



Universidad de las Ciencias
Informáticas

Universidad de las Ciencias Informáticas

Facultad 3

**Desarrollo de un cuadro de mando para la
Dirección Comercial de la XETID**

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

Autora:

Yoelsis Elisabeht Guerra Viltres

Tutor:

Ing. Olía Castellanos Leyva

La Habana, Cuba

Curso: 2019 – 2020

DECLARACIÓN DE AUTORÍA

Declaro ser el único autor de la presente tesis y reconozco a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo.

Y para que así conste, firmo la presente declaración de autoría en La Habana a los ___ días del mes de _____ del año 2020.

Yoelsis Elisabeht Guerra Viltres
Firma de Autor

Olia Castellanos Leyva
Firma de Tutor

AGRADECIMIENTOS

Agradezco a mis padres por ser mi motivación para alcanzar mis metas, por ser siempre mis guías y confiar en mí, por su dedicación y entrega durante estos cinco años de esfuerzo y sacrificio. Gracias por ser los mejores padres que la vida me pudo dar.

A mis mamis Felicia y Blanca por estar siempre pendientes de mí y brindarme su apoyo en todo momento.

A mi hermanito por confiar siempre en mí.

A mi tía Zoila y mi tío Carlos por ser otros padres para mí y estar presentes en cada paso que doy.

A mi novio Yasel Romero por convertirse en mi ejemplo a seguir desde que lo conocí, por apoyarme incondicionalmente y aconsejarme cuando creía que todo estaba perdido.

A mi abuelo Julio por apoyarme y ayudarme siempre.

A mi bisabuela China por brindarme su amor, aunque ya hoy no se encuentre conmigo.

A toda mi familia, mis tíos, primos que de una forma u otra siempre me han apoyado.

A mis queridas amigas Danays y Camila por compartir alegrías y llantos, pero sobre todo por siempre estar ahí juntas las tres en los momentos que de verdad nos necesitamos.

A mi tutora Olía por apoyarme en todo momento y confiar siempre en mí.

A mis amigos Angel y Lixanis por convertirse en parte de mi familia y darme ese sobrinito bello "Sebastián".

A mis compañeros de aula por formar parte de mi día durante mi paso por la Universidad.

A mis compañeras Dayana, Rachel y Yunisleidis por formar parte de mi equipo de estudio.

A todos los que de una forma u otra han contribuido a la realización de este trabajo.

DEDICATORIA

A mis padres por brindarme su apoyo incondicional y ser mi inspiración en la vida.

A mi hermanito por ser mi razón de vivir y uno de mis tesoros más preciados.

A mis mamis por brindarme su amor y apoyo en todo momento

A mi novio por ser el mejor compañero de vida y el que me acompaña cada día.

A mis grandes amigos por estar ahí cada día y apoyarme incondicionalmente.

RESUMEN

En las actividades diarias de cualquier organización se generan datos como resultado de todas las operaciones que se realizan. Es muy común que los mismos se almacenen y administren a través de sistemas transaccionales en bases de datos relacionales. La idea central que persiguen las organizaciones es que dejen de ser solo simples datos para convertirse en información que fortalezca la toma de decisiones. Precisamente la inteligencia de negocios entra a jugar un papel protagónico, y dentro de ella los Almacenes de Datos constituyendo las herramientas necesarias para gestionar y analizar la información contenida en estos sistemas transaccionales.

La presente investigación tiene como objetivo el desarrollo de un Cuadro de Mando para la Dirección Comercial de la Empresa de Tecnologías de la Información para la Defensa (XETID), incorporando un Almacén de Datos como sistema de solución. El desarrollo de este Almacén de Datos surge para facilitarle al cliente una forma más eficiente de poder almacenar la información que se gestiona durante el proceso de contratación, para su posterior explotación y análisis con vistas a facilitar la toma de decisiones. Durante la investigación se realizó un estudio de las metodologías, herramientas y tecnologías utilizadas en el desarrollo de este tipo de soluciones, y se seleccionaron las más adecuadas para que guiaran la propuesta de solución. Se utilizó para la validación de la solución las pruebas de volumen, carga y estrés para comprobar su capacidad y rendimiento en consecuencia con los requisitos definidos por el cliente.

Palabras claves: Cuadro de Mando, almacén de datos, metodología, indicadores, dimensiones

ÍNDICE

AGRADECIMIENTOS	II
DEDICATORIA	III
RESUMEN	IV
INTRODUCCIÓN	1
CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA	5
1.1 Definición de Cuadro de Mando	5
1.2 Introducción a los Almacenes de datos	6
1.2.1 Diferencias entre AD y una Base de Datos Operacional	7
1.3 Sistemas homólogos	7
1.4 Modelado multidimensional de los AD	11
1.5 Tipos de modelamiento de AD	12
1.6 Metodologías para el diseño e implementación de un AD.	13
1.7 Sistemas OLTP y OLAP	15
1.7.1 Procesamiento Transaccional en Línea (OLTP).	16
1.7.2 Procesamiento Analítico en Línea (OLAP).	16
1.8 Diseño del sistema de Extracción, Transformación y Carga	18
1.9 Herramientas utilizadas para el desarrollo del AD	18
1.9.1 Herramienta de modelado	18
1.9.2 Sistema Gestor de Base de Datos	19
1.9.3 Herramienta para el proceso de Extracción, Transformación y Carga	20
1.9.4 Herramienta para el proceso de Inteligencia de negocio	20
1.9.5 Herramienta para el análisis multidimensional	20
1.10 Conclusiones parciales	21
CAPÍTULO 2. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE LA SOLUCIÓN	22
2.1 Descripción de la propuesta de solución	22
2.2 Descripción de las fases de la Metodología Híbrida	23
2.2.1 Fase 1: Análisis de requerimientos.	23
2.2.2 Fase 2: Análisis de los OLTP	25
2.2.3 Fase 3: Modelo Lógico del AD	28
2.2.4 Fase 4: Proceso ETL	31
2.3 Conclusiones parciales	33
CAPÍTULO 3. VALIDACIÓN Y PRUEBAS	34

ÍNDICE

3.1	Pruebas de software	34
3.2	Diseño de casos de prueba	34
3.3	Pruebas de aceptación.....	36
3.4	Conclusiones parciales	38
CONCLUSIONES		40
RECOMENDACIONES		41
REFERENCIAS BIBLIOGRÁFICAS		42
ANEXOS.....		45

ÍNDICE DE FIGURAS

Figura 1. Diseño de la propuesta de solución.....	22
Figura 2. Modelo Conceptual.....	25
Figura 3. Modelo Conceptual Ampliado.....	28
Figura 4. Modelo Físico del Almacén de Datos.....	31
Figura 5. Representación de la cantidad de NC por iteraciones	36
Figura 6. Resultado de prueba de aceptación	38

ÍNDICE DE TABLAS

Tabla 1. Indicadores y perspectivas	24
Tabla 2. Determinación de Indicadores.....	26
Tabla 3. Nivel de granularidad	27
Tabla 4. Dimensiones del Almacén de Datos.....	29
Tabla 5. Hechos del Almacén de Datos.....	30
Tabla 6. Mapeo de datos.....	32
Tabla 7. Prueba de aceptación para Cantidad de contratos	37

INTRODUCCIÓN

Las herramientas de gestión y control son muy bien aceptadas por la mayor parte de las organizaciones. Muchos directivos de empresas comprenden la necesidad e importancia de poseer una herramienta que les permita evaluar el desempeño de su organización, sobre todo en grandes empresas que poseen un gran volumen de trabajo y se hace complejo llevar el control de forma manual. Actualmente las organizaciones requieren manejar adecuadamente información relevante de sus procesos, para monitorearlos y tomar decisiones objetivas en aras del desarrollo exitoso de su negocio. La adecuada presentación de la información es esencial para una evaluación más objetiva de los procesos de la institución y contribuir así a una efectiva toma de decisiones basada en datos.

Existen varias herramientas para mejorar la toma de decisiones en un negocio. El Cuadro de Mando (CM) es una de ellas y “es una herramienta de acción a corto plazo de implementación rápida y estrechamente ligado a los puntos claves de decisión y de responsabilidad de la empresa” (López Viñeglas, 1999). Los cuadros de mando son herramientas de extrema utilidad para la toma de decisiones en cualquier escenario, debido a que permiten apreciar el comportamiento tendencial de los principales indicadores que caracterizan los procesos de una institución. En el mundo empresarial este concepto se ha generalizado en los últimos años convirtiéndose en una herramienta útil para dirigir, medir resultados y reevaluar procedimientos en las organizaciones. En Cuba los modelos de gestión de muchas entidades fueron concebidos para un entorno más sencillo y estable al que se presenta hoy día, debido al proceso de informatización de la sociedad en el que se encuentra inmersa, por tanto, estas están forzadas a implementar sistemas de gestión que respondan a las exigencias actuales y futuras de dichas instituciones.

La XETID juega un papel primordial para fomentar el avance tecnológico en el país. La empresa centra la atención de sus productos en los sectores públicos, en especial el sector de la defensa. La misma cuenta con varias áreas de negocio, entre estas se encuentra el Sistema de Apoyo a la Toma de Decisiones (SATD). Este centro desarrolla herramientas que permiten obtener y visualizar oportunamente la información requerida durante el proceso de la toma de decisiones. Además, les permite a las entidades reutilizar de forma inteligente la información que está almacenada, reflejando equilibrio entre los objetivos a corto y a largo plazo de las mismas. Como parte del proceso de informatización en dicha entidad, es necesario crear productos que favorezcan el desarrollo de sus propias áreas, por tal motivo este centro tiene la tarea de

INTRODUCCIÓN

proveerle a la Dirección Comercial de la XETID un sistema que les permita realizar análisis multidimensionales y análisis de tendencias de los indicadores comerciales de la empresa. Dentro de los procesos que se manejan en la Dirección Comercial se encuentra el relacionado con la contratación, el cual cuenta con un sistema de gestión que presenta una Base de Datos (BD) relacional.

El sistema transaccional existente no facilita la búsqueda de información histórica con agilidad, ni la realización de comparaciones con la totalidad de los datos significativos en un determinado intervalo en el tiempo, debido a la forma en que está organizada dicha información, por lo que existe lentitud a la hora de tomar decisiones. Además, necesitan mejorar su capacidad de análisis y reducir el tiempo que se tarda en recopilar la información o mejorar el seguimiento y gestión de su estrategia. Hoy en día los especialistas no tienen la posibilidad de obtener reportes que muestren los datos históricos almacenados sobre la contratación. De este modo se impide la realización de análisis estadísticos complejos entre diferentes indicadores, análisis históricos y de tendencias, lo que conlleva a la pérdida de datos útiles y valiosos para la toma de decisiones.

A partir de estas necesidades se plantea el siguiente **problema a resolver**: ¿Cómo contribuir al análisis multidimensional sobre el comportamiento tendencial de los indicadores de desempeño de la Dirección Comercial de la XETID?

La investigación tiene como **objeto de estudio** la inteligencia de negocio, enmarcado en el **campo de acción** los Cuadros de Mando basados en indicadores de desempeño.

El **objetivo general** de esta investigación consiste en desarrollar un Cuadro de Mando para contribuir al análisis multidimensional sobre el comportamiento tendencial de los indicadores de desempeño de la Dirección Comercial de la XETID.

Para darle cumplimiento al objetivo definido se plantearon los siguientes **objetivos específicos**:

1. Establecer los referentes teóricos asociados a la inteligencia de negocio relativa al análisis multidimensional sobre el comportamiento tendencial de los indicadores de desempeño, a través de los cuadros de mando.
2. Realizar la identificación de requisitos, análisis y diseño del almacén de datos como aproximación a la implementación.
3. Implementar el almacén de datos y configurar los tableros de control, para contribuir al análisis multidimensional sobre el comportamiento tendencial de los indicadores de

INTRODUCCIÓN

desempeño de la Dirección Comercial.

4. Realizar pruebas de software que permitan validar el correcto funcionamiento de la solución.

Para dar respuesta a la interrogante y a los objetivos trazados, es necesario el cumplimiento de las siguientes **tareas de la investigación**:

- Realización de un estudio sobre el desarrollo de cuadros de mando para la gestión comercial.
- Realización de un estudio sobre el diseño y desarrollo de almacenes de datos.
- Selección de las tecnologías, herramientas y estándares que se necesitan para implementar la propuesta de solución.
- Selección y estudio de la metodología de desarrollo a utilizar.
- Análisis de los requerimientos.
- Análisis de las fuentes OLTP.
- Elaboración del modelo lógico de la estructura del almacén de datos.
- Desarrollo del proceso de extracción, transformación y carga.
- Diseño de los tableros de control del Cuadro de Mando.
- Validación de los tableros de control del Cuadro de Mando.

Para dar solución a las tareas antes expuestas se definen los siguientes métodos científicos:

Métodos teóricos:

- **Histórico-lógico:** permitió realizar un estudio crítico sobre la evolución de los diferentes enfoques relativos a tableros de control y las tecnologías empleadas para su implementación.
- **Analítico-sintético:** permitió el análisis y procesamiento de la información recopilada durante la investigación sobre los Almacenes de Datos, metodologías, tecnologías y herramientas de desarrollo.

