datos se utilizan, además, las cuatro áreas esenciales que proporciona la plataforma Pentaho BI.Server. Las mismas se explican a continuación:

- ✓ **Seguridad de acceso a datos de objetos:** incluye usuarios, contraseñas, autorizaciones permitidas, recursos web y protección a datos almacenados en el mercado.
- ✓ **Autenticación:** esta área 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:** ofrece protección a los localizadores de recursos uniformes (*Uniform Resource Locator - URL por sus siglas en inglés*) para responder a cada usuario si pueden o no acceder a una determinada página. Esto es decidido por el administrador de recursos web, el cual le brinda 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.

Además, para acceder al sistema se realiza a través del directorio activo de la UCI para garantizar un acceso centralizado y credenciales únicas. Para ello se utiliza el Servicio de Autenticación Central (CAS - *Central Authentication Services por sus siglas en inglés*). Los usuarios son administrados a través de la consola de administración del Pentaho BI Server.

Conclusiones del capítulo

Al describir las principales características del sistema propuesto, se puede concluir que el perfilado de datos realizado a la fuente de información permitió obtener una noción de su estado, así como el establecimiento de nuevas reglas del negocio y transformación aplicables durante el proceso de transformación. Además, las once tablas dimensionales y las cuatro tablas de hechos identificadas permitieron el diseño del modelo de datos para lograr el correcto funcionamiento del sistema. El diseño de las transformaciones para la carga de las dimensiones y los hechos constituye una aproximación a los pasos que se deben realizar para lograr la estandarización de la información y su almacenamiento.

Por su parte las políticas de recuperación y respaldo establecidas contribuyen a mantener la integridad de los datos almacenados. Con el diseño de los cubos multidimensionales para cada tabla de hecho y los reportes candidatos contenidos dentro del A.A Consulta Externa, se logra responder a las necesidades de información de los clientes identificadas durante el análisis de los procesos de negocio.

Capítulo 3: Implementación y validación del cuadro de mando integral

Una vez realizado el diseño multidimensional guiado por la propuesta de metodología utilizada, se procede a realizar la implementación física del sistema, lo que implica el desarrollo de cada uno de los subsistemas que conforman el sistema propuesto. Finalmente se exponen las pruebas realizadas al mismo para su validación, así como los resultados obtenidos en cada una de ellas luego de su aplicación para garantizar el buen funcionamiento del producto final.

3.1 Implementación del subsistema de almacenamiento

La implementación del subsistema de almacenamiento incluye la estandarización de los nombres. Luego de finalizado dicho proceso se define la nomenclatura a utilizar para la designación de las tablas, atributos y medidas dentro del mercado. Se procede entonces a la implementación de las estructuras físicas, la cual define la utilización de cuatro esquemas con diecinueve tablas, divididas en once tablas de dimensiones, cuatro tablas de hechos y cuatro tablas de metadatos. También contiene la estrategia de indexado descrita en el próximo epígrafe.

3.1.1 Indexación utilizada

En un *DM* una estrategia de indexado consiste en crear índices para las llaves primarias y foráneas. Esto se debe a que para la mayoría de las uniones entre tablas que se realizan, las columnas implicadas son precisamente las llaves y este tipo de operaciones es el que más tiempo de consulta consume. Para la búsqueda de datos mediante las llaves primarias y foráneas se utiliza el indexado *B-Tree*, que admite el gestor PostgreSQL, los cuales permiten la optimización de las consultas que se realizan mediante el uso de las llaves primarias. Las mismas son llaves subrogadas que presentan índices *B-Tree*. El indexado propuesto para las tablas de dimensiones es sobre la llave primaria de cada una de ellas.

3.2 Implementación del subsistema de integración

La implementación del subsistema de integración implica la realización de los procesos de extracción, transformación y carga. Antes de iniciar los procesos de integración de los datos se realiza un análisis previo de la fuente de datos. Dicha fuente es la base de datos relacional utilizada por el sistema xavia HIS, soportada por el SGDB PostgreSQL. De los subsistemas identificados por Kimball a continuación se describen algunos de los utilizados en el desarrollo de la solución:

Capítulo 3. Implementación y validación del cuadro de mando integral

- ✓ **Perfilado de datos:** permitió explorar los datos para verificar su calidad y el cumplimiento de los estándares conforme a los requerimientos especificados por el cliente. Mediante el uso de este subsistema fueron definidas nuevas reglas para las transformaciones.
- ✓ **Sistema de extracción:** permitió la extracción de los datos desde la fuente de origen para su transformación y posterior carga. Para ello se tuvo en cuenta la información relacionada con cada uno de los hechos y dimensiones.
- ✓ **Subsistema de transformación:** permitió realizar transformaciones como el mapeo de valores, el cambio de tipo de dato en algunos campos, la búsqueda de información en flujos de datos, el filtrado de valores, entre otras.
- ✓ **Subsistema de carga:** permitió realizar la carga de los datos a las tablas de dimensiones y hechos del mercado de datos Consulta Externa.
- ✓ **Captura de cambio en los datos:** permite detectar los cambios asociados a las fechas en que se realizará la carga de la información.
- ✓ **Rastreo de eventos de errores:** posibilita la captura de los errores que proporcionan información valiosa sobre la calidad de los datos y permiten la mejora de los mismos.
- ✓ **Cambio lento de las dimensiones:** implementa la lógica para crear atributos de variabilidad lenta a lo largo del tiempo. Fue utilizado el tipo dos de *SCD* para dar tratamiento al cambio de la información asociada a las dimensiones.
- ✓ **Llave subrogada:** permite crear claves subrogadas independientes para cada tabla.
- ✓ **Creador de cubos OLAP:** permite alimentar de datos a esquemas *OLAP* desde esquemas dimensionales relacionales.
- ✓ **Programador de trabajos:** permite gestionar los trabajos, los cuales se encargan de la ejecución de las transformaciones en un orden específico y atendiendo a la periodicidad definida para la carga de la información.
- ✓ **Control de versiones:** permite hacer control de versiones del proyecto *ETL* y de los metadatos asociados, lo cual permitirá conocer, entre otros elementos, cuando se realizó la última carga.

✓ **Repositorio de metadatos:** captura los metadatos de los procesos de *ETL* y de los aspectos técnicos.

3.2.1 Implementación de las transformaciones

Las transformaciones están conformadas por pasos, que constituyen el elemento más pequeño de la transformación y se encuentran acoplados a través de saltos. Una vez que se toma la decisión de qué reglas para las transformaciones serán establecidas, deben crearse e incluirse las definiciones en las rutinas de transformación.

Lo primero que se realiza es la conexión con la fuente de datos en el componente "Entrada Tabla", donde se escribe, además, una consulta *SQL* con el fin de obtener los campos relevantes para el mercado de datos Consulta Externa. Se efectúa también la limpieza y transformación de datos en la fecha para corregir su formato utilizando los componentes "Partir campo de fecha", "Partir campo día" y "Eliminar cero antes del día menor que diez". Después se ejecuta una búsqueda de las llaves primarias de las tablas de dimensiones con los componentes "Buscar" utilizado en todas las dimensiones, una vez encontradas se ordenan las llaves utilizando el "Ordenar campos" para facilitar su agrupación. El componente "Agrupar por id" cuenta las tuplas que contengan los mismos valores para todas las dimensiones.

Finalmente se validan con "Validar datos" para su posterior carga hacia la tabla de hechos del mercado de datos Consulta Externa haciendo uso del componente "cargar_cant_pacientes". Una vez completada la carga se gestionan los metadatos para el control de la transformación donde se captura la fecha, el nombre de la tabla del hecho y la dirección de protocolo de internet (*Internet Protocol – ip por sus siglas en inglés*) de la máquina donde se ejecutó la transformación. Estos datos serán guardados en la tabla de metadatos *gestion_carga_historica*.

De esta forma se realiza la extracción de los datos correspondientes a cada una de las tablas de hechos del mercado de datos Consulta Externa. La figura 14 muestra la transformación del hecho paciente.

