



UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE COMPUTACIÓN
CENTRO DE INVESTIGACIÓN EN SISTEMAS DE INFORMACIÓN

**IMPLEMENTACIÓN DE UN SISTEMA DE
INTELIGENCIA DE NEGOCIOS PARA
EL APOYO EN LA TOMA DE DECISIONES
EN UNA EMPRESAINTEGRADORA DE
SOLUCIONES DE TELECOMUNICACIÓN.**

Trabajo Especial de Grado presentado ante la ilustre
Universidad Central de Venezuela
por los bachilleres:

Carlos M. Buitrago M. C.I. 15.872.514

Juan A. González S. C.I. 17.402.439

para optar al título de:
Licenciado en Computación

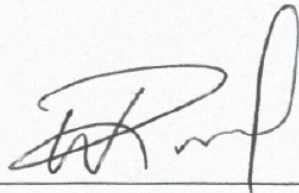
Tutor:
Prof. Wilfredo Rangel

Acta

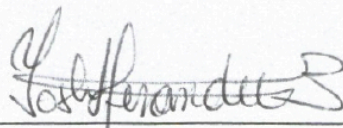
Quienes suscriben, miembros del jurado designado por el consejo de escuela de computación, para examinar el Trabajo Especial de Grado presentado por los bachilleres Carlos M. Buitrago M. y Juan A. González S., portadores de la cédula de identidad número V-15.872.514 y V-17.402.439 respectivamente, con el título: "IMPLEMENTACION DE UN SISTEMA DE INTELIGENCIA DE NEGOCIOS PARA EL APOYO EN LA TOMA DE DECISIONES EN UNA EMPRESA INTEGRADORA DE SOLUCIONES DE TELECOMUNICACIÓN." a los fines de optar al título de licenciado en computación, dejan constancia de lo siguiente:

Leído como fue, dicho trabajo por cada uno de los miembros del jurado, se fijó el día 02 de diciembre de 2016 a las 09:00 am para que los autores lo defiendan en forma pública, lo que se hizo en el salón del Centro de Computación de la Escuela de Computación, mediante una presentación oral de su contenido. Finalizada la defensa pública del Trabajo Especial de Grado, el jurado decidió aprobarlo con la nota de **19** puntos.

En fe de lo cual se levanta la presente Acta, en Caracas a los 02 días del mes de diciembre del año 2016.



Prof. Msc. Wilfredo Rangel – Tutor



Profa. Msc. Yosly Hernández – Jurado



Prof. Esmitt Ramirez – Jurado

Dedicatoria

A todos y cada uno de los miembros de nuestra familia,
por todo el amor, la amistad, la dedicación
y el apoyo que nos han brindado
incondicionalmente a lo largo
de nuestras vidas.

Resumen

Título:

Implementación de un sistema de inteligencia de negocios para el apoyo en la toma de decisiones en una empresa integradora de soluciones de telecomunicación.

Autores:

Carlos M. Buitrago M.

Juan A. González S.

Tutor:

Prof. Msc. Wilfredo A. Rangel A.

El presente Trabajo Especial de Grado tiene como finalidad describir y mostrar los resultados del proceso de implementación de una solución de Inteligencia de Negocios y el impacto logrado de la obtención de indicadores de gestión en una empresa integradora de soluciones de telecomunicación, que presentaba problemas a la hora de generar reportes. A través de esta solución, se generan reportes relacionados a los estados financieros, así como también indicadores que ayudan a la gerencia media y alta con información fundamental para la toma de decisiones. La metodología que gobernará los procesos metodológicos, aplicativos y a los efectos del desarrollo del presente TEG está basada en la de Ralph Kimball. La plataforma tecnológica utilizada como respaldo para el desarrollo de la solución, es la aplicación de Oracle Business Intelligence Standard Edition One. El uso del enfoque metodológico y la aplicación indicada, ofrece como resultado, una herramienta que integra la información vital para la toma de decisiones en la organización.