Métodos empíricos:

- **Entrevista:** se le realizaron entrevistas al cliente con el objetivo de comprender el proceso de negocio, recolectar datos como la estructura y funcionamiento de la organización, indicadores a medir, así como las deficiencias existentes que permitieron definir el problema a resolver y establecer el objeto de estudio

INTRODUCCIÓN

- **Análisis documental:** este método se utilizó en la revisión de la literatura especializada, tanto académica como empresarial, para extraer la información necesaria que permitió realizar el proceso de investigación.

El presente documento se encuentra estructurado en los 3 capítulos que a continuación se describen:

Capítulo 1. Fundamentación teórica: en este capítulo se definen aspectos relacionados con la fundamentación teórica, con el objetivo de obtener los elementos teóricos para la realización de la investigación. Además, se especifican los conceptos asociados al dominio del problema y se realiza un estudio de soluciones similares con el objetivo de definir sus características y la viabilidad de su uso en la solución de la problemática existente. Se define la metodología y se caracterizan las principales técnicas y herramientas que serán utilizadas durante el desarrollo de la solución.

Capítulo 2. Propuesta de solución: en este capítulo se presenta la propuesta de solución a través de la identificación de las necesidades de información del cliente. A partir de los requisitos se obtienen los artefactos correspondientes a las disciplinas de Análisis y Diseño, teniendo como guía la metodología seleccionada. Se realiza el diseño del modelo multidimensional propuesto a partir de los indicadores seleccionados. Además, se desarrolla el subsistema de visualización mediante una herramienta que permite consultar la información almacenada y analizarla mediante tablas de datos y gráficas.

Capítulo 3. Validación de la propuesta de solución: en este capítulo se lleva cabo un proceso de validación de requisitos a través de un conjunto de pruebas de volumen, carga y estrés, con el fin de garantizar la calidad de la solución propuesta.

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

En el presente capítulo se abordan los conceptos de interés asociados al dominio del problema y se brinda una panorámica de los aspectos relacionados con sistemas homólogos para la toma de decisiones. Se analizan las diferentes metodologías para el desarrollo de Amacenes de Datos (AD) existentes a nivel mundial, además de la definida en la XETID, haciendo la selección de la que regirá el proceso de desarrollo. Además, se profundiza en las características de las herramientas de implementación de un AD. Se define las herramientas que se utilizan para el proceso de desarrollo de solución, dejando plasmado los elementos relacionados con la elección de cada una de ellas. De igual forma se exponen las principales características del sistema gestor de bases de datos seleccionado.

1.1 Definición de Cuadro de Mando

Un Cuadro de Mando (CM), también conocido como tablero de control, es una herramienta de gestión que facilita la toma de decisiones y que recoge un conjunto coherente de indicadores que proporcionan a la alta dirección y a las funciones responsables una visión comprensible del negocio o de su área de responsabilidad. La información aportada por el cuadro de mando, permite enfocar y alinear los equipos directivos, las unidades de negocio, los recursos y los procesos con las estrategias de la organización (Norma UNE 66175: 2003, 2019)

Diferencias entre Cuadro de Mando Integral y Cuadro de Mando

El Cuadro de Mando Integral (CMI) conocido también como Balanced Scorecard- BSC, es un sistema y una herramienta por medio de la cual una organización logra ejecutar su plan estratégico trasladando el mismo a la acción. Su importancia radica en que –por medio de sus indicadores- muestrea de manera continua cuándo la organización y sus colaboradores alcanzan las metas del plan estratégico. El CM, en cambio, es una herramienta operativa útil para controlar lo que pasó en un área o departamento. Sus indicadores están focalizados a procesos y no al avance del plan estratégico de la organización. Es un instrumento útil para medir el avance de los resultados operativos en una organización, pero es insuficiente para lograr la ejecución de un plan estratégico. Por eso, es incorrecto denominarlos de la misma manera. Un CMI está compuesto por un mapa estratégico, formado por objetivos que provienen de la visión de la organización. Un cuadro de mando, en cambio, es un cuadro operacional, motivo por el cual sus objetivos no provienen de la estrategia sino de los diferentes programas y planes que tiene que

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

cumplir cada área o departamento de una organización. Los elementos que componen un CM son: objetivos operativos del área o departamento e indicadores de gestión que muestran lo que pasó durante un período. (Comando, 2020)

El CM de la Dirección Comercial de la XETID se desarrollará a partir de un almacén de datos con el objetivo de reducir los tiempos de respuestas y los costes de operaciones. Además de esta manera se podrán realizar análisis de los datos históricos y observar los comportamientos tendenciales del área comercial, proporcionando de esta manera la información clave para la toma de decisiones.

1.2 Introducción a los Almacenes de datos

Resulta de gran interés analizar e interpretar la información que se genera en una empresa, con el fin de contar con elementos que le permitan tomar decisiones para maximizar sus ganancias o mejorar su rendimiento. Con el surgimiento y desarrollo de las tecnologías, el volumen de información generada ha ido aumentando considerablemente, dando lugar a grandes cantidades de datos históricos, cuyos análisis resultan cada vez más complejos. Los AD conocidos por sus siglas en inglés como Data Warehouse, surgen como una solución a esta problemática, con el fin de apoyar el proceso de toma de decisiones en las empresas.

Bill Inmon, conocido como el padre de los AD, definió este término como *“la colección de datos, orientados a materias, integrados, cambiantes con el tiempo y no volátiles, para la ayuda al proceso de toma de decisiones de la dirección de una empresa”* (Inmon, 1992)

Según Ralph Kimball en la segunda edición de su libro *The Datawarehouse Toolkit*, *un AD es una copia de los datos transaccionales específicamente estructurada para la consulta y el análisis* (Kimball, 2002).

Entre las principales características de un AD se pueden mencionar las siguientes (Sinnexus Business intelligence, 2019):

- Orientado a temas: esta característica del AD permite responder a preguntas concretas de análisis de negocios acerca de un tema de interés. Con tema de interés, se refiere, a las áreas de la organización sobre las que se requiere hacer un análisis, como podrían ser las ventas, compras, productos, entre otros. Por ejemplo, para una compañía

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

manufacturera sus áreas de interés podrían ser: productos, órdenes, costos de materiales y materias primas.

- Integrado: de todos los aspectos del AD, éste es el más importante, y está estrechamente relacionado con la característica de orientado a temas. Consiste en poner en un formato consistente los datos provenientes de diversas fuentes, resolviendo problemas tales como: conflictos de nombres, inconsistencias en unidades de medidas, etc. Cuando todo lo anterior se ha llevado a cabo, entonces, se dice que está integrado.
- No volátil. esta característica significa que, una vez entrado los datos al almacén, estos no cambian, debido a que el propósito de un AD es permitir el análisis de los datos de lo que ha ocurrido. Los datos del AD son almacenados (en grandes cantidades) y accedidos, pero no actualizados.
- De tiempo variante. esta característica se enfoca en los cambios sobre el tiempo para descubrir tendencias e identificar patrones ocultos (sobre grandes cantidades de datos), útiles para los analistas de negocios. La variación en el tiempo implica que cada unidad de dato del AD es exacta para algún momento en el tiempo.

Después de analizar las definiciones y las características anteriormente planteadas se concluye que un AD es un conjunto de datos históricos seleccionados de distintas fuentes y coleccionados con el objetivo de ser analizados en un determinado momento para contribuir a la toma de decisiones estratégicas y tácticas de una organización.

1.2.1 Diferencias entre AD y una Base de Datos Operacional

El primer objetivo de una base de datos es registrar en tiempo real los datos con los que se alimenta, mientras que el AD está diseñado para realizar agregaciones de datos con fines analíticos. Los usuarios que acceden a los datos operacionales, comúnmente efectúan tareas predefinidas y actualizables que requieren acceso a una sola base de datos desde una aplicación y almacenan datos actuales. Por el contrario, los usuarios que acceden al almacén de datos, efectúan tareas complejas que requieren acceso a un conjunto de datos desde fuentes múltiples, frecuentemente no son predecibles y almacenan datos históricos. (Aguilar, 2011) (Anexo # 1)

1.3 Sistemas homólogos

Luego de haber analizado los principales conceptos asociados al dominio del problema, se realiza un estudio de cuadros de mando comerciales similares, con el objetivo de definir sus características y la viabilidad de su uso en la solución de la problemática existente.

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

BSC Designer: software privativo para Cuadro de Mando Integral desarrollado por la empresa Kítonik S.R.O. en la República Checa. (BSC Designer, 2020)



Principales características y funcionalidades:

- Gestión de mapas estratégicos con objetivos de negocios representados
- *KPI*¹ de actuación y de resultado, normalización de las métricas y cálculo del cuadro de mando.
- Iniciativas y planes de acción que incluyen detalles para la ejecución de objetivos.
- Notificaciones para mantener a los miembros del equipo informados respecto a las incidencias de la empresa.
- Reportes de rendimiento con planes de acción, datos de KPI y gráficos importantes.

Olucaro: es una herramienta de software privativa desarrollada en España que facilita la creación y gestión de CM. Permite representar de forma gráfica, y por áreas, todos los objetivos e indicadores de una empresa. (Olucaro Dashboard, 2020)



¹ KPI: son indicadores claves de desempeño que se utilizan para medir el grado de cumplimiento de los objetivos previamente establecidos. (Comenzando de cero, 2020)

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

Principales características y funcionalidades

- Aplicación web accesible desde cualquier dispositivo
- Permite la Importación y exportación de documentos en formatos de Excel.
- Consta de un Chat online que facilita la comunicación entre los distintos factores de la Empresa
- Facilita la compartición de documentos entre los miembros del equipo.
- Gestión de notificaciones que mantienen alerta a los miembros del equipo.

Corporater: es un software privativo que proporciona un conjunto de herramientas para la administración empresarial entre las que se encuentran la gestión de mapas estratégicos, CM, objetivos estratégicos, KPI e iniciativas estratégicas. Se encuentra desplegado en empresas de más de diez países tales como Estados Unidos, Reino Unido, Sudáfrica, Alemania y los Emiratos Árabes Unidos. (Corporate Business, 2020)



Principales características y funcionalidades

- Planificación y ejecución estratégica
- Gestión del portafolio de proyectos
- Gestión de indicadores claves de rendimiento (adelantos y retrasos)

SmartBPO: es una herramienta de software privativa desarrollada en España que dispone de un CM en el que figuran los indicadores más importantes de la labor comercial de una empresa. Haciendo uso de este software las empresas pueden obtener una visión general de la evolución de determinados indicadores claves relacionados con las ventas por periodos de tiempos, por zona geográfica, por departamento y por referencias de productos. (SmartBPO, 2020)

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA



Principales características y funcionalidades

- Panel de análisis con una visión general de la evolución de determinados indicadores clave relacionados con las ventas.
- Panel para analizar la red de ventas y conocer la evolución de las misma por comisionistas o agentes
- Panel permite visualizar el estado de las facturaciones emitidas en un período determinado y el importe pendiente de cobro.

ABAST: es un software privativo desarrollado por la empresa española *Abast Systems & Solutions S.L.* Esta solución se centra en un cuadro de mando con el que es posible visualizar de forma intuitiva los principales indicadores relacionados con la gestión comercial de las empresas y realizar análisis detallados para agilizar la toma de decisiones. (Abast Systems & Solutions S.L., 2020)



Principales características y funcionalidades

- Navegación sobre *Google Maps*² integrada en el Cuadro de Mando sobre las distintas sedes de la organización. (Abast Systems & Solutions S.L., 2020)

² Google Maps: es un servicio gratuito de Google de mapas a través de la web. (Google, 2019)

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

- Navegación sobre las diferentes zonas especificadas en un mapa, zonas coloreadas según cumplimentación de objetivos.
- Simulación de escenarios para cada vendedor.
- Visualización integrada vía RSS³ de contenidos actualizados con afectación para las ventas.

Conclusiones del estudio de los sistemas homólogos

El análisis anterior permitió estudiar algunas de las soluciones de CM más significativas a nivel global teniendo en cuenta aspectos tales como tipo de software (libre / privativo), país de desarrollo y características generales.

Como resultado se pudo apreciar que las soluciones estudiadas son de carácter privativo y se concentran en países desarrollados del primer mundo como España, Estados Unidos y Reino Unido. Como rasgo común la mayoría de dichas soluciones son aplicaciones webs que cuentan con sistemas de notificaciones, mecanismos de chats, mapas digitales y otras características que facilitan la gestión de indicadores estratégicos de las empresas.

Al ser soluciones extranjeras de software privativo no es posible tomarlas como parte de la solución de la presente investigación pues, de alguna manera, iría en contra de las políticas de soberanía tecnológicas que persigue la XETID como Empresa Estatal Socialista. Por otra parte, sería muy complejo lograr una integración de la solución obtenida en la Plataforma de Inteligencia de Negocios Damix perteneciente a la XETID.

1.4 Modelado multidimensional de los AD

Los AD se diseñan mediante un modelo multidimensional compuesto por dimensiones y hechos a diferencia de los sistemas clásicos de bases de datos que presentan estructuras diseñadas mediante un *Modelo Entidad Relación*⁴. Este modelo posee la misma información que el de

³ RSS (Really Simple Syndication): es una forma estandarizada de distribución de la información de las páginas web a los lectores de páginas.

⁴ Modelo Entidad Relación: es una herramienta que permite simplificar los procesos que interactúan en un negocio. (ESIC Business & Marketing School, 2018)

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

entidad-relación, pero la organiza de forma diferente para garantizar la velocidad y eficiencia en la recuperación de la misma.