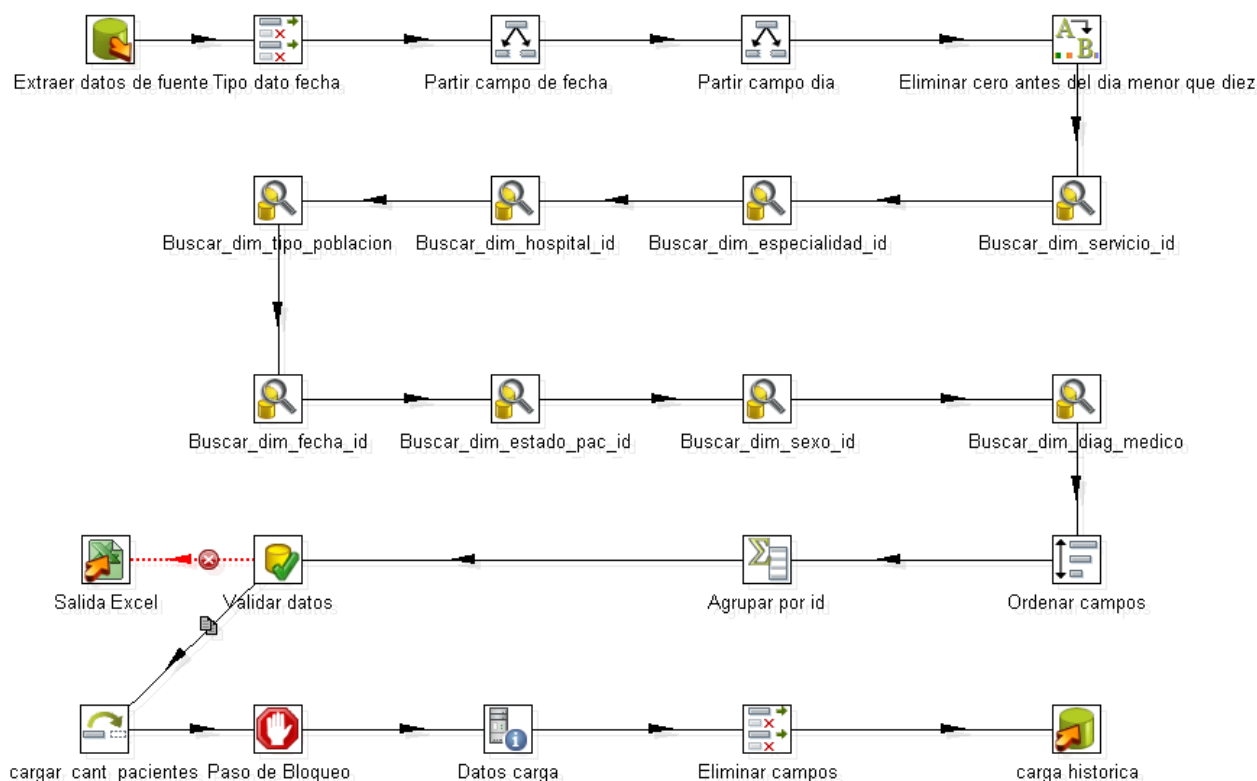


Figura 14: Transformación hecho pacientes
Fuente: Elaboración propia

Los datos que sean almacenados en el archivo excel error recibirán tratamiento luego de definir con el cliente si constituyen información relevante y cuáles reglas del negocio y transformación deben ser aplicadas para su carga al mercado de datos Consulta Externa.

3.2.2 Implementación de los metadatos

Los metadatos son la información referente a los datos, las técnicas que lo manipulan y otras estructuras como objetos y reglas de negocio. Existen tres tipos de metadatos (Kimball & Caserta, 2004) en dependencia de lo que describen: los metadatos de procesos, técnicos y de negocios. A continuación se explican los tipos de metadatos utilizados en la solución propuesta.

Metadatos de proceso

Los metadatos de procesos brindan estadísticas sobre los resultados de la ejecución del propio proceso de integración de datos, incluyendo aspectos tales como: filas cargadas con éxito, las filas rechazadas, y

el tiempo de carga. En el presente trabajo se utilizan tres tablas de metadatos de proceso, las mismas se muestran a continuación:

tb_metadatos_dim: almacena la información correspondiente a las acciones que se realizan sobre las tablas de dimensiones.

tb_metadatos_hech: almacena la información correspondiente a las acciones que se realizan sobre las tablas de hechos.

tb_metadatos_job: almacena la información correspondiente al resultado de la ejecución de los trabajos.

Metadatos técnicos

Los metadatos técnicos representan los aspectos técnicos de los datos. En el presente trabajo se utiliza una tabla de metadatos, mostrada a continuación:

gestion_carga_historica: almacena la fecha, el nombre de la tabla del hecho y el *ip* de la máquina donde se ejecutó la transformación.

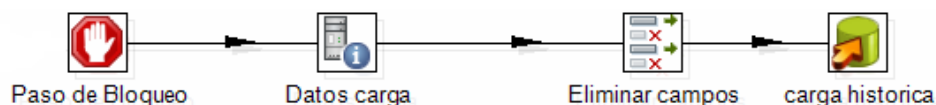


Figura 15: Metadato técnico gestión de la carga histórica

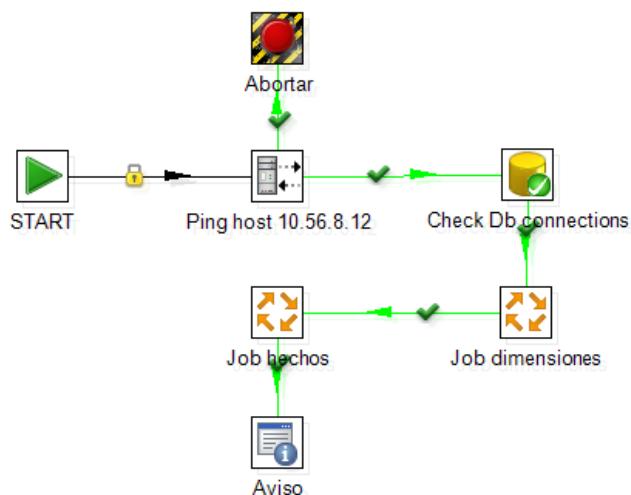
Fuente: Elaboración propia

Como muestra la figura 15 la implementación del metadato técnico contiene un componente llamado "Paso de Bloqueo", para esperar un tiempo luego de insertados los datos en la tabla de hechos. También se utiliza el componente "Datos carga" que permite seleccionar la información de interés sobre la carga. Otro componente es "Eliminar campos", que permite desechar los datos que no se desean guardar. Por último el componente "carga histórica" registra los datos seleccionados en la tabla de metadatos.

3.2.3 Implementación de los trabajos

En el contexto de integración de datos, el término trabajo o *job* se entiende como un conjunto sencillo o complejo de tareas cuyo objetivo consiste en realizar una acción determinada. La implementación de un trabajo define una secuencia lógica para la ejecución de las transformaciones, mediante el uso de pasos definidos, los cuales son diferentes a los disponibles en las transformaciones. Además, es posible ejecutar una o varias transformaciones de las que se hayan diseñado y guiar una secuencia de ejecución para ellas. Los trabajos se encuentran en un nivel superior a las transformaciones.

Para el sistema propuesto primeramente se implementó un trabajo que contiene todas las transformaciones de las dimensiones que serán cargadas en el mercado de datos Consulta Externa. Luego se implementó un trabajo que contiene todas las transformaciones de los hechos. Por último el trabajo general contiene los trabajos anteriores y es el encargado de ejecutarlos para realizar las cargas de las dimensiones y hechos automáticamente. La figura 16 muestra el trabajo general.



La implementación del trabajo general está compuesta por varios componentes. El primero es " STAR " que como su nombre indica inicia el trabajo, seguido a esto " Ping host " verifica si está encendida la máquina donde se encuentra la fuente y el mercado de datos. Después el componente " Check Db connections " verifica las conexiones y luego se procede a ejecutar los componentes " Job dimensiones " y " Job hechos ". Finalmente se muestra el mensaje: Trabajo finalizado con éxito.

Figura 16: Trabajo General

Fuente: Elaboración propia

3.2.4 Gestión del cambio lento en las dimensiones

En el mercado de datos Consulta Externa se implementó la gestión del cambio lento en las dimensiones mediante el tipo 2 de *SCD*. Teniendo en cuenta que el mismo permite mantener la información histórica a pesar de la ocurrencia de cambios en las dimensiones a lo largo del tiempo. Para su desarrollo se crean tres nuevos campos: la versión, la fecha de inicio y la fecha de fin. La versión contendrá un número secuencial que se incrementa al ocurrir un nuevo cambio sobre un valor en la fuente, inicialmente debe ser uno. Además, la fecha de inicio contiene el momento en que entró en vigencia el registro actual. Por último la fecha de fin guarda el registro actual que dejó de estar vigente.

3.3 Implementación del subsistema de visualización

Luego de realizada la carga de los datos y como parte de la implementación del tercer y último subsistema del mercado de datos Consulta Externa, se realizan los cubos *OLAP*, los reportes candidatos y los reportes que responden al CMI, así como la configuración del control de acceso al sistema.