Palabras Claves: Inteligencia de Negocios, OBIEE, Sistemas de Información

Tabla de Contenido

Acta	2
Dedicatoria	3
Resumen	4
Introducción	11
1. Capítulo I - Problema de la Investigación	12
1.1 Título	12
1.2 Situación Actual	12
1.3 Planteamiento del Problema.....	12
1.4 Solución Propuesta.....	13
1.5 Justificación	14
1.6 Alcance	14
1.7 Objetivo General	15
1.8 Objetivos Específicos.....	15
2. Capítulo II – Marco Conceptual.....	17
2.1 Sistemas de Información	17
2.1.1 Tipos de Sistemas de información	18
2.2 Base de Datos	19
2.2.1 Modelos de Base de Datos	19
2.2.2 Tipos de Base de Datos según su propósito	22
2.2.3 Sistemas Manejadores de Base de Datos	22
2.3 Almacén de Datos.....	22
2.3.1 Arquitectura de un Almacén de Datos	23
2.3.2 Características	23
2.4 Bodega de Datos	24
2.4.1 Características	24
2.5 Diferencias entre Bodega y Almacén de Datos	25
2.6 Extracción, Transformación y Carga (ETL).....	25
2.7 Modelo Dimensional	26
2.7.1 Dimensión	26
2.7.2 Tabla Dimensión	26
2.7.3 Hecho (Medida).....	27
2.7.4 Tabla de Hechos	27
2.7.5 Granularidad (Grano)	28
2.7.6 Jerarquía	28
2.7.7 Agregación	29
2.8 Inteligencia de Negocios.....	29
2.8.1 Características	29
2.8.2 Beneficios.....	29
2.8.3 Tipos de Soluciones.....	30
2.8.4 Arquitectura de una Solución	31
2.8.5 Herramientas de la Inteligencia de Negocios.....	32
2.9 Desarrollo de sistemas	34
2.9.1 Metodologías para el desarrollo de sistemas de Inteligencia de Negocios	35
2.9.2 Kimball Lifecycle	36
2.10 Metodología	38

2.11 Fases de la Investigación	38
2.11.1 Fase I: Planificación y Análisis	40
2.11.2 Fase II: Diseño	40
2.11.3 Fase III: Implementación	42
2.11.4 Fase IV: Despliegue	44
3. Capítulo III - Marco Aplicativo	45
3.1 Fase I: Planificación y Análisis	45
3.1.1 Requerimientos Capturados	46
3.2 Fase II: Diseño	49
3.2.1 Diseño de la aplicación de Inteligencia de Negocio	49
3.2.2 Diseño Técnico de la Arquitectura	49
3.2.3 Selección de productos e instalación	50
3.3 Modelado dimensional	51
3.3.1 Dimensiones	51
3.3.2 Hechos	53
3.4 Fase III: Implementación	55
3.4.1 Desarrollo de la aplicación BI	55
3.5 Diseño físico	55
3.5.1 SMS Generados	56
3.5.2 Servicios 0900	56
3.5.3 Infoline	57
3.5.4 Call center	57
3.5.5 Publicidad	58
3.6 Diseño y desarrollo de los procesos de ETL	59
3.6.1 Origen y destino de los datos	59
3.6.2 Procesos de extracción transformación y carga	59
3.7 Fase IV: Despliegue	94
4. Conclusiones	97
5. Observaciones y recomendaciones	98
6. Referencias Bibliográficas	99