Dimensiones: las dimensiones definen como están los datos organizados lógicamente y proveen el medio para analizar el contexto del negocio. Representan los ejes del *cubo*⁵ y los aspectos de interés mediante los cuales el usuario podrá filtrar y manipular la información almacenada en los hechos. (Reategui, 2016)

Hechos: los hechos son datos instantáneos en el tiempo, que son filtrados, agrupados y explorados a través de condiciones definidas en las dimensiones. Los datos presentes en los hechos constituyen el volumen del AD, y pueden estar compuestos por millones de registros dependiendo de su granularidad y de los intervalos de tiempo de los mismos. Los más importantes son los de tipo numérico. El registro del hecho posee una clave primaria que está compuesta por las claves primarias de las dimensiones relacionadas a este. (Reategui, 2016)

Indicadores: son aquellos conceptos cuantificables que permiten medir nuestro proceso de negocio. Son características cualitativas o cuantitativas de los objetos que se desean analizar en las empresas. Las medidas, siempre son numéricas, se almacenan en las tablas de hechos.

1.5 Tipos de modelamiento de AD

Esquema en estrella: este esquema se caracteriza por tener una tabla de hechos a la cual llegan las tablas de dimensiones relacionadas a esta a través de sus respectivas claves primarias. De esta manera concentra la información en una tabla de hechos. (Aguilar, 2011) (Ver Anexo # 2)

Esquema copo de nieve: este esquema tiene la particularidad de que una tabla de dimensiones puede estar formada por más de una tabla de datos lo que le hace diferente a el esquema de estrella permitiéndonos tener varios caminos para llegar a los datos influyendo directamente sobre el rendimiento de las consultas (Arturo, 2012).(Ver Anexo # 2)

Esquema constelación: este modelo está compuesto por una serie de esquemas en estrella, donde existe más de una tabla de hechos. Dichas tablas yacen en el centro del modelo y están relacionadas con sus respectivas tablas de dimensiones, en ocasiones dos o más hechos

⁵ Cubos: son los conceptos principales del negocio.

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

comparten una misma dimensión, aunque no es necesario que las diferentes tablas de hechos compartan las mismas tablas de dimensiones. (IBARRA, 2019)(Ver Anexo # 2)

Teniendo como antecedente las características de los distintos tipos de modelamientos, se selecciona para utilizar en el presente trabajo el esquema en estrella. El uso de este tipo de modelamiento está condicionado debido a que se adapta fácilmente a las necesidades específicas del cliente, pues toda la información está concentrada en una sola tabla de hechos.

1.6 Metodologías para el diseño e implementación de un AD.

Se entiende por metodología de desarrollo, una colección de documentación formal referente a los procesos, las políticas y los procedimientos que intervienen en el desarrollo del software. La finalidad de una metodología de desarrollo es garantizar la eficacia (p.ej. cumplir los requisitos iniciales) y la eficiencia (p.ej. minimizar las pérdidas de tiempo) en el proceso de generación de software. (Blanco-Cuaresma, 2019)

Metodología R. Kimball

La metodología se basa en lo que Kimball denomina Ciclo de Vida Dimensional del Negocio (Business Dimensional Lifecycle). Este ciclo de vida del proyecto de AD, está basado en cuatro principios básicos (Hernández, 2013):

- Centrarse en el negocio: hay que concentrarse en la identificación de los requerimientos del negocio y su valor asociado, y usar estos esfuerzos para desarrollar relaciones sólidas con el negocio, agudizando el análisis del mismo y la competencia consultiva de los implementadores.
- Construir una infraestructura de información adecuada: diseñar una base de información única, integrada, fácil de usar, de alto rendimiento donde se reflejará la amplia gama de requerimientos de negocio identificados en la empresa.
- Realizar entregas en incrementos significativos: crear el AD en incrementos entregables en plazos de 6 a 12 meses. Hay que usar el valor de negocio de cada elemento identificado para determinar el orden de aplicación de los incrementos. En esto la metodología tiene similitud con las metodologías ágiles de construcción de software.

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

- Ofrecer la solución completa: proporcionar todos los elementos necesarios para entregar valor a los usuarios de negocios. Esto significa tener un AD sólido, bien diseñado, con calidad probada, y accesible.

Metodología Hefesto

La metodología Hefesto fue creada por el Ing. Bernabeu Ricardo Dario en su libro "HEFESTO: Metodología para la Construcción de un Data Warehouse". Permite la construcción de un AD de forma sencilla, ordenada e intuitiva. Hefesto es una metodología bien fundamentada y explícita que permite desarrollar un AD de manera metódica y sencilla, guiándose por pasos lógicos relacionados sólidamente durante todas las etapas del proceso de confección. La ventaja principal de esta metodología es que especifica puntualmente los pasos a seguir en cada fase a diferencia de otras metodologías que mencionan los procesos, más no explican cómo realizarlos. (Bernabeu, 2010)

Metodología de W. Inmon

W. Inmon define su metodología en el año 1992 en el libro "Building the Data Warehouse", donde propone los mecanismos necesarios para llevar a cabo la correcta realización de un AD. A Inmon se le asocia con los AD a nivel empresarial, que involucran desde un inicio todo el ámbito corporativo, sin centrarse en un área específica hasta después de haber terminado completamente el diseño del AD. Afirma, que la creación de una *base de datos relacional*⁶ con una ligera normalización, es la base de los MD, o lo que es lo mismo, a partir de los esquemas relacionales, a los que se les irán añadiendo complejidad, se obtendrán finalmente los MD (Hernández, 2013)

Después de analizar las metodologías antes mencionadas se selecciona como metodología de desarrollo a utilizar en este trabajo un híbrido entre la metodología de Hefestos y la de Kimball debido a que esta es la metodología adoptada por el Centro de Datos (CD) de la Universidad de Ciencias Informáticas (UCI) adaptada a las necesidades propias de la empresa XETID para el desarrollo de almacenes de datos. Se decidió realizar un híbrido entre estas dos metodologías y no adoptar una de las dos con el objetivo aprovechar las facilidades que brinda Hefestos de una

⁶ Base de datos relacional: es una recopilación de elementos de datos con relaciones predefinidas entre ellos. (Services, 2020)

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

metodología sencilla y entendible con las de Kimball de generar artefactos para la documentación a la hora de realizar un AD.

Es importante destacar que algunos de los artefactos adoptados por el CD fueron modificados pues se determinó que no hacía falta documentar todo lo que este requería. Además, la metodología rectora fue la de Hefestos y no Kimball. Teniendo en cuenta esto se puede decir que las características de este híbrido de metodología son (Xetid, 2012):

- Los objetivos y resultados esperados en cada fase se distinguen fácilmente y son sencillos de comprender.
- Se basa en los requerimientos del usuario, por lo cual su estructura es capaz de adaptarse con facilidad y rapidez ante los cambios en el negocio.
- Reduce la resistencia al cambio, ya que involucra al usuario final en cada etapa para que tome decisiones respecto al comportamiento y funciones del AD.
- Utiliza modelos conceptuales y lógicos, los cuales son sencillos de interpretar y analizar.
- Es independiente del tipo de ciclo de vida que se emplee para contener la metodología.
- Es independiente de las herramientas que se utilicen para su implementación.
- Es independiente de las estructuras físicas que contengan el AD y de su respectiva distribución.
- Cuando se culmina con una fase, los resultados obtenidos se convierten en el punto de partida para llevar a cabo el paso siguiente.
- Se aplica tanto para AD como para Mercados de Datos⁷.

Esta metodología, al igual que la de Hefestos, cuenta con las mismas 4 fases de desarrollo, la diferencia radica en que se le han incorporado algunos pasos en cada una de estas fases y se han adoptado y creado artefactos para documentar lo que se realiza en cada fase, lo cual sirve para la reutilización del conocimiento en otros proyectos (Xetid, 2012).

1.7 Sistemas OLTP y OLAP.

Los sistemas de Procesamiento Transaccional y Analítico en Línea conocidos como OLTP y OLAP respectivamente, son conceptos de indispensable análisis para un mejor entendimiento y

⁷ Mercado de Datos: es una forma simple de un AD, ya que se especializa en el almacenamiento de los datos de un tema o área específica del negocio.

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

posterior desarrollo de los Data Warehouse, es de suma importancia reconocer cuando se está en presencia de uno u otro sistema. (Lisbet Marrero, 2015)

1.7.1 Procesamiento Transaccional en Línea (OLTP).

Los sistemas OLTP son bases de datos orientadas al procesamiento de transacciones. Una transacción genera un proceso atómico (que debe ser validado o invalidado), y que puede involucrar operaciones de inserción, modificación y borrado de datos. El proceso transaccional es típico de las bases de datos operacionales. (Sinnexus, 2020)

Las características principales de los sistemas OLTP son: (Angel González, 2011)

- Realizan transacciones en tiempo real del proceso de un negocio, con lo cual los datos almacenados cambian continuamente. Los sistemas OLTP en sus transacciones conducen procesos esenciales del negocio.
- Los sistemas OLTP son los responsables del mantenimiento de los datos, ya sea agregando datos, realizando actualizaciones o bien eliminándolos.
- Las estructuras de datos deben estar optimizadas para validar la entrada de los mismos, y rechazarlos si no cumplen con determinadas reglas de negocio.
- Para la toma de decisiones, proporciona capacidades limitadas ya que no es su objetivo, por lo tanto no es prioridad en su diseño. Si se quisiera obtener determinada información histórica relativa al negocio consultando un sistema OLTP, se produciría un impacto negativo en el funcionamiento del sistema

1.7.2 Procesamiento Analítico en Línea (OLAP).

Los sistemas OLAP (On-Line Analytical Processing) son bases de datos orientadas al procesamiento analítico. Este análisis suele implicar, generalmente, la lectura de grandes cantidades de datos para llegar a extraer algún tipo de información útil: tendencias de ventas, patrones de comportamiento, elaboración de informes complejos, etc. Este sistema es típico de los Mercados de Datos. (Sinnexus, 2020)

Las siguientes son características que los sistemas OLAP poseen: (Lisbet Marrero, 2015)

- Las bases de datos OLAP tienen un esquema que está optimizado para que las preguntas realizadas por los usuarios sean respondidas rápidamente.

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

- Las preguntas que se le hacen a un OLAP, deben permitir un uso interactivo con los usuarios.
- Un sistema OLAP está preparado para realizar informes complejos de una manera simple.
- OLAP proporciona una vista de datos multidimensional. Los cubos proporcionan una vista de los datos multidimensional que se extiende más allá del análisis de dos dimensiones que puede proporcionar una simple planilla de cálculo utilizada como tal.
- Los usuarios pueden cambiar fácilmente las filas, las columnas, y las páginas en informes de OLAP, pudiendo leer la información de la manera que se crea más conveniente para el análisis.

Existen varias arquitecturas para los sistemas OLAP como:

MOLAP (OLAP Multidimensional), ROLAP (OLAP Relacional) y HOLAP (OLAP Híbrido). MOLAP (Multidimensional On-line Analytical Process), estas herramientas utilizan bases de datos multidimensionales para proporcionar el análisis, su principal premisa es que el OLAP está mejor implantado almacenando los datos multidimensionalmente. El almacenaje de MOLAP, provee excelente rendimiento y compresión de datos. El sistema MOLAP utiliza una arquitectura de dos niveles: la bases de datos multidimensionales y el motor analítico. La base de datos multidimensional es la encargada del manejo, acceso y obtención del dato. El motor analítico es el responsable de la ejecución de los requerimientos OLAP y de proporcionar una interfaz a través de la cual los usuarios finales visualizan los análisis OLAP. (Lisbet Marrero, 2015)

ROLAP (Relacional On-line Analytical Process) arquitectura OLAP que crea vistas multidimensionales extrayendo los datos de bases de datos relacionales, simulan los datos multidimensionales usando sofisticadas técnicas de indexación. ROLAP utiliza una arquitectura de tres niveles. La base de datos relacional maneja el almacenamiento de los datos, el motor OLAP proporciona la funcionalidad analítica y alguna herramienta especializada es empleada para el nivel de presentación. (Molier, 2005.)

HOLAP. Un desarrollo un poco más reciente ha sido la solución OLAP híbrida (HOLAP), la cual combina las arquitecturas ROLAP y MOLAP para brindar una solución con las mejores características de ambas: desempeño superior y gran escalabilidad. Un tipo de HOLAP mantiene los registros de detalle (los volúmenes más grandes) en la base de datos relacional, mientras que mantiene las agregaciones en un almacén MOLAP separado. (Sinnexus, 2020)

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

Una vez analizadas las características de cada una de las arquitecturas, se propone utilizar la arquitectura ROLAP para la construcción del AD de la empresa XETID, debido a que los datos se extraen de una base de datos relacional soportada por un sistema gestor de base de datos PostgreSQL. Además esta arquitectura ofrecerá un sistema escalable para grandes volúmenes de datos que soportará en un futuro la empresa.