3.3.1 Implementación de los cubos multidimensionales

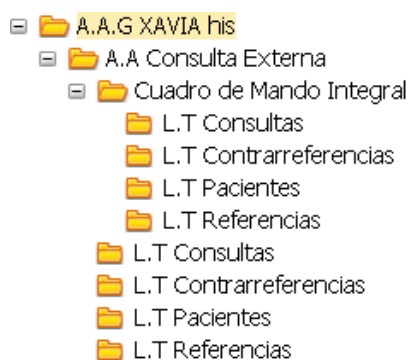
Los cubos son objetos *OLAP* que se componen de medidas, dimensiones relacionadas y son configurados en la base de datos. Como se había especificado en el capítulo 1, para la implementación de los cubos multidimensionales se decidió utilizar la herramienta Pentaho Schema Workbench. Primeramente se establece la conexión al *DM* y luego se crea el esquema que contiene los cubos con sus medidas y dimensiones asociadas.

A las medidas se les definió el tipo de dato, el origen de datos a la que se enlaza, el nombre que se muestra al usuario y la visibilidad. En el caso de las dimensiones, además de definir el nombre que será mostrado al usuario en los reportes, se configuraron otras propiedades de la dimensión como los atributos, las relaciones entre los atributos y las jerarquías definidas por el usuario.

Finalmente se genera un fichero *XML* que contiene los cuatro cubos multidimensionales correspondientes a cada uno de los hechos descritos en el modelo de datos. Existen tres dimensiones comunes para todos los hechos: *dim_tiempo*, *dim_servicio* y *dim_especialidad*, es decir, estas dimensiones están contenidas en cada uno de los cubos.

3.3.2 Arquitectura de información

En el capítulo 2 y como parte del diseño del subsistema de visualización, se describe la forma en que queda la arquitectura de información del mercado de datos Consulta Externa. A continuación se especifica cada uno de los elementos que componen las estructuras de navegación.



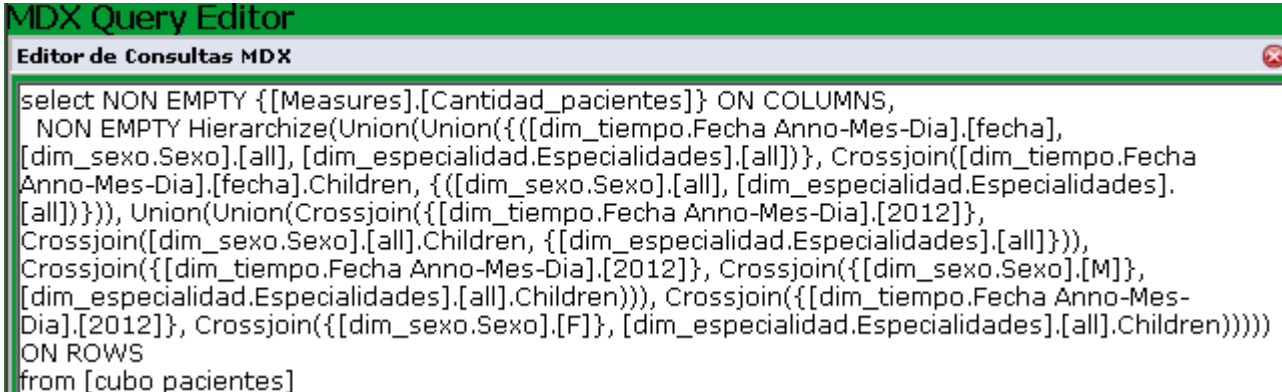
De esta manera se presenta la capa de visualización del mercado de datos Consulta Externa; la misma contiene un Área de Análisis (A.A), cuatro Libros de Trabajos (L.T) y cincuenta y dos reportes multidimensionales asociados a los libros de trabajos correspondientes. Incluye, además, el CMI con cuatro Libros de Trabajos (L.T) y veinte reportes asociados. La figura 17 muestra la composición de la arquitectura de información.

Figura 17: Arquitectura de información

Fuente: Elaboración propia

3.3.3 Implementación de los reportes candidatos

Cada uno de los libros de trabajo contiene un conjunto de reportes candidatos, los cuales son realizados mediante consultas *MDX* a la base de datos. Las consultas *MDX* constituyen un lenguaje de consulta para bases de datos multidimensionales sobre cubos *OLAP*. Son utilizadas en *BI* para la manipulación de datos y la definición de uso de los objetos de una base de datos multidimensional, permitiendo la generación de reportes hacia la toma de decisiones basada en el análisis de datos históricos. Se define una consulta *MDX* para cada uno de los reportes y gráficos, a continuación se muestra un ejemplo.



```
MDX Query Editor
Editor de Consultas MDX
select NON EMPTY {[Measures].[Cantidad_pacientes]} ON COLUMNS,
NON EMPTY Hierarchize(Union(Union({[dim_tiempo.Fecha Anno-Mes-Dia].[fecha],
[dim_sexo.Sexo].[all], [dim_especialidad.Especialidades].[all]}}, Crossjoin([dim_tiempo.Fecha
Anno-Mes-Dia].[fecha].Children, {[dim_sexo.Sexo].[all], [dim_especialidad.Especialidades].
[all]}))), Union(Union(Crossjoin({[dim_tiempo.Fecha Anno-Mes-Dia].[2012]},
Crossjoin([dim_sexo.Sexo].[all].Children, {[dim_especialidad.Especialidades].[all]})),
Crossjoin({[dim_tiempo.Fecha Anno-Mes-Dia].[2012]}, Crossjoin({[dim_sexo.Sexo].[M]},
[dim_especialidad.Especialidades].[all].Children))), Crossjoin({[dim_tiempo.Fecha Anno-Mes-
Dia].[2012]}, Crossjoin({[dim_sexo.Sexo].[F]}, [dim_especialidad.Especialidades].[all].Children))))
ON ROWS
from [cubo pacientes]
```

Figura 18: Consulta *MDX* para obtener la cantidad de pacientes atendidos por fecha, según especialidad y sexo

Fuente: Elaboración propia

A continuación se muestra el contenido que se visualiza en el Pentaho BI Server al ejecutar la consulta correspondiente al reporte "Obtener la cantidad de pacientes atendidos por fecha, según servicios, especialidades y sexo". El mismo muestra que para el año 2012 la mayor cantidad de pacientes atendidos corresponde al sexo femenino.

Capítulo 3. Implementación y validación del cuadro de mando integral

Fecha Año-Mes-Día	Especialidades	Servicios	Medidas		
			Cantidad pacientes		
			Sexo		
<input type="checkbox"/> Todos	<input type="radio"/> F	<input type="radio"/> M			
<input type="checkbox"/> 2012	<input type="checkbox"/> Todas	<input type="checkbox"/> Todos los servicios	84	45	39
		UCI Adulto	3		3
		Medicina general	3	3	
		Salud ocupacional	9	3	6
	Medicina general	<input type="checkbox"/> Todos los servicios	3	3	
		Medicina general	3	3	
	ECOR	<input type="checkbox"/> Todos los servicios	3	3	
		Salud ocupacional	3	3	
	UCI Adulto	<input type="checkbox"/> Todos los servicios	3		3
		UCI Adulto	3		3

Figura 19: Reporte Cantidad de pacientes atendidos por fecha, según servicios, especialidades y sexo
Fuente: Elaboración propia

La inclusión de gráficos en tercera dimensión en cada reporte facilita el proceso de toma de decisiones. La figura 20 muestra el gráfico correspondiente al reporte anterior.

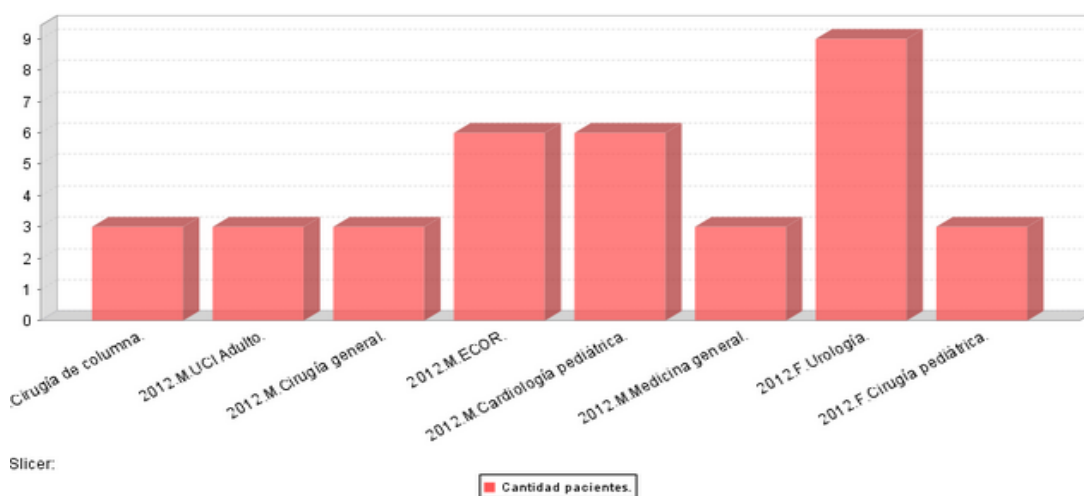


Figura 20: Gráfico Cantidad de pacientes atendidos por fecha, según especialidades y sexo
Fuente: Elaboración propia

3.3.4 Implementación del cuadro de mando integral

Existen varios tipos de CMI (Kaplan & Norton, 2005), de ellos se implementa un CMI departamental, puesto que es específico para el área de consulta externa. Como herramienta de análisis el CMI es un

Capítulo 3. Implementación y validación del cuadro de mando integral

concepto dinámico que favorece la toma de decisiones, lo que facilita el entendimiento de la información de forma resumida.