Índice de Figuras

Figura 1.1: Arquitectura de la solución planteada	13
Figura 2.1: Sistema de información. (Laudon K. & Laudon J., 2004).....	17
Figura 2.2: Tipos de Sistema de información. (Laudon K. & Laudon J., 2004).....	18
Figura 2.3: Modelo Jerárquico. (Moraga, 2001)	20
Figura 2.4: Modelo Red. (Moraga, 2001)	20
Figura 2.5: Modelo Relacional. (Moraga, 2001)	21
Figura 2.6: Modelo Orientado a Objetos. (Moraga, 2001).....	21
Figura 2.7: Arquitectura de un almacén de Datos. (EcuRed.CU, 2012)	23
Figura 2.8: Ejemplo: Tabla Dimensión. (Darmawikarta, 2007).....	26
Figura 2.9: Ejemplo: Tabla de Hechos. (Darmawikarta, 2007)	27
Figura 2.10: Sistemas de IN dentro de los tipos Sistemas de Información (Inteligencia de Negocios, 2001).	31
Figura 2.11: Arquitectura de la solución de Inteligencia de Negocios (NASE, 2010)	31
Figura 2.12: Ejemplos de Reportes (SIPEC, 2011).....	33
Figura 2.13: Ejemplos de Indicadores (Vélez, 2000).	33
Figura 2.14: Diagrama de la metodología de Kimball Lifecycle (Kimball, R., 2011).	38
Figura 2.15: Fases del desarrollo utilizando la metodología Kimball Lifecycle	39
Figura 2.16: Back Room y Front Room de un DW. (Kimball & Caserta, 2004).....	41
Figura 2.17: Fases del proceso de extracción transformación y carga	44
Figura 3.1: Arquitectura conceptual del Sistema de Inteligencia de Negocios	50
Figura 3.2: SMS Generados.....	56
Figura 3.3: Servicios 0900.....	56
Figura 3.4: Infoline.....	57
Figura 3.5: Call Center	57
Figura 3.6: Publicidad.....	58
Figura 3.7: Oracle BI Administrator tool	58
Figura 3.8: Proceso Actualizar Almacén diario	61
Figura 3.9: Proceso Parte 1 Extracción Transformación Diaria	62
Figura 3.10: Proceso Dimensión Asociado extracción	63
Figura 3.11: Proceso Dimensión Cliente 0900 extracción	64
Figura 3.12: Proceso Dimensión Compañía extracción	65
Figura 3.13: Proceso Dimensión Numero	66
Figura 3.14: Proceso Dimensión Palabra Clave	67
Figura 3.15: Proceso Dimensión Rubro	68
Figura 3.16: Proceso Dimensión Servicio	69
Figura 3.17: Proceso Dimensión Call Center	70
Figura 3.18: Dim Asociado Carga	71
Figura 3.19: Dim Cliente 0900 Carga	72
Figura 3.20: Dim compañía carga	72
Figura 3.21: Dim Numero Carga	73
Figura 3.22: Dim Palabra Clave Carga	73
Figura 3.23: Dim rubro	74
Figura 3.24: Dim Servicio Carga	74
Figura 3.25: Dim Call Center Carga	75
Figura 3.26: Parte 2 Carga Diaria	76
Figura 3.27: Proceso tabla de hechos 0900.....	77
Figura 3.28: Proceso tabla de hechos SMS Generados	78
Figura 3.29: Proceso tabla de hechos Call Center.....	79

Figura 3.30: Fact 0900 Carga	80
Figura 3.31: Fact SMS Generados.....	80
Figura 3.32: Fact Call Center Carga	81
Figura 3.33: Fact Registro Clientes 0900	81
Figura 3.34: Fact RC900 SMS Calcular Días a Procesar	82
Figura 3.35: Fact RC900 SMS Generar Registro Día a Procesar.....	83
Figura 3.36: Actualizar Almacén Mensual	84
Figura 3.37: Parte1 ETL Mensual	85
Figura 3.38: Proceso Dimensión Ejecutivo de Ventas	86
Figura 3.39: Proceso tabla de hechos Infoline	87
Figura 3.40: Dim Cliente Infoline	88
Figura 3.41: Dim Ejecutivo Ventas	89
Figura 3.42: Parte2-Carga-Mensual	90
Figura 3.43: Fact Publicidad Extracción	91
Figura 3.44: Fact Snapshot Infoline	92
Figura 3.45: Dim Ejecutivo Ventas	93
Figura 3.46: Cuadro de mandos Demo Total de llamadas por operadora y total de SMS por operadora.....	94
Figura 3.47: Clientes nuevos y desafiliados	95
Figura 3.48: Meta	95
Figura 3.49: Catálogo.....	96