1.8 Diseño del sistema de Extracción, Transformación y Carga

El sistema de Extracción, Transformación y Carga por sus siglas en inglés (ETL) es la base sobre la cual se alimenta el almacén de datos. ETL es el proceso que organiza el flujo de los datos entre diferentes sistemas en una organización y aporta los métodos y herramientas necesarias para mover datos desde múltiples fuentes a un almacén de datos, reformatearlos, limpiarlos y cargarlos en otra base de datos, mercado de datos o almacén de datos para analizar. (Aguilar, 2011)

La primera parte del proceso ETL consiste en localizar y extraer los datos desde los sistemas de origen. Cada sistema separado puede usar una organización diferente de los datos o formatos distintos. La extracción convierte los datos a un formato homogéneo y consolidado para iniciar el proceso de transformación.

La fase de transformación de un proceso de ETL aplica una serie de reglas de negocio o funciones sobre los datos extraídos para convertirlos en datos que serán cargados, como pueden ser: filtrado, limpieza, depuración, homogeneización, agrupación de la información y seleccionar campos específicos que se consideren necesarios para el almacén de datos. Por último la fase de carga que es la que se encarga de organizar y actualizar los datos en los mercados de datos.

1.9 Herramientas utilizadas para el desarrollo del AD

Para el desarrollo de AD es necesario elegir correctamente las herramientas a utilizar, basado en las empleadas por el equipo de desarrollo de SATD para la creación de los mismos, se seleccionaron las que a continuación se detallan teniendo como premisa que se podrá dar solución al objetivo de la investigación.

1.9.1 Herramienta de modelado

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

Las herramientas de modelado de sistemas informáticos, son herramientas que se emplean para la creación de modelos de sistemas que ya existen o que se desarrollarán. Permiten concentrarse en ciertas características importantes del sistema, prestando menos atención a otras. Los modelos resultados, son una buena forma de determinar si están representados todos los requerimientos del software, como también saber si el analista comprendió qué hará el sistema. (ALEGSA, 2020)

Visual Paradigm for UML Community Edition v8.0: es una plataforma de modelado, diseñada para apoyar a los arquitectos de sistemas, desarrolladores y diseñadores UML (*Unified Modeling Language*⁸) para acelerar el proceso de análisis y diseño de aplicaciones empresariales complejas. La UCI cuenta con licencia para su uso, por lo que se contribuye al cumplimiento de las políticas de migración hacia software libre que sigue nuestro país.

1.9.2 Sistema Gestor de Base de Datos

Un Sistema Gestor de Base de Datos (SGBD) es un sistema que permite la creación , gestión y administración de bases de datos , así como la elección y manejo de las estructuras necesarias para el almacenamiento y búsqueda de la información del modo más eficiente posible. (Marin, 2019)

PostgreSQL 10.1: es un sistema gestor de base de datos objeto relacional libre, debido a que incluye características de la orientación a objetos, como puede ser la herencia, tipos de datos, funciones, restricciones, disparadores, reglas e integridad transaccional, liberado bajo la licencia BSD⁹

PgAdmin 1.22: es una aplicación de diseño y manejo de bases de datos para su uso con PostgreSQL. Está diseñado para responder a las necesidades de todos los usuarios, desde escribir consultas SQL sencillas hasta el desarrollo de bases de datos complejas. La interfaz

⁸ UML (*Unified Modeling Language*): Lenguaje Unificado de Modelado que propicia un conjunto de ayudas para el desarrollo de programas informáticos.

⁹ Licencia BSD: licencia de software otorgada principalmente para los sistemas BSD (Berkeley Software Distribution). Es una licencia de software libre permisiva. La licencia BSD permite el uso del código fuente en software no libre (SBD, 2016)

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

gráfica soporta todas las características de PostgreSQL y hace fácil administración. Es un software libre publicado bajo la licencia PostgreSQL. (PgAdmin., 2019)

1.9.3 Herramienta para el proceso de Extracción, Transformación y Carga

Pentaho Data Integration Community Edition (PDI) v 4.2: está desarrollado íntegramente en *Java*¹⁰ bajo licencia *LGPL*¹¹. Se utiliza para la integración de datos, carga de almacenes de datos y mercados de datos, limpieza de datos, análisis y perfilado de datos, migración de datos entre bases de datos y exportar datos de bases de datos a archivos planos y viceversa. Transforma e integra datos entre sistemas de información existentes y los mercados de datos que compondrán el sistema de *inteligencia de negocio*¹².

1.9.4 Herramienta para el proceso de Inteligencia de negocio

Pentaho Schema Workbench 3.10: esta herramienta de la suite Pentaho tiene como objetivo facilitar la tarea de diseño de cubos OLAP. Su interfaz permite modelar un *XML*¹³ con el diseño del cubo a través de opciones lógicas e intuitivas que no requieren de un manejo avanzado de este formato de archivo. Dentro de sus características se destacan:

- Diseñador intuitivo de esquemas OLAP
- Permite crear, editar, actualizar y publicar esquemas OLAP para ser desplegados por aplicaciones de visualización Pentaho.
- Acelera de manera considerable la construcción e implementación de este tipo de soluciones.

1.9.5 Herramienta para el análisis multidimensional

Damix es una herramienta que permite visualizar la información resumida de una forma sencilla facilitando la toma de decisiones estratégicas mediante un cuadro de mando o tablero de control que muestra vistas en tablas dinámicas, mapas temáticos y gráficas, así como un sistema de

¹⁰ Java: es un lenguaje de programación

¹¹ LGPL (Lesser General Public Lincence) es una licencia de software libre.

¹² Inteligencia de negocio: es la combinación de tecnología, herramientas y procesos que permiten transformar los datos almacenados en información. (Services O., 2020)

¹³ XML: formato simple basado en texto para representar la información de manera estructurada. (Servicio de Administración Tributaria, 2020)

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

planificación estratégica y representa la relación causa_efecto entre los objetivos estratégicos a indicadores que se visualizan en el tablero de control. (XETID, 2019)

1.10 Conclusiones parciales

El capítulo expuso una panorámica general del proceso de desarrollo de los AD, así como el estudio de las tendencias actuales en cuanto a la metodología de desarrollo, las tecnologías y herramientas que se utilizarán en este Trabajo de Diploma. Luego de esta investigación se arribaron a las siguientes conclusiones:

- Se realizó un estudio de las principales metodologías existentes para el desarrollo de los almacenes de datos, el cual permitió identificar la metodología híbrida como la más adecuada para la realización de los mismos, debido a que cubre las fases por las que pasa la construcción de un AD y brinda ventajas que facilitan su desarrollo. Además, está adaptada a las necesidades de la XETID.
- Se conocieron los distintos tipos de modelamiento y según las características de cada uno se decidió utilizar el esquema en estrella, debido a que este modelo es soportado por casi todas las herramientas de consulta y análisis, y los metadatos son fáciles de documentar y mantener
- Se identificaron las herramientas para el desarrollo de los almacenes de datos. Teniendo en cuenta sus características y funcionalidades se puede asegurar que son adecuadas para implementar de manera eficiente la solución propuesta, para satisfacer las necesidades del cliente.

CAPÍTULO 2. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE LA SOLUCIÓN

En el presente capítulo se describe el proceso de análisis, diseño e implementación del AD para el desarrollo del Cuadro de Mando para la Dirección Comercial de la XETID, siempre tomando en consideración las etapas y pasos que brinda la metodología de desarrollo escogida en el capítulo anterior. Se hace una descripción detallada de cada una de las fases de dicha metodología, partiendo inicialmente de la identificación de los requerimientos de información del usuario, hasta llegar a la definición de los procesos ETL para poblar completamente el AD.

2.1 Descripción de la propuesta de solución

Para dar solución al problema planteado se propone el desarrollo de un cuadro de mando para la Dirección Comercial, para así poder satisfacer las necesidades y carencias con que cuenta esta área.

La siguiente imagen muestra de manera gráfica los pasos generales que se llevarán a cabo durante el desarrollo de la solución.

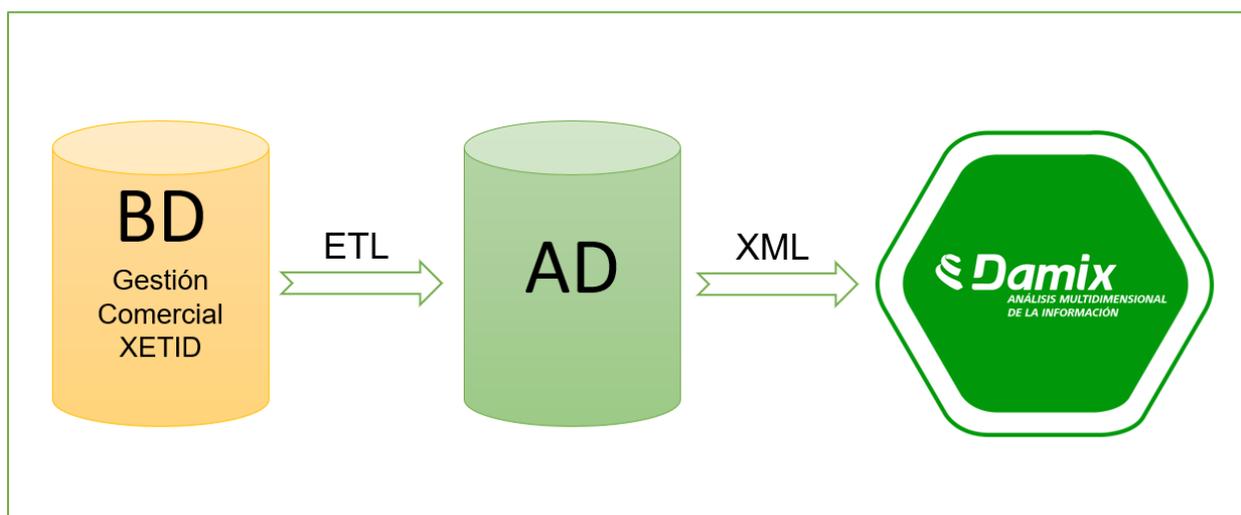


Figura 1: Diseño de la propuesta de solución

Fuente: Elaboración propia

- 1- El proceso comienza a partir del análisis de la base de datos del sistema de Gestión Comercial de la XETID.
- 2- Luego en la herramienta PDI se desarrollan los procesos de Extracción, Transformación y Carga obteniendo como resultado el almacén ya poblado.

CAPÍTULO 2. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE LA SOLUCIÓN

- 3- Posteriormente en la herramienta Pentaho Schema Workbench se desarrolla el esquema, que no es más que un fichero en formato XML que le traduce a la plataforma Damix la estructura del almacén.
- 4- Finalmente, en la plataforma Damix se configuran los análisis de la información que se quieren mostrar y a su vez los tableros de control que conforman el CM.

De esta manera quedará elaborado el CM para la Dirección Comercial de la XETID, el cual permitirá realizar análisis multidimensionales relacionados con los contratos recepcionados en la empresa, permitiendo así poder tomar decisiones oportunas. Además posibilitará observar el comportamiento de los indicadores claves de la actividad comercial en la entidad, brindando la posibilidad de obtener la cantidad, monto y promedio del monto de todos los contratos que se recepcionen en la empresa.

2.2 Descripción de las fases de la Metodología Híbrida

La metodología consta de 4 fases y cada una de ellas con una serie de pasos que guiarán todo el proceso de desarrollo.

2.2.1 Fase 1: Análisis de requerimientos.

En esta primera fase se comienza planificando cada una de las entrevistas con el cliente, para recolectar las necesidades de información de los usuarios y así obtener las preguntas claves del negocio que permitan cumplir los objetivos trazados. Luego se analizan las preguntas a fin de identificar cuáles son los indicadores y perspectivas que se van a tener en cuenta para el desarrollo del almacén de datos. Posteriormente se confecciona un modelo conceptual para visualizar el resultado obtenido durante las entrevistas. Esta fase se concluye con la validación del mismo por el cliente, donde se obtendrá un acta de validación para poder continuar con las demás fases.

Identificar preguntas claves

Para el desarrollo del segundo paso de la metodología se realizó una entrevista a el jefe del departamento de Comercial de la División de Inteligencia de Negocio con el propósito de obtener las preguntas claves, los resultados que se esperan obtener y los reportes que considere importantes.

CAPÍTULO 2. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE LA SOLUCIÓN

De la entrevista realizada se obtuvieron las necesidades de información del cliente, a continuación, se muestran estos requerimientos agrupados para su estudio teniendo en cuenta el interés del cliente.

- 1- Se desea realizar un análisis acerca de la cantidad de contratos con clasificación (X) del sector (X) con estado (X) del proveedor o cliente (X) y tipo de contrato (X) en un tiempo (inicio-fin) (X).
- 2- Se desea realizar un análisis acerca del monto de los contratos con clasificación(X) del sector (X) con estado (X) del proveedor o cliente (X) y tipo de contrato (X) en un tiempo (inicio-fin) (X).
- 3- Es de interés realizar un análisis acerca del promedio del monto de contratos con clasificación (X) del sector (X) con estado (X) del proveedor o cliente (X) y tipo de contrato (X) en un tiempo (inicio-fin) (X).

Identificar indicadores y perspectivas de análisis

De las entrevistas realizadas se obtuvieron las preguntas claves para el negocio, y a partir de las mismas se obtuvieron los siguientes indicadores con sus correspondientes perspectivas.