Los reportes en forma de gráficas son implementados mediante consultas *MDX* realizadas al mercado de datos, los cuales responden a los indicadores identificados en el diseño en la sección 2.10.4. Los distintos tipos de gráficos dinámicos que brinda el CMI permiten mostrar la información en forma de resumen. Si se desea ver el origen se puede entrar en detalle de forma dinámica. Un ejemplo de lo expresado se muestra en la figura 21.

El siguiente gráfico permite realizar análisis con respecto a la mayor o menor cantidad de pacientes atendidos en el año 2012, teniendo en cuenta una determinada especialidad. También si se compara esta cantidad real con la esperada por la entidad, se puede determinar si se cumplió o no con el plan de pacientes atendidos por especialidades. Además, se brinda la opción de analizar estos mismos parámetros por el promedio y el porcentaje de pacientes.

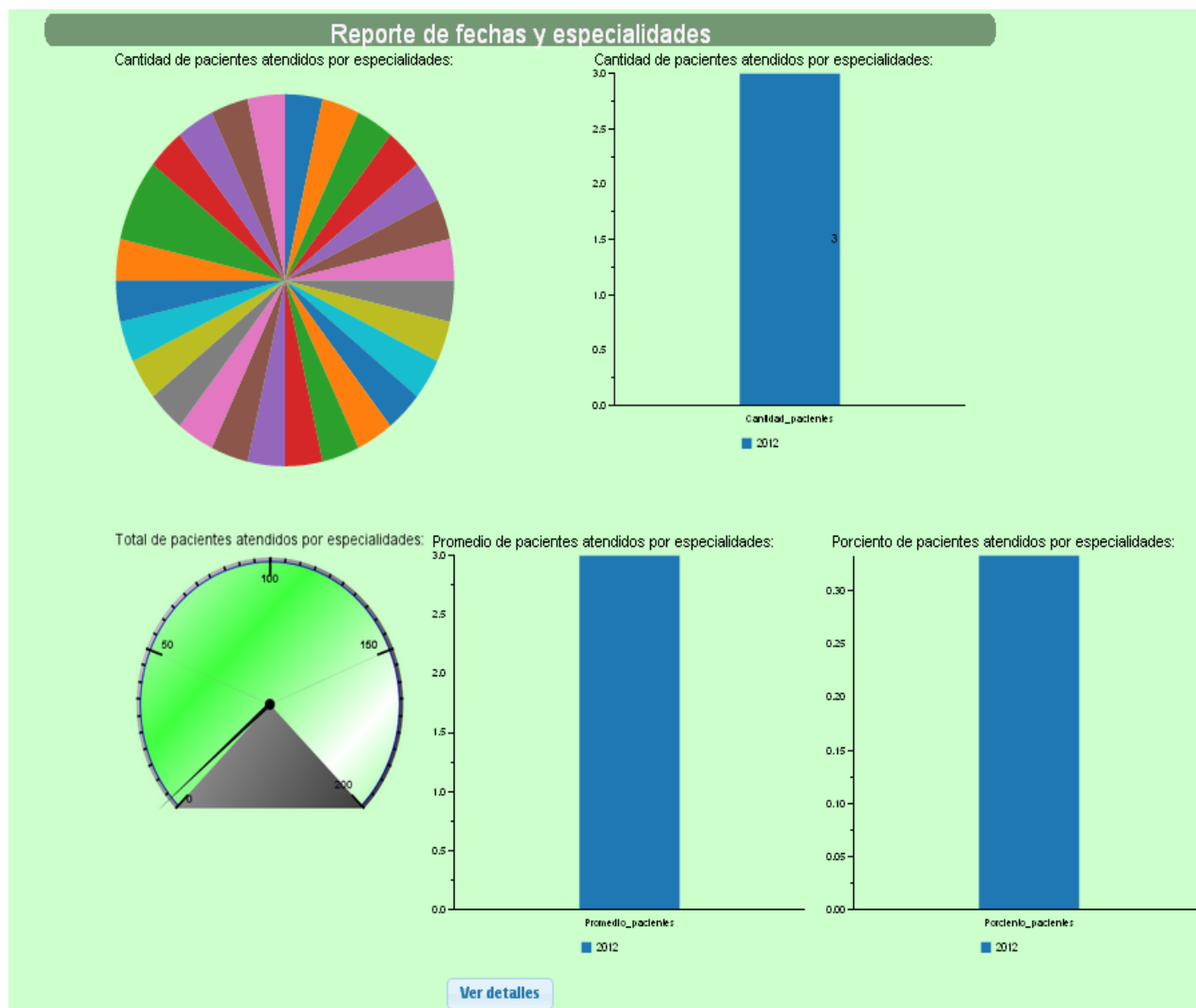


Figura 21: Reporte CMI Cantidad, promedio y porciento de pacientes atendidos por fecha, según especialidades

Fuente: Elaboración propia

A continuación se muestra el gráfico que permite obtener la cantidad, porciento y promedio de pacientes atendidos por fecha según tipo de paciente. El reloj muestra el total de pacientes atendidos por el mismo parámetro.

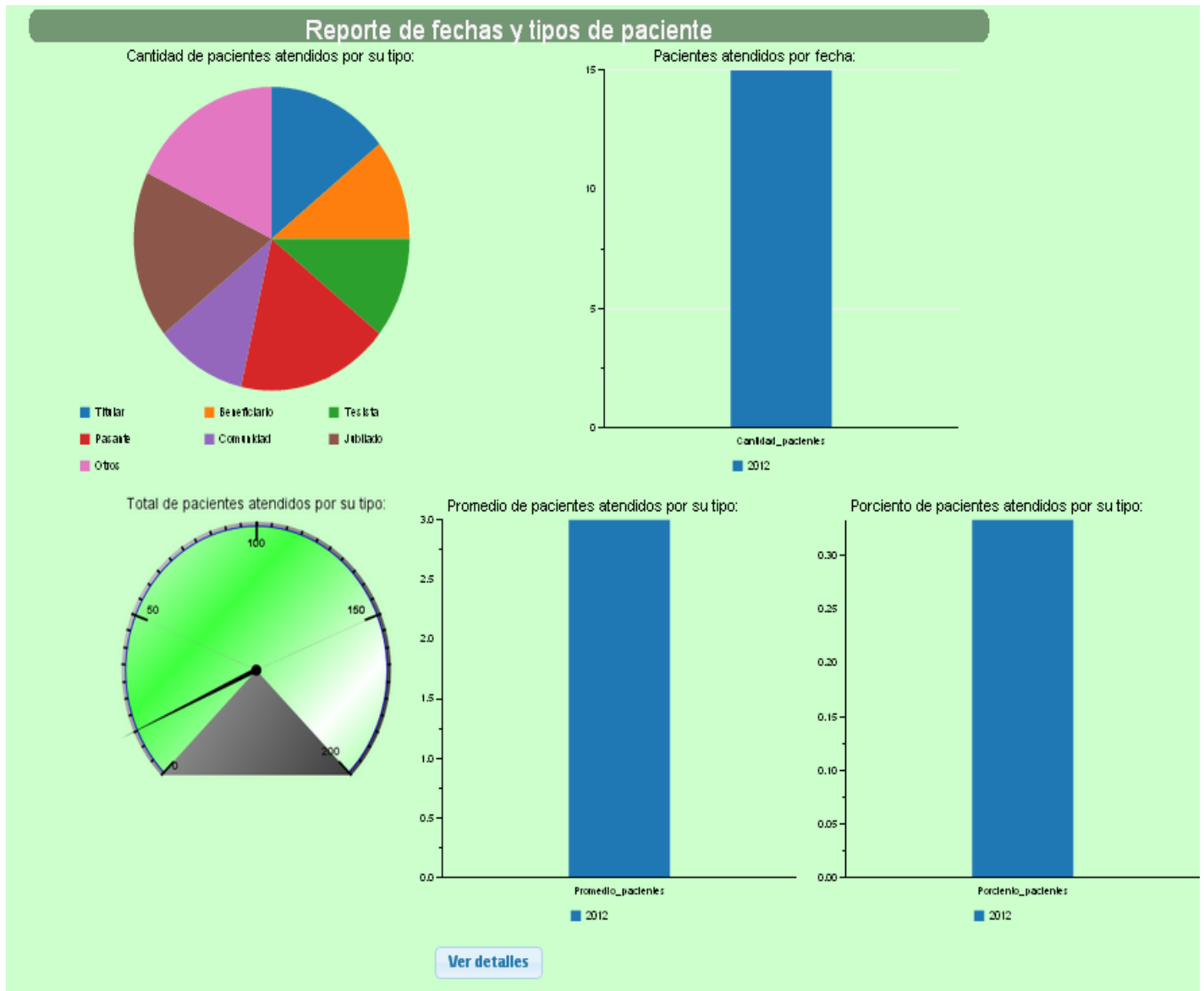


Figura 22: Reporte CMI Cantidad, promedio y porcentaje de pacientes atendidos por fecha según tipo de paciente.
Fuente: Elaboración propia