Índice de Tablas

Tabla 2.1: Cuadro comparativo de Sistemas Manejadores de Bases de Datos	22
Tabla 2.2: Cuadro comparativo entre Almacén de Datos y Bodega de Datos	25
Tabla 2.3: Metodologías para el desarrollo de Sistemas de IN (Eckerson, s.f.)	36
Tabla 2.4: Ejemplo de la Matriz de Bus Empresarial	37
Tabla 3.1: Plantilla para el registro de actividades a realizar	45
Tabla 3.2: Extracto plantilla para el registro de actividades a realizar	45
Tabla 3.3: Matriz del Bus Empresarial	51
Tabla 3.4: Dimensión Date	52
Tabla 3.5: Dimensión Compañía	52
Tabla 3.6: Dimensión Asociado	52
Tabla 3.7: Dimensión Servicio	53
Tabla 3.8: Dimensión Call Center	53
Tabla 3.9: Dimensión Gerente	53
Tabla 3.10: Tabla de hechos SMS generados	53
Tabla 3.11: Tabla de hechos 0900	54
Tabla 3.12: Tabla de hechos Infoline	54
Tabla 3.13: Tabla de hechos Call Center	54
Tabla 3.14: Tabla de hechos Publicidad	55
Tabla 3.15: Estructura de archivos del Sistema de Inteligencia de Negocios	55
Tabla 3.16: Lista de los procesos de extracción, transformación y carga	60
Tabla 3.17: Formato de presentación de los procesos ETL	60
Tabla 3.18: Proceso Actualizar Almacén Diario	61
Tabla 3.19: Proceso Parte 1 Extracción Transformación Diaria	62
Tabla 3.20: Proceso Dimensión Asociado extracción	63
Tabla 3.21: Proceso Dimensión Cliente 0900 extracción	64
Tabla 3.22: Proceso Dimensión Compañía extracción	65
Tabla 3.23: Proceso Dimensión Numero	66
Tabla 3.24: Proceso Dimensión Palabra Clave	67
Tabla 3.25: Proceso Dimensión Rubro	68
Tabla 3.26: Proceso Dimensión Servicio	69
Tabla 3.27: Proceso Dimensión Call Center	70
Tabla 3.28: Dim Asociado Carga	71
Tabla 3.29: Dim Cliente 0900 Carga	72
Tabla 3.30: Dim compañía carga	72
Tabla 3.31: Dim Numero Carga	73
Tabla 3.32: Dim Palabra Clave Carga	73
Tabla 3.33: Dim rubro	74
Tabla 3.34: Dim Servicio Carga	74
Tabla 3.35: Dim Call Center Carga	75
Tabla 3.36: Parte 2 Carga Diaria	76
Tabla 3.37: Proceso tabla de hechos 0900	77
Tabla 3.38: Proceso tabla de hechos SMS Generados	78
Tabla 3.39: Proceso tabla de hechos Call Center	79
Tabla 3.40: Fact 0900 Carga	80
Tabla 3.41: Fact SMS Generados	80
Tabla 3.42: Fact Call Center Carga	81
Tabla 3.43: Fact Registro Clientes 0900	81
Tabla 3.44: Fact RC900 SMS Calcular Días a Procesar	82

Tabla 3.45: Fact RC900 SMS Generar Registro Día a Procesar	83
Tabla 3.46: Actualizar Almacén Mensual	84
Tabla 3.47: Parte1 ETL Mensual.....	85
Tabla 3.48: Proceso Dimensión Ejecutivo de Ventas	86
Tabla 3.49: Proceso tabla de hechos Infoline	87
Tabla 3.50: Dim Cliente Infoline	88
Tabla 3.51: Dim Ejecutivo Ventas	89
Tabla 3.52: Parte2-Carga-Mensual	90
Tabla 3.53: Fact Publicidad Extracción	91
Tabla 3.54: Fact Snapshot Infoline.....	92
Tabla 3.55: Dim Ejecutivo Ventas	93

Introducción

Las organizaciones han incrementado el uso de sistemas de información para obtener, además de los datos asociados a la operación del negocio, información relevante en forma de indicadores que miden hechos del negocio, mejorando así la eficiencia de sus procesos y la efectividad de sus resultados mediante la toma oportuna de decisiones, obteniendo mejoras visibles en su rentabilidad y mostrando una ventaja competitiva en el contexto donde se desenvuelve.

Generalmente, se ha recurrido a los sistemas transaccionales que llevan un registro detallado de los datos que se producen en los procesos de la organización, para la elaboración de informes de gestión, sobrecargando de esta manera su actividad. Estos informes son fundamentales ya que registran el resultado de los indicadores, el análisis experto de estos resultados y el conjunto de acciones recomendadas para corregir las desviaciones. Este informe es la base para la toma de decisiones. Mientras más datos tengan que ser analizados, más evidente es la necesidad de contar con herramientas automatizadas de procesamiento analítico de la información para la obtención de estos indicadores.