Tabla 1. Indicadores y perspectivas

INDICADORES	PERSPECTIVAS
Cantidad de contratos	Clasificación de contrato, sector, estado, tipo de contrato, proveedores/clientes, tiempo(inicio-fin)
Monto de contratos	Clasificación de contrato, sector, estado, tipo de contrato, proveedores/clientes, tiempo(inicio-fin)
Promedio de monto de contratos	Clasificación de contrato, sector, estado, tipo de contrato, proveedores/clientes, tiempo(inicio-fin)

Construcción del modelo conceptual

CAPÍTULO 2. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE LA SOLUCIÓN

Una vez identificados los indicadores y perspectivas se confeccionó el modelo conceptual de los datos, en el que se reflejaron las perspectivas a la izquierda y los indicadores a la derecha, como se muestra en la Figura 2, proporcionando una idea del alcance del AD.

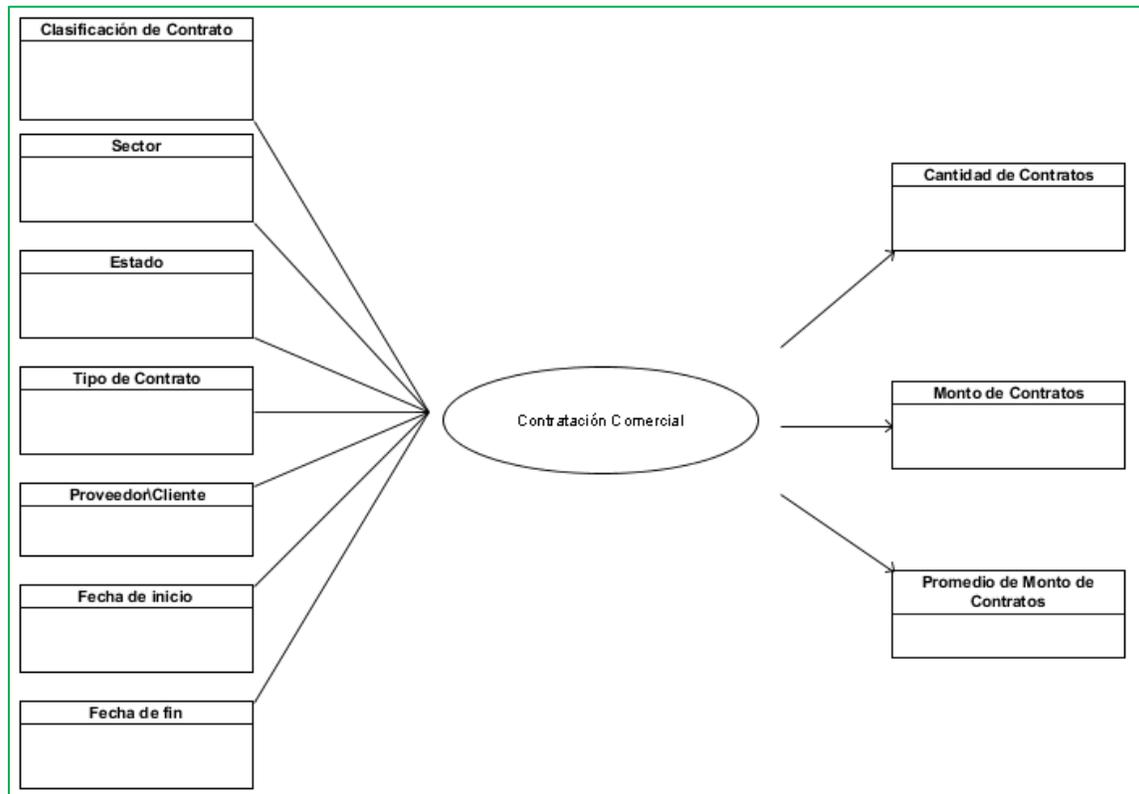


Figura 2. Modelo Conceptual

Fuente: Elaboración propia

Aprobación del Modelo Conceptual

Para asegurar que el modelo conceptual se construyó correctamente y que reflejó lo que el cliente necesitaba, se aprobó por el mismo en una reunión donde se expuso el Modelo Conceptual, utilizando siempre un lenguaje entendible para ambas partes. Luego se dio paso a la firma del acta de validación, donde quedó reflejado el mutuo acuerdo con los resultados por ambas partes.

2.2.2 Fase 2: Análisis de los OLTP

En esta fase primeramente se realiza un análisis de las fuentes OLTP para determinar cómo serán calculados los indicadores y establecer las respectivas correspondencias entre el modelo

CAPÍTULO 2. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE LA SOLUCIÓN

conceptual creado en el paso anterior y las fuentes de datos. Luego se definen los campos que se incluirán en cada perspectiva. Se ampliará el modelo conceptual con la información en este paso, y se finalizará definiendo las reglas del negocio.

Determinación de indicadores

Tabla 2. Determinación de indicadores

INDICADOR	HECHOS	FUNCIÓN DE SUMARIZACIÓN	ACLARACIÓN
Cantidad de contratos	Cantidad de contratos	COUNT	El indicador Cantidad de contratos representa la cantidad de contratos formalizados por la dirección comercial de la Xetid.
Monto de contratos	Monto de contrato	SUM	El indicador Monto de contratos representa la sumatoria de los montos de los contratos formalizados por la dirección comercial de la Xetid.
Promedio de monto de contrato	AVG (Monto de contratos)	AVG	El indicador Promedio de monto de contratos representa el promedio de los montos de los contratos formalizados por la dirección comercial de la Xetid

Fuente: Elaboración propia

Establecer correspondencias

En este paso se establecen las correspondencias entre el modelo conceptual y los sistemas fuentes, para una mejor especificación y entendimiento del AD. Para el desarrollo de la investigación se requirió del uso de la base de datos de la Dirección Comercial de la XETID como sistema fuente de la investigación. Este proceso queda documentado en la planilla “Especificación de las correspondencias”.

CAPÍTULO 2. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE LA SOLUCIÓN

Nivel de granularidad

Una vez establecidas las relaciones entre los sistemas fuentes y el modelo conceptual definido y luego de determinar la forma en que se calculan los indicadores, en este paso se procede a definir los campos que contendrán cada perspectiva, ya que será a través de estos por donde se filtrarán y manipularán los diferentes indicadores. En la Tabla 3 se muestra el resultado obtenido:

Tabla 3. Nivel de granularidad

TABLAS	CAMPOS	PERSPECTIVAS
Clasificación de contratos	<i>Id, denominacion</i>	Dimensión clasificación de contratos
Sector	<i>Id, denominacion</i>	Dimensión sector
Estado	<i>Id, denominacion</i>	Dimensión estado
Tipo de contrato	<i>Id, denominacion</i>	Dimensión tipo de contrato
Proveedor/Cliente	<i>Id, denominacion</i>	Dimensión proveedor o cliente
Fecha de inicio	<i>Id, ano, trimestre, mes, semana</i>	Dimensión Tiempo
Fecha de fin	<i>Id, ano, trimestre, mes, semana</i>	Dimensión Tiempo

Fuente: Elaboración propia

Modelo Conceptual Ampliado

Con el fin de mostrar los resultados obtenidos en los pasos anteriores, se ampliará el modelo conceptual colocando debajo de cada perspectiva los campos escogidos y debajo de cada indicador su fórmula de cálculo, quedando este debidamente documentado en la planilla “Modelo Conceptual Ampliado”. A continuación, en la Figura 3 se muestra el modelo conceptual ampliado.

CAPÍTULO 2. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE LA SOLUCIÓN

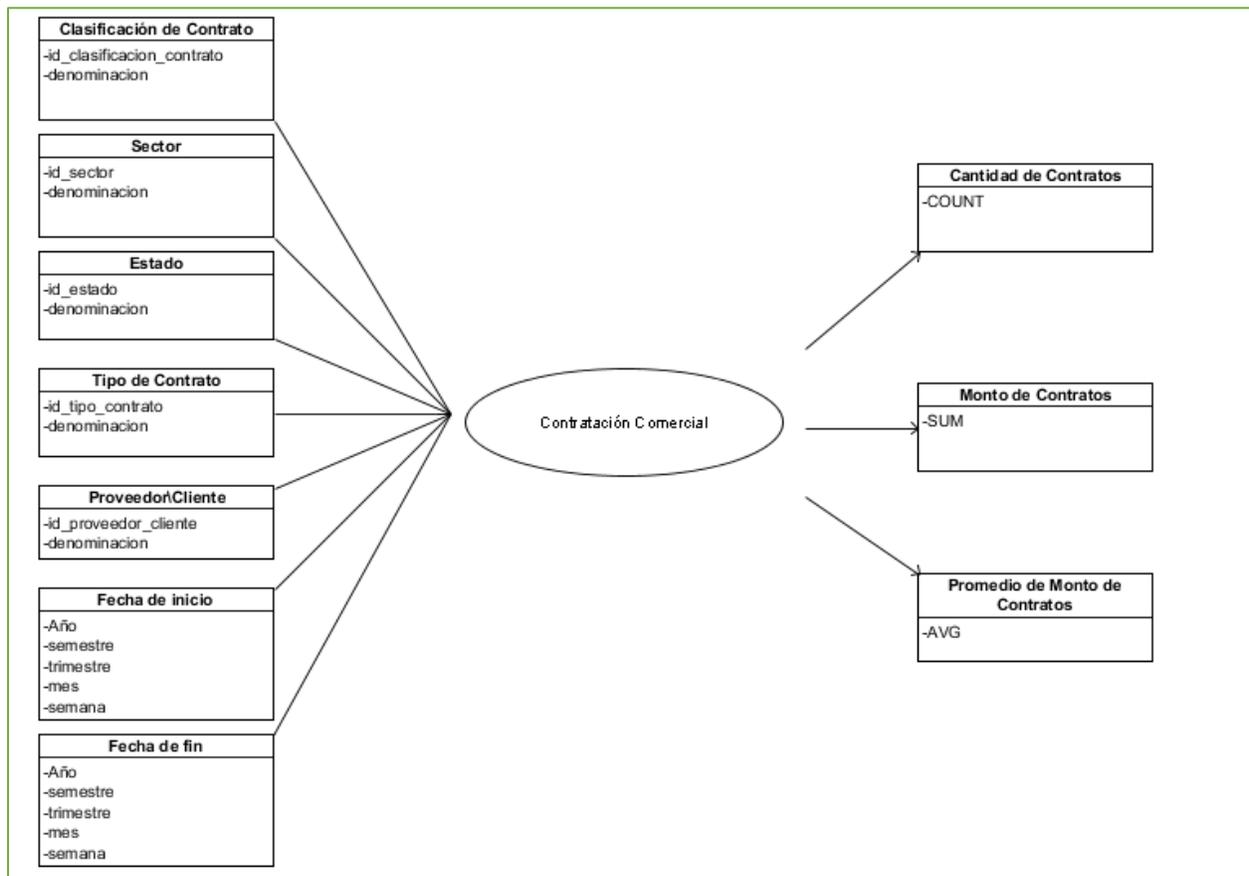


Figura 3. Modelo Conceptual Ampliado

Fuente: Elaboración propia

Definir reglas del negocio

Este proceso queda documentado en la planilla "Reglas del negocio". Según esta planilla lo que se realizó en la empresa fue lo siguiente:

Reglas del negocio:

1. Los identificadores de los indicadores no pueden estar repetidos.
2. Los identificadores de los indicadores no pueden ser nulos.
3. Los identificadores de las dimensiones no pueden ser nulos.
4. Los valores que indiquen cantidad tienen que ser mayores o igual que cero.

2.2.3 Fase 3: Modelo Lógico del AD

CAPÍTULO 2. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE LA SOLUCIÓN

En esta fase se confecciona el modelo lógico de la estructura del AD teniendo como base el Modelo Conceptual creado anteriormente. Para ello primero se definirá el tipo de modelo que se utilizará, luego se llevarán a cabo las acciones propias al caso, para diseñar las tablas de dimensiones y de hechos. Se realizarán las uniones pertinentes entre estas tablas y finalmente se diseñarán las tablas y columnas físicas del AD.

Definir tipo de modelo lógico del AD

En esta investigación se decidió utilizar el esquema en de estrella debido a que se puede dar solución a los requerimientos actuales del cliente con una única tabla de hechos.

Definición de estándares para objetos físicos

Luego de escoger el modelo lógico que se utilizará en el AD, se procede a definir algunos estándares físicos a tener en cuenta para la construcción del mismo, tales como la nomenclatura de la BD, así como las que tendrá las tablas de dimensiones y hechos. El equipo de desarrollo siempre que construya un AD deberá utilizar el estándar que propone el comité de roles de arquitecto de datos en aras de ayudar a la uniformidad y entendimiento común del equipo.

Identificar dimensiones

En este paso se diseñan las tablas de dimensiones que formarán parte del AD. Cada perspectiva definida en el modelo conceptual constituirá una tabla de dimensiones.

Se define como estándar de codificación para las dimensiones la estructura "dim_nombre", para las tablas de hecho "cub_nombre" y para los campos se separan las palabras que se vayan a incluir con guión bajo. Se utiliza el patrón llaves subrogadas para el tratamiento de los identificadores en las dimensiones. Los campos "id_nombre_perspectiva" almacenan las llaves primarias de las tablas de origen asociadas a cada perspectiva. A continuación, en la Tabla 4 se muestra el resultado de este proceso.