3.4 Pruebas

Una vez concluido el proceso de construcción del sistema propuesto, es necesario realizar una serie de pruebas y validaciones para comprobar el correcto funcionamiento del producto desarrollado. Resulta de gran importancia el diseño de pruebas de manera independiente durante cada fase del proceso de construcción de la solución. Existen diferentes tipos de pruebas que pueden ser aplicadas con el objetivo de obtener un producto con calidad y comprobar la completitud de respuesta del sistema a los requisitos definidos inicialmente (González, 2012). A la solución propuesta se le aplican pruebas a través del modelo V, el cual se usa como base de procesos dentro del ciclo de vida del software. El mismo tiene en cuenta las pruebas en paralelo con el ciclo de vida del proyecto y no como una función aislada que se realiza una vez concluido. Uno de sus principales objetivos es mejorar y garantizar la calidad del producto a desarrollar. En el presente trabajo se aplicaron casos de pruebas, listas de chequeo y la técnica de ladov que constituye una prueba de aceptación. En las siguientes secciones se describen cada una de las pruebas mostradas en la figura 23.

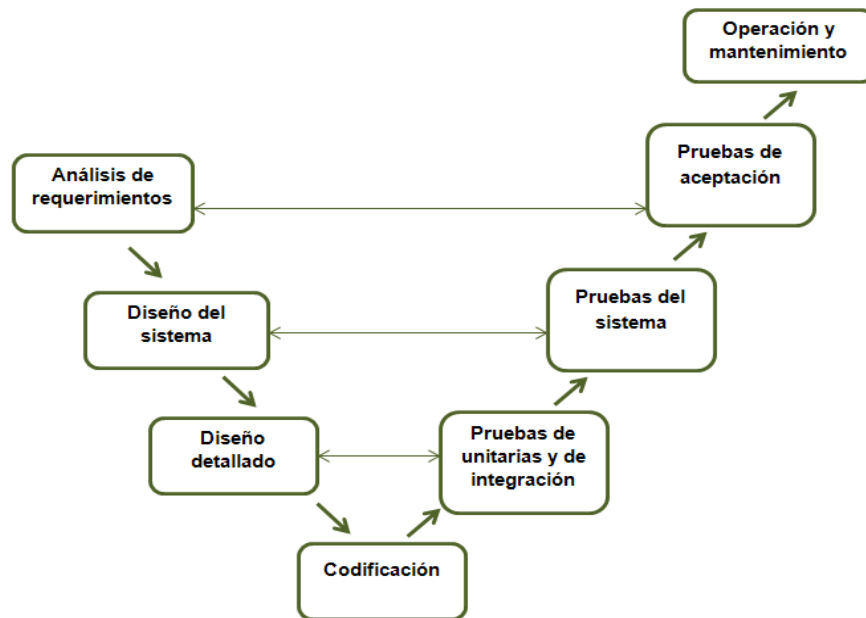


Figura 23: Modelo V
Fuente: Elaboración propia

3.4.1 Pruebas unitarias

Las pruebas unitarias se refieren a la prueba individual de unidades aisladas de un software. El objetivo de las pruebas unitarias es separar cada parte del programa y mostrar que las partes individuales funcionen tal y como lo define la especificación del programa. Las pruebas unitarias facilitan que el programador cambie el código en caso de encontrar algún error y luego vuelva a hacerle pruebas para asegurarse de que no ha introducido errores con los nuevos cambios.

Al concluir la implementación se efectúan pruebas unitarias por parte del equipo de desarrollo. Las mismas se realizan al subsistema de integración y visualización, aplicando pruebas internas al código generado por las consultas *MDX* y *SQL* sin ocurrencia de problemas. Además, se aplican pruebas a los componentes relacionados con la capa de visualización, donde fueron identificadas tres no conformidades (NC), las cuales se resolvieron en su totalidad. Las mismas se relacionan a continuación:

NC1. Ajustar en los nombres de los reportes las tildes y letras ñ.

NC2. Ajustar en los nombres de los gráficos del CMI el tamaño de letra para su mejor visibilidad.

NC3. Eliminar los gráficos en reportes que cuentan con muchas dimensiones, teniendo en cuenta que en estos casos el gráfico pierde su objetividad.

3.4.2 Pruebas de integración

Las pruebas de unidad por sí solas no descubren errores de integración, para ello se realizan pruebas donde se analiza la interacción entre las diferentes partes del sistema y se asegura que la comunicación, enlaces y los datos compartidos ocurran apropiadamente. El objetivo principal de las pruebas de integración realizadas en el presente trabajo es identificar errores en la integración de los subsistemas que conforman la solución propuesta. Se comprueba que los mismos se integran correctamente al extraerse los datos de la fuente de forma adecuada. Además, luego de los procesos de transformación estos son visualizados obteniendo los resultados esperados.

3.4.3 Pruebas del sistema

Las pruebas del sistema son similares a las pruebas de integración, solo que estas buscan probar el sistema como un todo, sin error. Dichas pruebas examinan que tan bien están los requerimientos generales y abarcan todas las partes combinadas del sistema. Constituyen actividades en las que el sistema o uno de sus componentes, se ejecuta bajo circunstancias previamente especificadas con el

objetivo de estudiar su funcionamiento. De esta manera, es posible observar y registrar los resultados para realizar una evaluación de los mismos. También se realizan a la documentación del sistema.

La estrategia definida para realizar las pruebas del sistema incluye la confección de una lista de chequeo y casos de prueba. Ambos resultan de gran importancia para demostrar la funcionalidad del mismo.

Listas de chequeo

La lista de chequeo consta de una serie de preguntas, en forma de cuestionario, mediante el cual se verifica el grado de cumplimiento de determinadas reglas establecidas para los procesos de desarrollo del sistema, además de medir la calidad de los artefactos generados durante la realización del producto. Esta evaluación se desarrolla a través del análisis de un grupo de indicadores, distribuidos en tres secciones fundamentales:

- ✓ **Estructura del documento:** abarca todos los aspectos definidos por el expediente de proyecto o el formato establecido por el proyecto.
- ✓ **Indicadores definidos:** abarca todos los indicadores a evaluar durante la etapa.
- ✓ **Semántica del documento:** contempla todos los indicadores a evaluar respecto a la ortografía, redacción y demás.

La estructura de la lista de chequeo está formada por los siguientes elementos:

- ✓ **Peso:** define si el indicador a evaluar es crítico o no.
- ✓ **Indicadores a evaluar:** son los indicadores a evaluar en las secciones **Estructura del documento, Semántica del documento e Indicadores definidos por la etapa.**
- ✓ **Evaluación (Eval):** es la forma de evaluar el indicador en cuestión. El mismo se evalúa de 1 en caso de que exista alguna dificultad sobre el indicador y 0 en caso de que el indicador revisado no presente problemas.
- ✓ **No Procede (N.P.):** se usa para especificar que el indicador no es necesario evaluarlo en ese caso.
- ✓ **Cantidad de elementos afectados:** especifica la cantidad de errores encontrados sobre el mismo indicador.

Capítulo 3. Implementación y validación del cuadro de mando integral

- ✓ **Comentario:** especifica los señalamientos o sugerencias que quiera incluir la persona que aplica la lista de chequeo. Pueden o no existir señalamientos o sugerencias.

Luego de quedar definida la estructura de la lista de chequeo, esta es aplicada al mercado de datos Consulta Externa, la misma se encuentra en el artefacto lista de chequeo. Los resultados arrojados se muestran en la figura 24, donde se visualiza el comportamiento de los veinte y cuatro indicadores. Los mismos son divididos en secciones, para un total de 2 no conformidades en la sección indicadores definidos por etapa.

NC1. No se utiliza el menor número de transformaciones posibles al cargar los datos hacia el área de trabajo.