Las plataformas de Inteligencia de Negocios traen consigo una variedad de componentes que permiten extraer los datos, ya sea de fuentes internas o externas, transformarlos y cargarlos en almacenes o bodegas de datos, que son estructuras sencillas pero poderosas en cuanto a la forma de sintetizar la información. Asimismo, permiten ver los indicadores desde distintas perspectivas o criterios de clasificación.

El presente documento del Trabajo Especial de Grado tiene como objetivo principal implementar indicadores de gestión, a través de una plataforma de Inteligencia de Negocios, orientados a apoyar a la alta gerencia en general y específicamente a la gerencia de mercadeo en la toma de decisiones que faciliten la creación de estrategias de mercado y como consecuencia impulsen el negocio. Adicionalmente, explica las bases teóricas investigadas y estudiadas para lograr una visión más amplia y contribuir a la comprensión del ambiente que emplea este tipo de soluciones.

El documento está estructurado en 3 capítulos. En el capítulo I se plantea el problema de la investigación planteando la situación actual, el problema, solución propuesta, alcance, justificación y objetivos. El capítulo II define el marco conceptual y en el capítulo III se describe el marco aplicativo con la solución planteada y conclusiones.

Capítulo I - Problema de la Investigación

En este capítulo se detalla el contexto del problema describiendo la situación actual de una empresa seleccionada como caso de estudio, exponiendo la problemática a resolver y la solución adecuada a las necesidades de la misma; también se expone el alcance del Trabajo Especial de Grado (T.E.G.), el objetivo general, los objetivos específicos, la importancia y la justificación del mismo.

1.1 Título

Implementación de un sistema de inteligencia de negocios para el apoyo a la toma de decisiones en una empresa integradora de soluciones en tecnologías de comunicación.

1.2 Situación Actual

La empresa seleccionada como caso de estudio en el año 2013 es una organización integradora de soluciones de tecnologías de comunicación ubicada en el Estado Miranda que inició sus operaciones como proveedor de servicios de información principalmente mediante la vía telefónica. Con el objetivo de estar a la par en la tecnología de las telecomunicaciones, crea posteriormente una nueva unidad de negocios para el manejo de audio texto (líneas 0-900) y servicios de mensajería de texto SMS (Short Message Service). Años después y apoyándose en su amplia experiencia, se establece como una de las empresas líderes como proveedor de centro de llamadas.

El departamento de mercadeo de la organización posee la importante función de manejar y coordinar estrategias de venta. Así mismo está encargado de satisfacer los requerimientos y necesidades de los clientes basado en investigaciones realizadas para la formulación y lanzamiento de productos. Adicionalmente y como consecuencia de lo anterior, es el encargado de comunicar al departamento de finanzas los precios de ventas de los productos y servicios ofrecidos para que de esta forma el departamento de finanzas genere los presupuestos de ingresos y egresos necesarios para que la organización funcione de forma efectiva económica, financiera y contablemente.

La organización cuenta con diversos sistemas manejadores de base de datos y sistemas independientes (normalmente desarrollos propios) para el registro de las operaciones diarias. Sin embargo, debido a la gran cantidad de datos, la alta complejidad de la información que se maneja y el escaso tiempo para el análisis de la misma; a la gerencia de mercadeo se le complica el hecho de tomar decisiones con acierto y rapidez en un entorno sujeto a constantes cambios, más aún cuando la empresa se ve enfrentada a situaciones en las que debe tomar decisiones que determinarán su futuro próximo.

1.3 Planteamiento del Problema

Tal y como se expuso anteriormente, las organizaciones disponen de sistemas de información (también llamados sistemas transaccionales), cuya finalidad es preservar cada una de las transacciones generadas por la operación de los procesos diarios.

La organización cuenta con una gran cantidad de datos los cuales están almacenados en diferentes bases de datos, archivos de texto y archivos de Excel. Para la gerencia de mercadeo

obtener los reportes e indicadores conlleva mucho tiempo ya que las consultas toman largos periodos en generar resultados.

Para obtener reportes e indicadores, el gerente debe acudir a un conjunto de fuentes que le permita recolectar datos puntuales y agruparlos, esto puede venir luego de días de espera por consultas en las bases de datos o de hojas de cálculo que guardan registros de forma manual de la operación. Estas fuentes y la manera de construir información analítica adolecen de una serie de riesgos que van desde la captura y registro errado de datos, hasta el uso de exceso de tiempo útil para su desarrollo.