Tabla 4. Dimensiones del Almacén de Datos

DIMENSIÓN	CAMPOS
dim_clasificacioncontrato	Id_clasificacioncontrato, denominacion

CAPÍTULO 2. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE LA SOLUCIÓN

dim_estado	id_estado, denominacion
dim_proveedorcliente	id_proveedorcliente, denominacion
dim_tiempo	id_tiempo, fecha, ano, semestre, trimestre, mes, semana, dia, numsemestre, numtrimestre, nummes, numsemana, numdia
dim_sector	id_sector, denominacion
dim_tipocontrato	id_tipocontrato, denominacion

Fuente: Elaboración propia

Identificar hechos

En este paso se definen las tablas de hechos, que son las que contendrán los hechos a través de los cuales se construirán los indicadores de estudio.

Tabla 5. Hechos del Almacén de Datos

HECHOS	INDICADORES Y LLAVES
cub_contratacioncomercial	Montocontrato, id_dimensiones_asociadas

Fuente: Elaboración propia

Realizar uniones entre dimensiones y hechos

El campo “id_dimensiones_asociadas” representa a los id de cada dimensión relacionada a la tabla de hecho. La llave primaria de la tabla de hechos la constituyen la unión de todas estas llaves y “montocontrato” es la constante donde se va a almacenar el valor resultante de los requisitos.

CAPÍTULO 2. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE LA SOLUCIÓN



Figura 4. Modelo físico del AD

Fuente: Elaboración propia

Diseñar tablas y columnas físicas

En este paso quedarían diseñadas e implementadas físicamente las tablas de hechos y dimensiones en el AD según el modelo lógico.

2.2.4 Fase 4: Proceso ETL

Una vez construido el modelo lógico e implementado en el AD, se procede a poblarlo con datos a través de los procesos ETL. Primeramente, se realizará un mapeo de los datos fuentes hacia los destinos detallando claramente de donde salen y hacia donde se dirigen, luego se establecerán algunas restricciones y condiciones adicionales en caso que lo requiera. Una vez definidas estas transformaciones se procede a la carga incremental de los datos hacia las

CAPÍTULO 2. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE LA SOLUCIÓN

dimensiones y hechos según corresponda. Finalmente se diseñará y construirá la automatización de todo el proceso.

Mapeo de datos Fuente-A-Destino

Tabla 6. Mapeo de datos

FUENTE DE DATOS			DESTINO		
Denominación de la fuente de datos: Gestion_empresarial_dev			Denominación del destino de los datos: DwhContratacion		
TABLA	CAMPO	TIPO DE DATO	TABLA	CAMPO	TIPO DE DATO
sector	name	varchar	mod_contratacion. dim_sector	denominacion	varchar
estado	id	numeric	mod_contratacion. dim_estado	denominacion	varchar
contrato	id	numeric	mod_contratacion. dim_contrato	denominacion	varchar
clasificación_contrato	name	varchar	mod_contratacion. dim_clasificacionc ontrato	denominacion	varchar
proveedorcliente	name	varchar	mod_contratacion. dim_proveedorocli ente	denominacion	varchar
tipo_contrato	name	varchar	mod_contratacion. dim_tipocontrato	denominacion	varchar

Fuente: Elaboración propia

Cargas incrementales de los datos

En este paso se realiza la carga de los datos almacenados en la fuente de datos para ser guardados en las dimensiones correspondientes. Para esto se llevó a cabo una serie de transformaciones y trabajos con la herramienta Pentaho Data Integration.

Diseño, construcción de la automatización del sistema de ETL

CAPÍTULO 2. ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE LA SOLUCIÓN

En este proceso se llevaron a cabo una serie de pasos los cuales tuvieron como objetivo fundamental, el traspaso de datos desde el Sistema de Contratación Comercial al sistema para la toma de decisiones. La aplicación ETL Pentaho Data Integration fue la encargada de leer los datos primarios, realizar transformaciones, hacer un proceso cualitativo y una correcta filtración para en un final tener dichos datos disponibles.

2.3 Conclusiones parciales

En el presente capítulo se llevó a cabo el proceso de análisis y diseño del AD que va a dar solución a las interrogantes actuales del cliente. Para la realización del AD se utilizó como guía el diseño de la metodología híbrida teniendo en cuenta sus 4 fases, así como el desarrollo del Cuadro de mando a través de la plataforma Damix. Luego de esta investigación se arribaron a las siguientes conclusiones:

- El empleo de la metodología híbrida permitió organizar el desarrollo de la solución y generar los productos de trabajo necesarios correspondientes a la etapa de análisis y diseño.
- La utilización de la entrevista como vía para obtener e identificar las necesidades de información claves de alto nivel, sirvió como guía en la investigación para llevar a cabo las metas y estrategias de la entidad.
- El uso de la herramienta Damix permitió visualizar de una forma sencilla la información resumida a través de los elementos (tablas dinámicas, gráficas) que conforman el cuadro de mando, optimizando la inmediatez de la información y ganando en independencia.

CAPÍTULO 3. VALIDACIÓN Y PRUEBAS

En este capítulo se realiza la validación de los requerimientos con el objetivo de verificar si estos describen la solución que se desea. Se evalúa el grado de calidad y fiabilidad de los resultados obtenidos en el desarrollo de la presente investigación. Además, se aplican pruebas de software para verificar la calidad del CM y satisfacer las necesidades del cliente.

3.1 Pruebas de software

Las pruebas del software son la actividad más común de control de la calidad realizada en los proyectos para asegurar el correcto funcionamiento del software. Tienen como objetivos la verificación de la correcta implementación de los requisitos explícitamente establecidos, la adecuada integración de los componentes que conforman el sistema y la ejecución de casos de prueba que permitan detectar el mayor número de no conformidades (NC) y corregirlas antes de la entrega del software al cliente. (XETID, 2013)

Las pruebas de software comprenden el conjunto de actividades que se realizan para identificar posibles fallos de funcionamiento, configuración o usabilidad de un programa o aplicación, por medio de pruebas sobre el comportamiento del mismo. (Testing Colombia, 2016)

3.2 Diseño de casos de prueba

Un diseño de casos de pruebas se describe como un conjunto de entradas, condiciones de ejecución y resultados esperados desarrollados para determinar si el requisito de una aplicación está parcial o completamente satisfecho. Su propósito es especificar una forma de probar el sistema que incluya las entradas, los resultados esperados y las condiciones bajo las que ha de probarse. (XETID, 2012)

Para realizar el proceso de pruebas es necesario el análisis y diseño de las mismas. En esta actividad los objetivos de las pruebas generales se transforman en condiciones de pruebas y casos de pruebas tangibles. Para ello se ejecutaron las siguientes tareas:

- Revisar la base de pruebas, los informes de análisis de riesgos, la arquitectura, el diseño y las especificaciones de interfaz
- Evaluar la “testabilidad” de la base de pruebas y de los objetos de pruebas
- Identificar y priorizar las condiciones de prueba en base al análisis de los elementos, la especificación, el comportamiento y la estructura del software.
- Diseñar y priorizar los casos de prueba

CAPÍTULO 3. VALIDACIÓN Y PRUEBAS

- Identificar los datos de pruebas necesarios para soportar las condiciones y los casos de pruebas

Teniendo en cuenta lo anterior, para evaluar la calidad del producto obtenido, se realizaron pruebas de caja negra, pruebas funcionales y de aceptación. Las mismas son descritas a continuación.

Pruebas funcionales

Las pruebas funcionales son pruebas diseñadas tomando como referencia las especificaciones funcionales de un componente o sistema (lo que se va a testear, el software o una parte de él). Se realizan para comprobar si el software cumple las funciones esperadas. (Pressman, 2010)

Método de Caja negra

Las pruebas de Caja negra permiten obtener un conjunto de condiciones de entrada que ejerciten completamente todos los requisitos funcionales de un programa. En ellas se ignora la estructura de control, concentrándose en los requisitos funcionales del sistema y ejercitándolos (Pérez, 2016). Estas pruebas permiten encontrar (Pressman, 2010):

- Funciones incorrectas o ausentes.
- Errores de interfaz
- Errores en estructuras de datos o en accesos a las bases de datos externas
- Errores de inicialización y terminación.

Para llevar a cabo el método de Caja negra a continuación se describe la técnica de Partición equivalente a utilizar:

Técnica de prueba: partición equivalente

Esta técnica divide el dominio de entrada de un programa en clases de datos a partir de las cuales pueden derivarse casos de prueba. El método se esfuerza por definir un caso de prueba que descubra ciertas clases de errores, reduciendo así el número total de casos de prueba que deben desarrollarse. Una clase de equivalencia representa un conjunto de estados válidos y no válidos para las condiciones de entrada. (Pressman, 2010)

Para aplicar esta técnica se debe desarrollar los Diseños de casos de pruebas (DCP) con el objetivo de obtener un conjunto de pruebas que tenga la mayor probabilidad de detectar los defectos del software. Un DCP es un conjunto de condiciones o variables bajo las cuales un

CAPÍTULO 3. VALIDACIÓN Y PRUEBAS

analista determinará si una aplicación o una característica de éstos es parcial o completamente satisfactoria (Pressman, 2010).

Con el objetivo de comprobar el funcionamiento de los componentes, se realizaron 3 iteraciones de pruebas para poder alcanzar resultados satisfactorios, obteniéndose los resultados que se muestran en la siguiente figura:

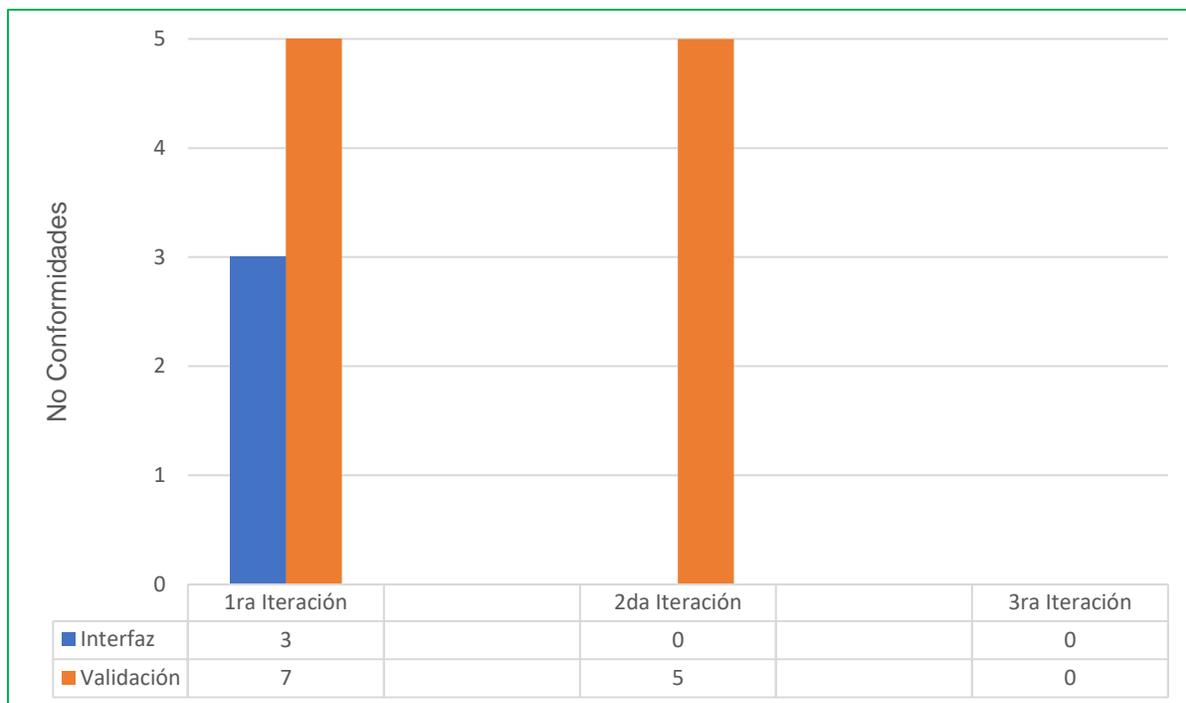


Figura 5. Representación de la cantidad de NC por iteraciones

Fuente: Elaboración propia

En una primera iteración se detectaron un total de 10 NC (7 de validación y 3 de interfaz). En la segunda iteración se manifestaron 5 NC de validación, quedando resueltas las NC de interfaz. Finalmente, en una tercera iteración se obtuvieron resultados satisfactorios al no mostrar NC.

3.3 Pruebas de aceptación

Las pruebas de aceptación (CPA), también llamadas pruebas del cliente, son especificadas por el cliente y se centran en las características y funcionalidad generales del sistema que son visibles y revisables por parte del cliente (Pressman, 2010). Se utilizan para validar que cada

CAPÍTULO 3. VALIDACIÓN Y PRUEBAS

requisito implementado funciona como se había especificado. La Tabla 7 muestra el CPA asociado al indicador 1 Cantidad de contratos, los demás casos de prueba se pueden encontrar en el Anexo # 4.