NC2. No se tuvo en cuenta los formatos fuentes y tipos de datos de las perspectivas de análisis.

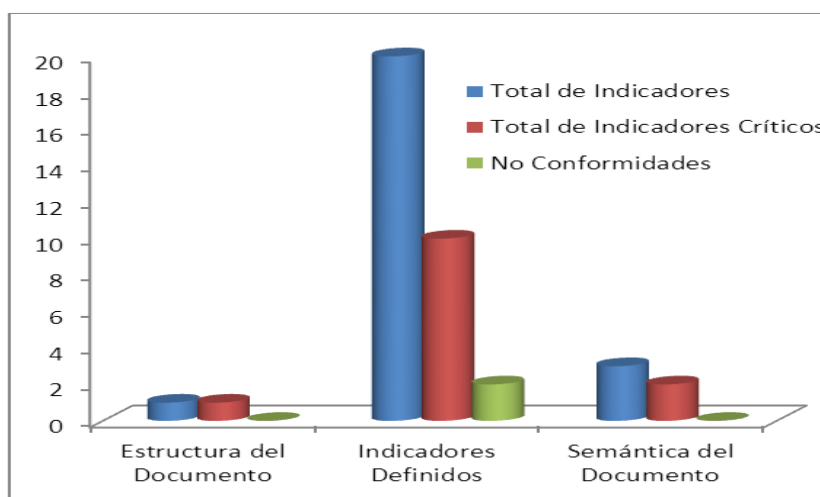


Figura 24: Comportamiento de los indicadores por secciones
Fuente: Elaboración propia

Casos de prueba

Se aplica la técnica partición de equivalencia basada en casos de prueba como parte de las pruebas de caja negra. Los casos de prueba son utilizados para identificar posibles fallos de implementación y comprobar el grado de cumplimiento de los requisitos especificados para el sistema. Fueron diseñados cuatro casos de prueba en correspondencia con los casos de uso de información identificados. Los mismos están regidos por los reportes candidatos asociados a cada tema de análisis, con el fin de

Capítulo 3. Implementación y validación del cuadro de mando integral

comprobar que sean visualizados con las variables correspondientes. La tabla 8 describe el caso de prueba asociado al CUI Obtener datos de pacientes.

Tabla 8: Caso de prueba basado en CUI Obtener datos de pacientes

Escenario	Descripción	Perfiles de análisis	Indicadores a medir	Respuesta del sistema	Flujo central
EC1. Obtener datos de pacientes.	Muestra la cantidad, porcentaje y promedio de pacientes atendidos en una fecha determinada.	<ul style="list-style-type: none"> ✓ Tiempo ✓ Servicios ✓ Hospital ✓ Especialidades ✓ Tipo de paciente ✓ Diagnóstico ✓ Tipo población ✓ Sexo 	Cantidad, porcentaje y promedio de pacientes.	El sistema muestra todas las variables disponibles para el análisis, ubicado en las filas y las columnas que pueden ser visualizadas en el reporte.	<p>El usuario se autentica. Despliega la carpeta A.A.G xavia HIS.</p> <p>Luego selecciona la carpeta A.A Consulta Externa y dentro de esta el L.T Pacientes.</p> <p>Finalmente selecciona el reporte que desea visualizar en relación a los datos del paciente.</p>

Las pruebas del sistema fueron supervisadas por especialistas del Departamento Almacenes de Datos. Los cuatro casos de prueba basados en CUI permitieron identificar un total de cuatro NC, dos de complejidad baja, una de complejidad media y una de complejidad alta, las mismas fueron resueltas por el equipo de desarrollo. Los resultados obtenidos se muestran en la figura 25 que contiene el total de NC y su distribución según la complejidad.

NC1. Uso excesivo de mayúsculas en el A.A Consulta Externa y los L.T.

NC2. El reporte “Obtener la cantidad de pacientes atendidos en un mes y año según la especialidad” no muestra correctamente la perspectiva de análisis especialidad.

NC3. Error ortográfico en el nombre del reporte “Obtener la cantidad de contrarreferencias emitidas en un mes y año dado por médicos, según servicios”.

NC4. Uso inadecuado del idioma inglés.

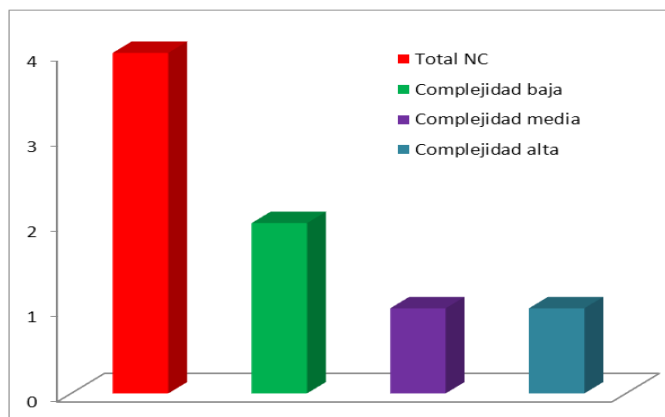


Figura 25: Resultados de la aplicación de los casos de prueba basados en CUI al sistema
Fuente: Elaboración propia

Pruebas de aceptación

Las pruebas de aceptación evalúan el grado de calidad del software con relación a todos los aspectos relevantes para que el producto sea aceptado por los usuarios finales. La aplicación de las pruebas de aceptación al mercado de datos Consulta Externa fue realizada mediante la técnica de ladov (Tejedor, 2005) para determinar el grado de satisfacción de los usuarios y el criterio de los especialistas.

La aplicación de la técnica de ladov constituye un instrumento de gran valor para el estudio de la satisfacción - insatisfacción de los usuarios. Se realiza un cuestionario conformado por tres preguntas cerradas y dos preguntas abiertas intercaladas con otras preguntas (ver Anexo1). Las posibles respuestas de las preguntas cerradas son: si, no sé y no. Las posibles combinaciones de estas respuestas tienen un valor entre 1 y 6 determinado en el Cuadro Lógico de ladov (ver Anexo 2) y representan la posición de cada sujeto en la siguiente escala de satisfacción.

- | | | |
|--------------------------------------|--------------------------------------|-------------------------|
| 1. Clara satisfacción | 3. No definida | 5. Clara insatisfacción |
| 2. Más satisfechos que insatisfechos | 4. Más insatisfechos que satisfechos | 6. Contradictoria |

A dichas posiciones les corresponde un valor que se expresa en la tabla 9, que oscila entre -1 y 1.

Tabla 9: Niveles de satisfacción

Niveles de satisfacción	
1	Máximo de satisfacción
0,5	Más satisfechos que insatisfechos
0	No definido y contradictorio
- 0,5	Más insatisfechos que satisfechos
-1	Máxima insatisfacción

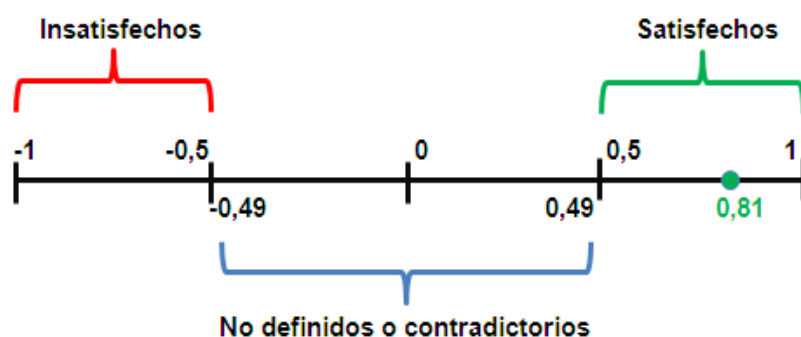
Por último utilizando la siguiente fórmula se calcula el índice de satisfacción grupal:

$$ISG = \frac{A(1) + B(0,5) + C(0) + D(-0,1) + E(-1)}{N}$$

Fórmula para calcular el índice de satisfacción grupal

En esta fórmula A, B, C, D, E, representan el número de sujetos con índice individual 1; 2; 3; 4; 5 y 6 donde N representa el número total de sujetos que responden el cuestionario.

El resultado de calcular el índice de satisfacción grupal para el sistema propuesto arroja un valor de 0,81 ubicado en el rango entre 0,5 y 1, lo cual indica que existe satisfacción. La figura 26 muestra lo antes mencionado.



A pesar de que se haya comprobado que el sistema satisface los requisitos definidos, resulta imprescindible proveer una correcta preparación a los especialistas que darán uso al mismo. Luego la aplicación se encuentra lista para su ejecución e instalación.