1.4 Solución Propuesta

Debido a que en los sistemas transaccionales se encuentra dispersa la información que necesitan los gerentes para desarrollar reportes e indicadores necesarios para la toma de decisiones de manera consistente y oportuna, se propone el desarrollo de un sistema de Inteligencia de Negocios, cuya arquitectura se ilustra en la siguiente Figura 1.1

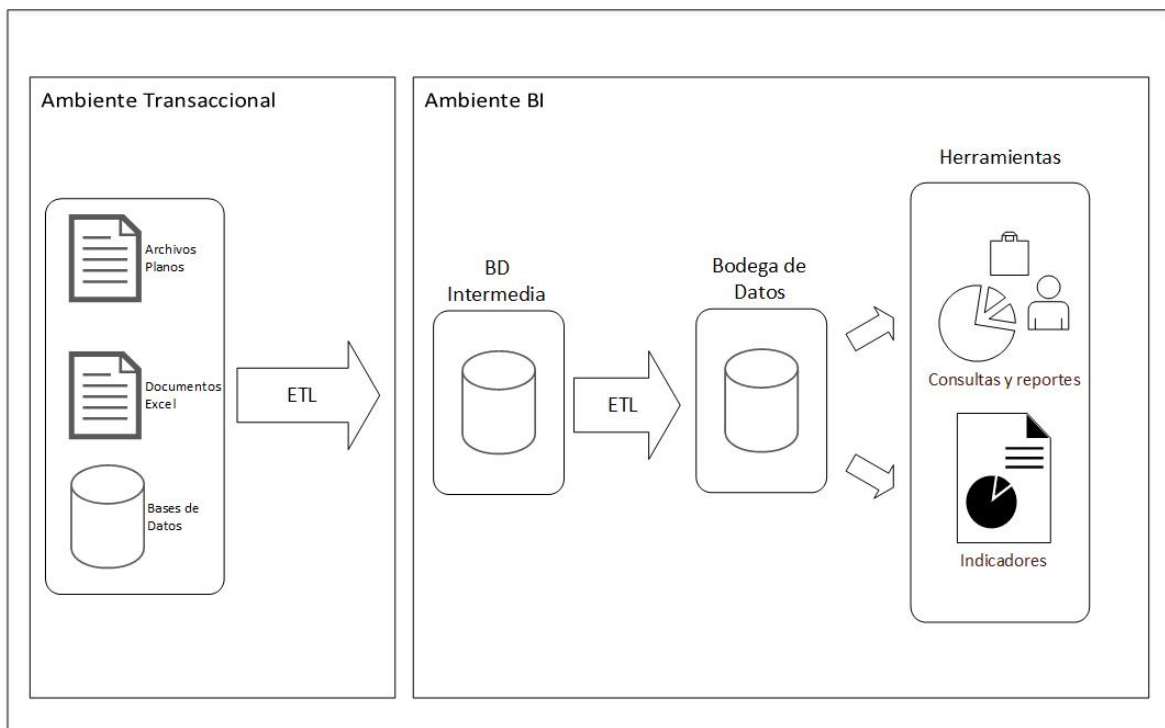


Figura 1.1: Arquitectura de la solución planteada

Los componentes ilustrados anteriormente podemos definirlos como:

Ambiente Transaccional: hace referencia a la información transaccional existente de la organización, allí se identifican las fuentes de datos (Hojas de Excel, Bases de Datos, Etc.) y se extrae sólo de él, la información necesaria para alimentar al sistema de Inteligencia de Negocios.

Ambiente BI: Es el nuevo ambiente a implementar y está compuesto por los siguientes componentes:

- **Base de Datos Intermedia:** a través de procesos de extracción, transformación y carga, se transfiere los datos pertinentes desde la base de datos transaccional a esta base de datos intermedia.
- **ETL, Extracción, transformación y carga:** a través de estos procesos propios del ambiente BI se modelan los datos desde la base de datos intermedia al Almacén de Datos.
- **Almacén de Datos o DWH (Data Warehouse):** es el repositorio principal, el cual contiene los datos necesarios de la organización de una manera íntegra, depurada, consistente y completa que permiten crear los cuadros de mando y reportes que son utilizados por las personas que toman decisiones para ser analizados.
- **Cuadros de Mando:** espacios de presentación que muestran los reportes, indicadores, gráficos y tablas dinámicas, que se utilizan para la toma de decisiones.