Tabla 7. Prueba de aceptación para Cantidad de contratos

CASO DE PRUEBA DE ACEPTACIÓN		
Código: 1		
Nombre: Cantidad de contratos		
Descripción: el sistema muestra la cantidad de contratos con clasificación (X) del sector (X) con estado (X) del proveedor o cliente (X) y tipo de contrato (X) en un tiempo (inicio-fin) (X)		
Condiciones de ejecución: el trabajador responsable debe estar registrado en el sistema y contar con los permisos requeridos para poder realizar estas acciones		
ACCIÓN	DATOS DE ENTRADAS	RESULTADO ESPERADO
El especialista selecciona la opción Cantidad de contratos		El sistema muestra la cantidad de contratos según los datos especificados
CE1: El especialista selecciona la opción Cantidad de contratos dejando todos los campos vacíos	Datos nulos	El sistema muestra un mensaje de error “Llene los campos para realizar la operación correctamente”
CE2: El especialista introduce en la opción tiempo un rango de fechas donde la fecha de inicio es mayor que la fecha fin	Datos incorrectos	El sistema valida los datos, al ser incorrectos muestra un mensaje “El rango de fechas es incorrecto”
CE3: El especialista introduce los datos para obtener la cantidad de contratos de manera correcta	Datos correctos	El sistema valida que los datos sean correctos, seguidamente se muestra la cantidad de contratos según los datos especificados
Resultados esperados: el sistema muestra la cantidad de contratos		

CAPÍTULO 3. VALIDACIÓN Y PRUEBAS

Evaluación de la prueba: prueba satisfactoria

Fuente: Elaboración propia

Análisis de los resultados

La figura 6 muestra los resultados obtenidos luego de haber realizado las pruebas de aceptación.

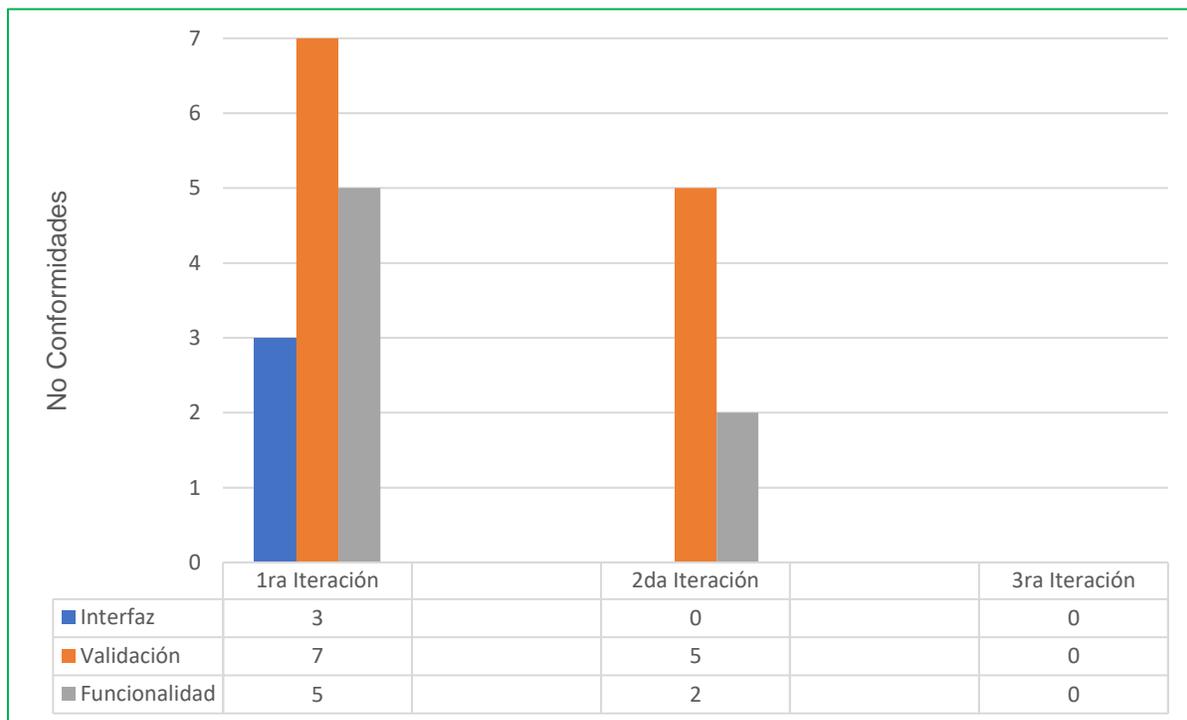


Figura 6. Resultado de prueba de aceptación.

Fuente: Elaboración propia

Se realizaron un total de 3 iteraciones, donde se ejecutaron cada uno de los casos de prueba de aceptación diseñados para cada funcionalidad del sistema. En una primera iteración se encontraron 15 NC divididas en 7 de validación, 3 de interfaz y 5 de funcionalidad. En la segunda iteración se resolvieron las NC de interfaz y se obtuvieron 5 de validación y 2 de funcionalidad. Finalmente fueron resueltas en una tercera iteración obteniéndose resultados satisfactorios de la herramienta, evaluándose de esta manera todos los casos de prueba como satisfactorios.

3.4 Conclusiones parciales

CAPÍTULO 3. VALIDACIÓN Y PRUEBAS

En el presente capítulo se llevó a cabo el proceso de validación y pruebas del CM con el objetivo de verificar que la solución cumple con los requisitos planteados. Luego de esta investigación se arribaron a las siguientes conclusiones:

- Con la utilización de los diferentes tipos de pruebas realizados para validar los procesos de ETL y de Inteligencia de Negocio respectivamente, se logró probar el AD, y así determinar que el sistema cumpliera con los requisitos del cliente.
- La ejecución de pruebas funcionales mediante el método Caja negra demostró que las funciones son operativas a través de la interfaz del software, manteniendo así la integridad de la información externa.
- Los resultados alcanzados durante la fase de pruebas demostraron que se obtuvo un software confiable que satisface los requisitos planteados para el CM.

CONCLUSIONES

Al concluir la investigación para el desarrollo del Cuadro de Mando para la Dirección Comercial de la XETID, se pudo arribar a las siguientes conclusiones:

- El estudio de los referentes teóricos durante la investigación facilitó la adquisición de los conocimientos necesarios para la solución propuesta.
- El análisis de las metodologías y tecnologías para el desarrollo de los mercados de datos permitió identificar a la metodología híbrida (Hefestos + Kimball) como una de metodologías más ágiles y fáciles de usar para personas con poca experiencia en el tema.
- El análisis de los requerimientos del cliente posibilitó la obtención del modelo conceptual del almacén de datos con la información del proceso de contratación a partir de la definición de los indicadores y sus perspectivas de análisis.
- Se realizó el diseño e implementación de la solución utilizando como guía la metodología seleccionada la cual permitió a través de sus fases la construcción del mercado de datos de manera ágil e intuitiva.
- El diseño del cubo de información en la herramienta Schema Workbench, permitió la visualización de los datos deseados por el cliente mediante tablas dinámicas y gráficas, haciendo uso de la herramienta Damix.
- La validación de la solución permitió determinar que el sistema cumplió con los objetivos propuestos por el cliente.

RECOMENDACIONES

A partir del análisis de los resultados alcanzados con la presente investigación se recomienda:

- Ampliar las funcionalidades del almacén de datos teniendo como fuente de información la base de datos del Sistema de Contratación de la XETID.
- Se propone para la implementación de futuros mercados de datos con características similares en otras áreas, utilizar la metodología y las herramientas seleccionadas en la investigación.

REFERENCIAS BIBLIOGRÁFICAS

- (s.f.). Obtenido de <https://blog.powerdata.es/el-valor-de-la-gestión-de-datos/que-es-un-gestor-de-datos-y-para-que-sirve>
- Abast Systems & Solutions S.L. (02 de 2020). Obtenido de <http://www.abast.es/business-intelligence-y-big-data/soluciones-bi-desarrolladas-por-abast/cuadro-de-mando-comercial/>
- Aguilar, Y. N. (2011). *Construcción de un Data Mart para el Sistema Integral de Gestión de Medicamentos*. La Habana: Universidad de Ciencias Informáticas.
- ALEGSA. (15 de febrero de 2020). www.alegsa.com.ar. Obtenido de www.alegsa.com.ar: [en línea] https://www.alegsa.com.ar/Dic/herramienta_de_modelado.php
- Angel González, O. H. (2011). *Diseño e implementación de un mercado de datos para el análisis de eventos*. La Habana: Universidad de Ciencias Informáticas.
- Arturo, B. V. (2012). *Desarrollo de una Aplicación de Business Intelligence para la Empresa Empaqplast. Carrera Ingeniería en Sistemas e Informática*. ESPE. Sede Sangolquí. Disponible en : <http://repositorio.espe.edu.ec/handle/21000/5819>.
- Bernabeu, R. D. (2010). *HEFESTO: Metodología para la Construcción de un Data Warehouse*. Córdoba: s.n.
- Blanco-Cuaresma, S. (15 de diciembre de 2019). *Metodologías de Desarrollo*. [En línea]. Obtenido de <http://www.marblestation.com/?p=644>.
- BSC Designer. (02 de 2020). Obtenido de <https://bscdesigner.com/es/sobre-nosotros>
- Chimbote-Perú, D. d. (2017). *Metodología de Desarrollo de Software*.
- Comando, C. T. (15 de febrero de 2020). www.tablerodecomando.com. Obtenido de [en línea] <https://www.tablerodecomando.com>
- Comenzando de cero. (25 de febrero de 2020). comenzandodecero.com. Obtenido de <https://comenzandodecero.com/que-es-un-kpi>
- Corporate Business. (02 de 2020). Obtenido de <https://corporater.com/en/business-solutions/strategy-management/balanced-scorecard-software/>
- Cuadro de mando comercial*. (15 de noviembre de 2019). Obtenido de Centro de conocimiento: <https://help.cowcrm.com/crm-cliente/gestion-cuentas/cuadro-mando-comercial/>
- Cuadro de Mando Comercial*. (15 de noviembre de 2019). Obtenido de SmartBPO: <http://www.smartbpo.es/cuadro-de-mando-comercial-analitica-de-ventas/>
- ESIC Business & Marketing School. (enero de 2018). www.esic.edu. Obtenido de [en línea] <https://www.esic.edu/rethink/tecnologia/modelo-entidad-relacion-descripcion-aplicaciones>
- Globetesting. (15 de 7 de 2020). www.globetesting.com. Obtenido de [en línea] <https://www.globetesting.com/glosario/pruebas-de-carga/>

REFERENCIAS BIBLIOGRÁFICAS

- Globetesting. (7 de 25 de 2020). *www.globetesting.com*. Obtenido de [en línea] <https://www.globetesting.com/glosario/prueba-de-estres/>
- Google, S. (19 de diciembre de 2019). *sites.google.com*. Obtenido de <https://sites.google.com/sites/rmipasico/-que-es-google-maps>
- Hernández, Y. G. (2013). *METODOLOGÍA DE DESARROLLO PARA PROYECTOS DE ALMACENES DE DATOS*. La Habana: s.n.
- IBARRA, M. d. (diciembre de 15 de 2019). *Procesamiento Analítico en Línea (OLAP)*. Universidad Nacional del Nordeste. 2006,. Obtenido de <http://exa.unne.edu.ar/informatica/SO/OLAPMonog.pdf>
- Inmon, B. (1992). *Building the Data Warehouse*. Wiley: s.l.
- I-Sales Beauty*. (15 de noviembre de 2019). Obtenido de INFCO: http://www.infco.com/wp-content/uploads/2015/06/Folleto_i-Sales_Beauty.pdf
- Kimball, R. (2002). *The Data Warehouse Toolkit: the complete guide to*. New York: John Wiley & Sons.
- Lisbet Marrero, P. P. (2015). "*Mercado de datos para la toma de decisiones a partir de la información de los estudiantes generada por el sistema de Gestión Universitaria y el Entorno Virtual de Enseñanza y Aprendizaje de la Universidad de Ciencias Informáticas*". La Habana.
- López Viñeglas, A. (1999). *El Cuadro de Mando y los Sistemas de Información para la Gestión Empresarial. Posibilidad de Tratamiento Hipermedia*. Madrid: Editora AECA.
- Marin, R. (16 de abril de 2019). *revistadigital.inesem.es*. Obtenido de revistadigital.inesem.es: [en línea] <https://revistadigital.inesem.es/informatica-y-tics/los-gestores-de-bases-de-datos-mas-usados/>
- Microfocus. (febrero de 2017). *www.microfocus.com*. Obtenido de www.microfocus.com: [en línea] <https://www.microfocus.com>
- Molier, F. J. (2005.). *Grupo A y B Informatica :Bloque específico*. Valencia: S.L. Valencia: S.L: EDITORIAL MAD.
- Norma UNE 66175: 2003*. (15 de noviembre de 2019). Obtenido de Academia: www.academia.edu/11089570/Guía_para_la_implantación_de_sistemas_de_indicadores_Norma_UNE_66175_2003
- Olucaro Dashboard. (02 de 2020). Obtenido de <https://www.olucaro.com/olucaro-dashboard/>
- Pérez, C. S. (2016). *Tablero de Control para la Dirección de Posgrado de la Universidad de las Ciencias Informáticas*. La Habana: Universidad de las Ciencias Infomáticas.
- PgAdmin*. (15 de noviembre de 2019). Obtenido de <http://www.pgadmin.org/index.php>
- Polanco Garay, L. W., & Moré Soto, D. (2016). *Guía para el desarrollo de Almacenes de Datos*. La Habana.