Figura 26: Índice de satisfacción grupal

Fuente: Elaboración propia

3.4.4 Calidad de datos en el mercado de datos consulta externa

El término "calidad de datos" generalmente es asociado a los sistemas de información que manejan grandes volúmenes de datos, ya que puede encontrarse información incompleta e inconsistente. En un mercado de datos, este tipo de situaciones pueden ser más comunes, puesto que son integradas distintas fuentes de datos. Se definen a continuación uno de los procesos que permite comprobar la calidad de la información almacenada en el mercado de datos Consulta Externa.

Perfilado de datos

Se realiza un nuevo análisis de perfilado, esta vez sobre el mercado de datos Consulta Externa, para conocer el estado de los datos luego de haberse cargado y así verificar si cumplen con las reglas del negocio y transformación definidas anteriormente. Los resultados arrojados por el proceso indican que la carga de los datos correspondientes a cada uno de los hechos se realizó correctamente. No fueron almacenados valores vacíos ni nulos y los hechos contienen únicamente valores enteros, a pesar de que la fuente presenta varios tipos de datos. El siguiente gráfico muestra los resultados correspondientes al perfilado de datos realizado a la tabla hech_paciente.

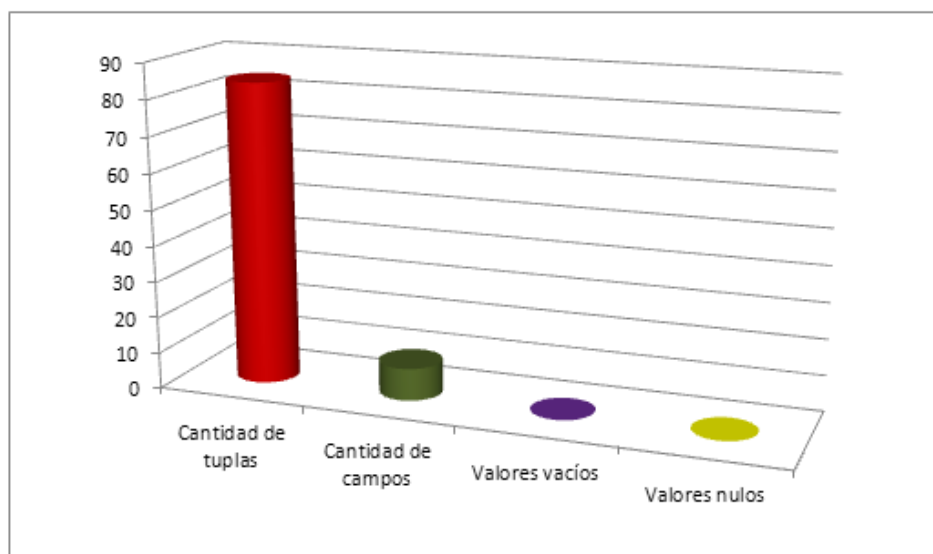


Figura 27: Resultados perfilado de datos hech_paciente
Fuente: Elaboración propia

3.5 Diagrama de despliegue de la solución

Los diagramas de despliegue exponen las relaciones físicas de los diferentes nodos que conforman un sistema y la distribución de los componentes sobre dichos nodos, además, proporcionan una vista arquitectónica de alto nivel del sistema. También ayudan a visualizar el camino de la implementación y muestran cómo y dónde se desplegará el sistema. Los nodos constituyen recursos de cómputos entre los cuales existen relaciones que representan el medio de comunicación. Por último permiten comprender la correspondencia entre la arquitectura de software y la arquitectura de hardware.

En el presente trabajo los usuarios se conectan desde la estación cliente por el protocolo seguro de transferencia de hipertexto (*Hypertext Transfer Protocol Secure - HTTPS por sus siglas en inglés*) hacia el servidor de inteligencia de negocios que posee los subsistemas de visualización e integración de los datos. Este a su vez accede al servidor que contiene el SGBD con el subsistema de almacenamiento. Además los usuarios pueden imprimir los reportes *OLAP* mostrados a través de una impresora conectada mediante el puerto bus universal en serie (*Universal Serial Bus - USB por sus siglas en inglés*) a sus estaciones de trabajo La figura 28 muestra el diagrama descrito.

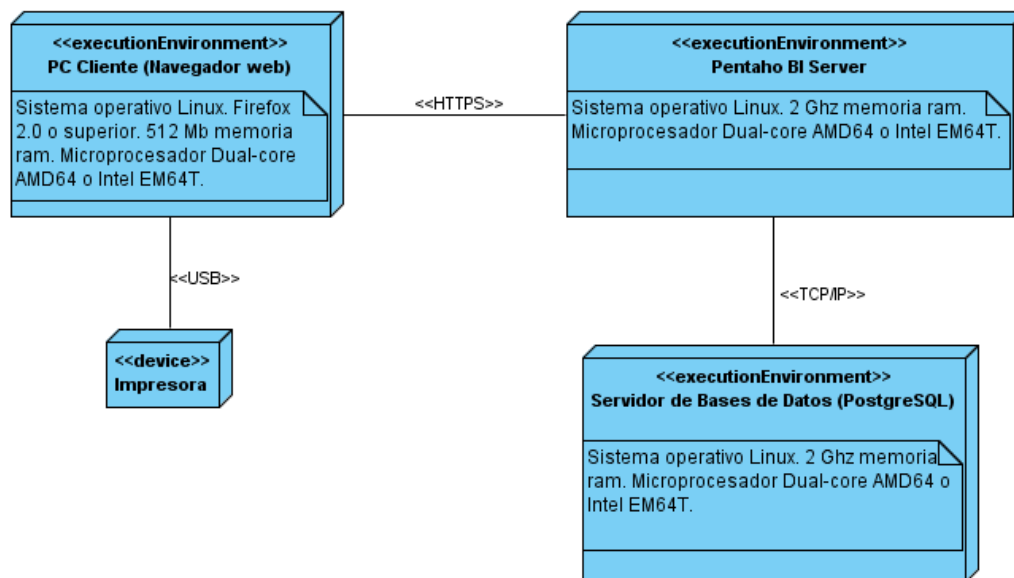


Figura 28: Diagrama de despliegue
Fuente: Elaboración propia

Conclusiones del capítulo

En el presente capítulo se abordó la implementación y prueba del sistema propuesto concluyendo el proceso de construcción del mismo. Además las reglas del negocio y transformación identificadas en la fase de requisitos, fueron aplicadas durante las transformaciones de las dimensiones y los hechos.

Los metadatos de procesos implementados almacenan información acerca de la ejecución de las transformaciones, permitiendo a los administradores la gestión del sistema y de esta manera el correcto funcionamiento del mismo.

La utilización del modelo V permitió considerar las actividades de prueba como un proceso que ocurre durante todo el ciclo de vida, en lugar de establecer una fase independiente al concluir el proceso de desarrollo.

Conclusiones

La investigación realizada cumple los objetivos planteados inicialmente mediante el desarrollo de la solución propuesta, puesto que resultó viable la aplicación del CMI sobre el mercado de datos Consulta Externa, como herramienta de análisis para los directivos médicos. Además, el estudio de las metodologías y herramientas realizado, garantizó que la propuesta de metodología seleccionada guiara el proceso de desarrollo del sistema a través de cada fase del ciclo de vida.

Por su parte en el perfilado de datos realizado a las 14076 tuplas de la fuente se detectaron 11749 valores duplicados. Sin embargo al realizar dicho proceso al mercado de datos Consulta Externa no se detectaron errores, asegurando que los datos mostrados en los reportes sean fiables. Finalmente la inclusión de la técnica ladov dentro del modelo V permitió que las pruebas efectuadas al sistema durante la etapa de aceptación validaran el cumplimiento de los requisitos establecidos, pues el índice de satisfacción grupal es de 0,81.

Recomendaciones

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

- ✓ Actualizar el CMI implementado con información referente a los recursos materiales empleados para las distintas especialidades de las consultas externas para obtener un control de la cantidad necesaria de los mismos, reducir los gastos y evitar el desvío de recursos.