REFERENCIAS BIBLIOGRÁFICAS

- Pressman, R. S. (2010). *Ingeniería de Software: Un Enfoque Práctico*. 7ma ed. McGraw Hill.
- Que es gestión de la información*. (16 de abril de 2017). Obtenido de <https://instituciones.sld.cu>
- Reategui, F. U. (2016). <http://es.scribd.com>. [En línea]. Obtenido de <https://es.scribd.com/doc/55234594/Data-Warehousing-Hefesto>.
- SBD. (2016). umh2820.edu.umh.es. Obtenido de umh2820.edu.umh.es: <http://umh2820.edu.umh.es/wp-content/uploads/sites/885/2016/02/Licencia-BSD.pdf>
- Services, A. W. (2020). *Qué es una base de datos relacional*. Obtenido de [aws.amazon.com](https://aws.amazon.com/es/relational-database/): [en línea] <https://aws.amazon.com/es/relational-database/>
- Services, O. (18 de marzo de 2020). www.onegolive.com. Obtenido de [www.onegolive.com](https://www.onegolive.com/faq/guia-practica-de-inteligencia-de-negocio/que-es-inteligencia-de-negocio): <https://www.onegolive.com/faq/guia-practica-de-inteligencia-de-negocio/que-es-inteligencia-de-negocio>
- Servicio de Administración Tributaria. (15 de enero de 2020). www.sat.go.mx. Obtenido de <https://www.sat.go.mx>
- Sinnexus. (17 de febrero de 2020). www.sinnexus.com. Obtenido de Bases de datos OLTP y OLAP: [en línea] http://www.sinnexus.com/business_intelligence/olap_vs_oltp.aspx
- Sinnexus Business intelligence*. (15 de noviembre de 2019). Obtenido de http://www.sinnexus.com/business_intelligence/datawarehouse.aspx.
- SmartBPO. (02 de 2020). Obtenido de <https://www.smartbpo.es/cuadro-de-mando-comercial-analitica-de-ventas/>
- Sommerville, I. (2011). *Libro de ingeniería de software*. Mexico: PEARSON EDUCACIÓN. Recuperado el 1 de 2020
- Testing Colombia*. (18 de 1 de 2016). Obtenido de Testing Colombia: <https://testingcolombia.com>
- XETID. (2012). *Proceso de desarrollo de Software*. La Habana.
- XETID. (2013). *"Proceso de desarrollo de software"*. La Habana.
- XETID. (enero de 2019). www.xetid.cu. Obtenido de [en línea] <https://www.xetid.cu/producto/35>
- Xetid, E. (2012). *Metología UCID para el desarrollo de Data Warehouse*. Habana.

ANEXOS

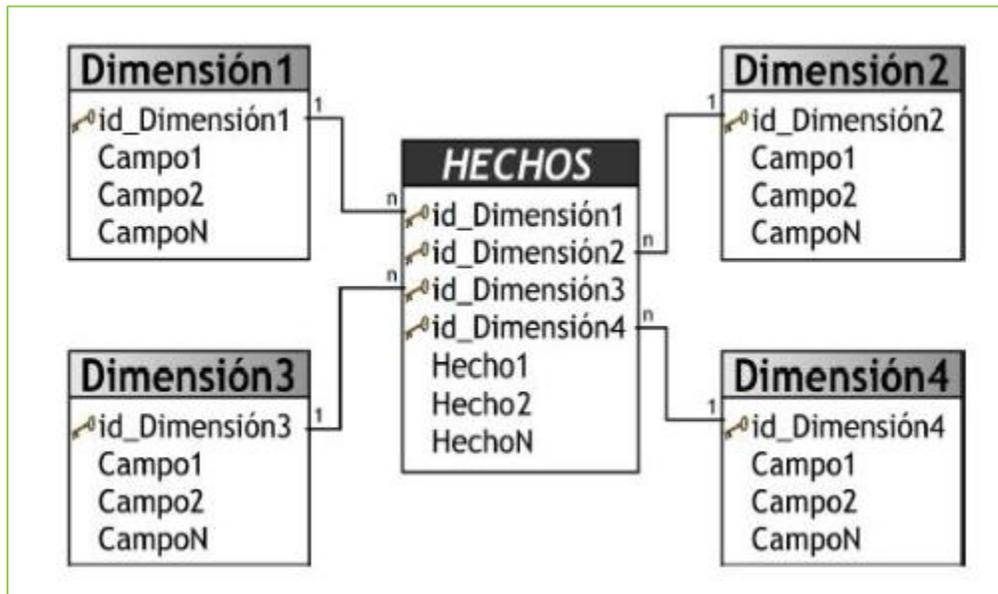
Anexo # 1: Tabla de comparación entre AD y Bases de Datos Operacional

Aspectos	BD operacional	Almacenes de Datos
Objetivo	De tipo operativo (operaciones del día a día)	Análisis y toma de decisiones
Proceso	De transacciones. Repetitivo y conocido.	De consultas masivas. Puntual y no conocido.
Actividad	Predomina la actualización.	Predomina la consulta
Rendimiento	Importancia del tiempo de respuesta de la transacción instantánea	Importancia de la respuesta masiva.
Organización	Organización -Estructura normalmente relacional	Visión multidimensional
Horizonte histórico	30 a 90 días	5 a 10 años
Volumen de datos	Pequeño/medio. Del orden del Mb a Gb.	Medio/grande. Del orden del Gb a Tb.

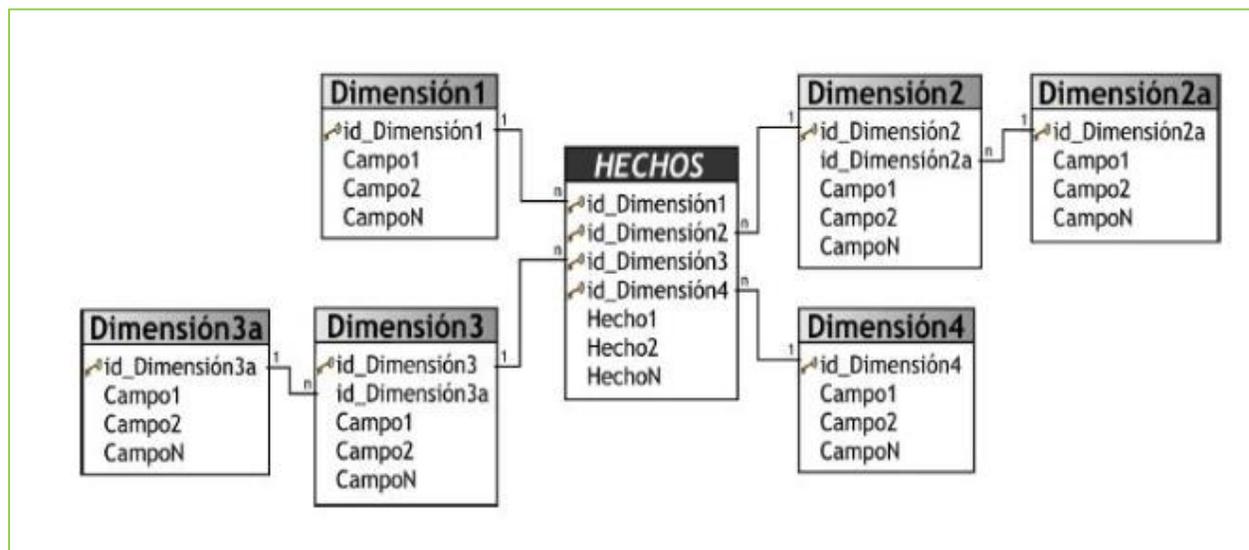
Tabla comparativa entre Base de Datos Operacional y AD

ANEXOS

Anexo # 2: Tipos de modelamiento de los AD

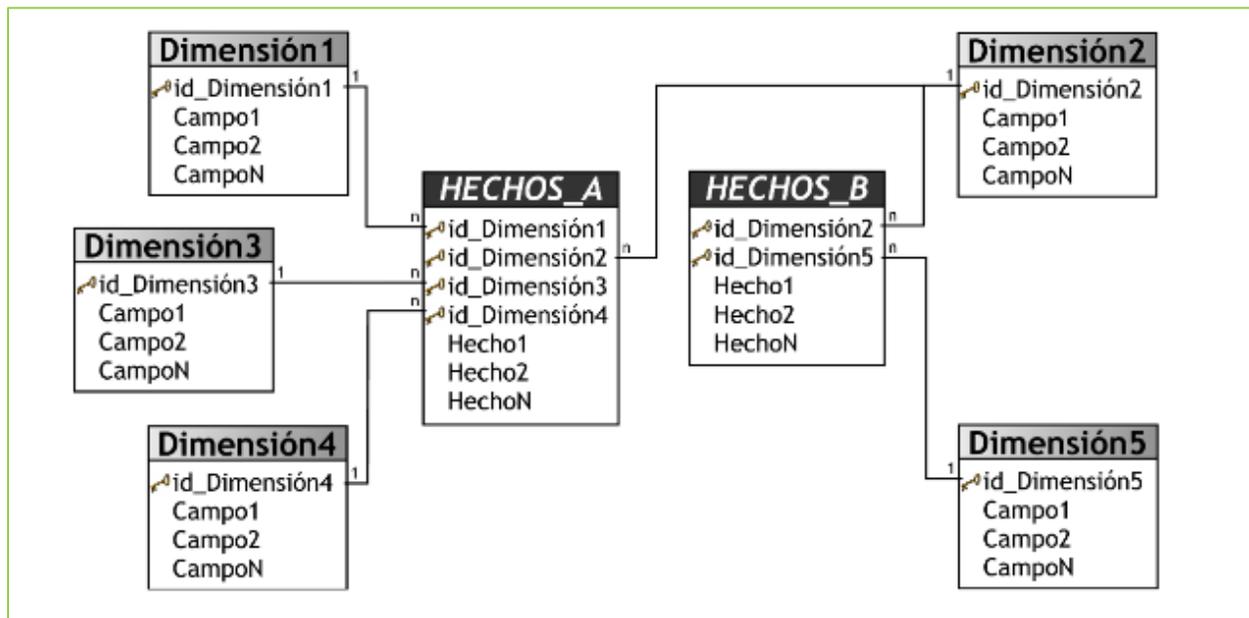


Esquema en estrella



Esquema copo de nieve

ANEXOS



Esquema constelación

Anexo # 3: Fases de la metodología híbrida



Fase 1 de la metodología híbrida



Fase 2 de la metodología híbrida



Fase 3 de la metodología híbrida



Fase 4 de la metodología híbrida

Anexo # 4: Casos de prueba de Aceptación

CASO DE PRUEBA DE ACEPTACIÓN		
Código: 2		
Nombre: Monto de contratos		
Descripción: el sistema muestra el monto de contratos con clasificación (X) del sector (X) con estado (X) del proveedor o cliente (X) y tipo de contrato (X) en un tiempo (inicio-fin) (X)		
Condiciones de ejecución: El trabajador responsable debe estar registrado en el sistema y contar con los permisos requeridos para poder realizar estas acciones		
ACCIÓN	DATOS DE ENTRADAS	RESULTADO ESPERADO
El especialista selecciona la opción Monto de contratos		El sistema muestra el monto de los contratos según los datos especificados

ANEXOS

CE1: el especialista selecciona la opción Monto de contratos dejando todos los campos vacíos	Datos nulos.	El sistema muestra un mensaje de error “Llene los campos para realizar la operación correctamente”
CE2: el especialista introduce en la opción tiempo un rango de fechas donde la fecha de inicio es mayor que la fecha fin	Datos incorrectos.	El sistema valida los datos, al ser incorrectos muestra un mensaje “El rango de fechas es incorrecto”
CE3: el especialista selecciona los campos para obtener el monto de contratos de manera correcta	Datos correctos.	El sistema valida que los datos sean correctos, seguidamente se muestra el monto de los contratos según los datos especificados
Resultados esperados: El sistema muestra el monto de los contratos		
Evaluación de la prueba: Prueba satisfactoria		

CASO DE PRUEBA DE ACEPTACIÓN		
Código: 3		
Nombre: Promedio de monto de contratos		
Descripción: el sistema muestra el promedio de monto de contratos con clasificación (X) del sector (X) con estado (X) del proveedor o cliente (X) y tipo de contrato (X) en un tiempo (inicio-fin) (X).		
Condiciones de ejecución: El trabajador responsable debe estar registrado en el sistema y contar con los permisos requeridos para poder realizar estas acciones.		
ACCIÓN	DATOS DE ENTRADAS	RESULTADO ESPERADO
El especialista selecciona la opción Promedio de monto de contratos		El sistema muestra promedio del monto de los contratos según los datos especificados

ANEXOS

CE1: el especialista selecciona la opción Promedio de monto de contratos dejando todos los campos vacíos	Datos nulos	El sistema muestra un mensaje de error “Llene los campos para realizar la operación correctamente”
CE2: el especialista introduce en la opción tiempo un rango de fechas donde la fecha de inicio es mayor que la fecha fin	Datos incorrectos	El sistema valida los datos, al ser incorrectos muestra un mensaje “El rango de fechas es incorrecto”
CE3: el especialista introduce los datos para obtener Promedio del monto de los contratos de manera correcta	Datos correctos	El sistema valida que los datos sean correctos, seguidamente se muestra promedio del monto de los contratos según los datos especificados
Resultados esperados: el sistema muestra el promedio del monto de los contratos		
Evaluación de la prueba: prueba satisfactoria		