Referencias Bibliográficas

Bibliografías

- Bernabeu, R.D. (2007). *Data Warehousing _ Hefesto*. Argentina. p. 5-9 y 21-25.
- Bernabeu, R.D. (2009). *Metodología propia para la Construcción de un Data Warehouse*. Córdoba, Argentina.
- Cabero, A.J. (2005). Las TICs y las Universidades: retos, posibilidades y preocupaciones. *Revista de la Educación Superior*, Vol. 34. 0185-2760. p. 1-2.
- Cristiá, A.A., Morgado, Y.S., & et al. (2010). *Guía para la personalización de PostgreSQL 8.4*. La Habana , Cuba. p. 6.
- Codd, E.F. & Salley, C.T. (1993). *Providing OLAP (on-line Analytical Processing) to User-analysts: An IT Mandate*. US : Codd & Date, Inc, 1993. Vol. 32. 0718-3305. p. 31
- Dávila, A. (1999). Nuevas herramientas de control: El Cuadro de Mando Integral. *Revista de antiguos alumnos.*, 23(3), ISSN (0213-9111). p. 2-4.
- Eunice L.B. & Ripoll V.F. (2003). *Una aproximación a las implicaciones del cuadro de mando integral en las organizaciones del sector público*. España. p. 23-25.
- Gascho, M., & Salteiro, S. (2000). *The balanced scorecard Judgmental effects of common and unique performance measures*. *The Accounting Review*.
- González, Y. H. (2010). *Metodología para el Desarrollo de Soluciones de Almacenes de Datos e Inteligencia de Negocios en DATEC*. La Habana, Cuba. p.1-4 y 35.
- Inmon, W. H. (2002). *Building the Data Warehouse*.(3ra Edición ed.) Canadá: Wiley Publishing, Inc. p. 19 y 77-122.
- Kaplan, R., Norton, D. (junio 2005). *El cuadro de mando integral. "Los directivos, como los pilotos, necesitan de un instrumental que mida su entorno y su rendimiento para conducir el viaje hacia la excelencia futura."* p. 2
- Kimball, R., & Caserta, J. (2004). *The Data Warehouse ETL Toolkit* (1ra ed.). 10475 Crosspoint Boulevard Indianapolis: Wiley Publishing, Inc. p. 113-116 y 351-357.
- Kimball, R., & Ross, M. (2002). *The Data Warehouse Toolkit*. (2da ed.). (R. Elliott, Ed.) Canada: Wiley Computer Publishing.

- Medina, D.N., Iznaga, Y.G., & et al. (2012). *Procesos de integración de datos. Tema 2: Herramientas de Integración de Datos Funcionamiento, Arquitectura Y Componentes del Pentaho Data Integration*. Cuba. p. 21.
- Nader, J. (2003). *Tesis de Magister en Ingeniería del Software. Sistema de Apoyo Gerencial Universitario*. p. 55-57.
- Övergaard G., & Palmkvist, K. (2004). *Use Cases Patterns and Blueprints*. Addison Wesley Professional. 0-13-145134-0.
- Santos, M. C. & Fidalgo E.C. (2006). *Un análisis de la flexibilidad del Cuadro de Mando Integral en su adaptación a la naturaleza de las organizaciones*. Madrid. p. 2-5 y 8-10
- Stratebi. (2010). *Comparativa B.I. Open Source*. España, 11.
- Tamborero, G.C. (2008). *¿ Es útil para la atención primaria un sistema de información de base poblacional*. España. p. 174.
- Thomsen, E. (2002). *OLAP Solutions Building Multidimensional Information Systems*. New York.
- Vogelmann , E. M. (2008). *Políticas y Modelos de Seguridad*. Argentina. p. 16-17.
- Zepeda, L. (8 de junio de 2008). *Tesis doctoral Metodología para el Diseño Conceptual de Almacenes de Datos*. España. p. 148-159
- Zorrilla, M. (2007). *Data werehouse y OLAP*. España, Cantabria. p. 21

Webgrafía

- Bernabeu, D. (20 de 5 de 2010). Dataprix. Recuperado el 20 de 1 de 2013, de Dataprix: <http://www.dataprix.com/blogs/bernabeu-dario/dimensiones-degeneradas>
- Flores, E.P., & Uris,J.S., & et al. (Jun de 2009). scielo.isciii.es. Recuperado el 29 de 10 de 2012, de scielo.isciii.es: http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S0213-91112009000300017&lng=es.
- López, E.H., Cuevas, L.R., & et al. (15 de abril de 2013). bvs.sld.cu. Recuperado el 19 de noviembre de 2012, de bvs.sld.cu: http://www.bvs.sld.cu/revistas/san/vol_16_7_12/san20712.htm
- Martínez, A. O. (2007). EcuRed. Recuperado el 27 de 4 de 2013, de EcuRed: <http://www.gestiopolis.com/canales8/ger/olap-online-analytic-processing.htm>

- hosting diario. (17 de enero de 2013). hosting diario. Recuperado el 29 de abril de 2013, de hosting diario:
<http://hostingdiario.com/los-distintos-tipos-de-backups/>
- Pentaho. (2007). Pentaho. Recuperado el 11 de 2012, de Pentaho:
<http://mondrian.pentaho.com/documentation/workbench.php>
- pgAdmin. (2010). pgAdmin. Recuperado el 20 de 1 de 2013, de pgAdmin:
<http://www.pgadmin.org/features.php>
- Sánchez, M.A., Días, F.J. (25 de 03 de 2008). www.portalesmedicos.com. Recuperado el 12 de 12 de 2012, de www.portalesmedicos.com:
<http://www.portalesmedicos.com/publicaciones/articulos/999/1/Sistema-de-Soporte-a-la-Decisiones-para-la-deteccion-automatizada-de-Sindrome-Metabolico-en-un-Laboratorio-de-Fisiologia-Humana.html>
- Sanchez, M., Natalini, J., & et al. (2010). bibliotecadigital.uns.edu.ar. Recuperado el 26 de 10 de 2012, de bibliotecadigital.uns.edu.ar: <http://bibliotecadigital.uns.edu.ar/scielo.php?script=sci_arttext&pid=S1853-20552010000100004&lng=en&nrm=iso>. ISSN 1853-2055
- Tejedor, L. J. C. (junio de 2005). efdeportes.com. Recuperado el 20 de abril de 2013, de efdeportes.com:
<http://www.efdeportes.com/efd85/iadov.htm>
- Visual Paradigm. (2008). Visual Paradigm. Recuperado el 11 de 2012, de Visual Paradigm:
<http://www.visual-paradigm.com/support/vpuml/releasenotes/640.jsp>
- (Buenas Tareas, 2010). Cuadro De Mando Integral. BuenasTareas.com. Recuperado 10, 2010, de Buenas Tareas.com:
<http://www.buenastareas.com/ensayos/Cuadro-De-Mando-Integral/952617.html>

Glosario de Términos

Referencias: Documentos clínicos creados cuando un paciente es referido a otra entidad o especialidad.

Contrarreferencias: Documentos clínicos creados cuando el paciente referido es enviado a otra entidad o especialidad.

Anexos

Anexo 1

Cuestionario aplicado

- 1) ¿Encuentras factible el desarrollo del Cuadro de Mando Integral para el módulo Consulta Externa?
_Si _No _No sé
- 2) ¿El Cuadro de Mando Integral para el módulo Consulta Externa facilita la interacción con la información referida al mismo?
_Si _No _No sé
- 3) ¿Al consultar la información referente al módulo Consulta Externa preferirías hacerlo en otra aplicación que no fuera el Cuadro de Mando Integral para el módulo Consulta Externa?
_Si _No _No sé
- 4) ¿Las funcionalidades del Cuadro de Mando Integral para el módulo Consulta Externa se ajustan a los requisitos definidos en el análisis?
_Si _No _No sé
- 5) ¿Qué es lo que más te disgusta del Cuadro de Mando Integral para el módulo Consulta Externa?
- 6) ¿Crees importante el uso de diferentes herramientas para la toma de decisiones?
_Si _No _No sé
- 7) ¿Qué es lo que más te gusta del Cuadro de Mando Integral para el módulo Consulta Externa?
- 8) ¿Para tomar decisiones tácticas y estratégicas escogerías el módulo Consulta Externa del xavia HIS o el Cuadro de Mando Integral para el módulo Consulta Externa?
_Si _No _No sé
- 9) ¿El Cuadro de Mando Integral para el módulo Consulta Externa es fácil de manipular?
_Si _No _No sé
- 10) ¿Te gusta utilizar el Cuadro de Mando Integral para el módulo Consulta Externa?
_Me gusta mucho
_Me gusta más de lo que me disgusta
_Me da lo mismo
_Me disgusta más de lo que me gusta
_No me gusta nada
_No sé decir

Anexo 2

Cuadro Lógico de ladov

Cuadro Lógico de ladov									
	3. ¿Quisieras ir a estudiar otra asignatura o hacer otra cosa en el horario de educación física?								
	NO			NO SÉ			SI		
10. ¿Te gusta la clase de educación física?	8. Si tu pudieras escoger entre asistir o no asistir a las clases de educación física ¿Irías a esas clases?								
	Si	No sé	No	Si	No sé	No	Si	No sé	No
Me gusta mucho	1	2	6	2	2	6	6	6	6
No me gusta tanto	2	2	3	2	3	3	6	3	6
Me da lo mismo	3	3	3	3	3	3	3	3	3
Me disgusta más de lo que me gusta	6	3	6	3	4	4	3	4	4
No me gusta nada	6	6	6	6	4	4	6	4	5
No se que decir	2	3	6	3	3	3	6	3	4