La metodología que gobernará los procesos de desarrollo e implementación del presente TEG será la descrita por Ralph Kimball en el libro *The data warehouse life cycle* y la plataforma tecnológica utilizada para el desarrollo de la solución será la aplicación Oracle Business Intelligence.

1.5 Justificación

Se plantea el desarrollo de una solución de Inteligencia de Negocios que permita a la organización disponer de una plataforma automatizada, orientada a sintetizar la información relevante para la toma de decisiones utilizando la tecnología de Oracle Business Intelligence, que contiene las herramientas necesarias para generar un modelo de datos orientado al análisis del negocio, integrar datos transaccionales, crear el Almacén de Datos y generar cuadros de mando con las consultas que muestran la información que será analizada.

En consecuencia, se dispondrá de una gestión optimizada para la gerencia de mercadeo de la organización, donde la toma de decisiones será rápida y confiable.

1.6 Alcance

Esta solución de Inteligencia de Negocios contendrá los siguientes reportes e indicadores, que según La gerencia de Mercadeo y la Alta Gerencia de la organización, constituyen un grupo de indicadores importantes y necesarios para la toma de decisiones.

Reportes:

- **Productividad por centros de llamadas:** mide por cada Centro de Llamada que atiende la organización, la cantidad de llamadas recibidas, atendidas y perdidas.
- **Total de ventas generadas por código corto:** indica el total de bolívares fuertes facturados por el lado de la organización por código corto, durante un periodo de tiempo determinado.
- **Top 1000 de cliente:** identifica a los mejores 1000 clientes, los cuales se medirán sumando el total en BsF consumidos mediante el consumo de mensajes SMS y el total de minutos de llamadas 0900, incluyendo datos del consumo por asociado.
- **Duración promedio de las llamadas cortas y largas a los 0900:** Este indicador muestra el promedio de minutos consumidos en cada línea 0900, dividido en llamadas cortas que

son aquellas que su duración es menor a 3 minutos y la llamada larga las cuales tienen una duración mayor a 3 minutos.

- **Comparativa de metas mensuales por unidad de negocios vs. facturación mensual por unidad de negocios:** compara el presupuesto estimado para las unidades de negocios (SMS, 0900 y pégate 503) con el facturado por unidad de negocios (SMS- 0900) en bolívares fuertes.
- **Cantidad de clientes activos y desafiliaciones mensuales en Infoline:** indica la cantidad de clientes activos en el servicio Infoline y la cantidad de desafiliaciones.
- **Clientes nuevos para la empresa:** mide los números ANI (Associate Number Identifier) que empiecen a usar los servicios ofrecidos por la empresa.
- **Números (ANI) en mensajes por el 530:** Monitorea los números que empiezan a mandar mensajes a través del 530 (pégate) y los números (ANI) que dejaron de enviar SMS por las operadoras y lo empezaron a hacer mediante el servicio 530.
- **Contenido consumido por número de teléfono:** Monitorea el número de mensajes y el número de minutos que consume un cliente por asociados, filtrados por etiqueta (palabra clave) y por rubros.
- **Inversión publicitaria invertida por asociado:** indica la inversión publicitaria invertida por asociado en donde un asociado está vinculado a un ejecutivo de ventas.
- **Desafiliaciones de clientes para la empresa:** mide los números que dejan de usar los servicios ofrecidos por la empresa.

1.7 Objetivo General

Implementar un sistema de inteligencia de negocios para el apoyo a la toma de decisiones en una empresa integradora de soluciones en tecnologías de telecomunicación.

1.8 Objetivos Específicos

- Generar los indicadores de gestión requeridos por la Gerencia de Mercadeo.
- Diseñar el Modelo Dimensional orientado al negocio, identificando los hechos que se van a medir, la granularidad, las dimensiones y las jerarquías.
- Implementar la base de datos intermedia, que contenga los datos transaccionales pertinentes para el desarrollo del Almacén de Datos, que respalde los requerimientos solicitados y formalizados.
- Incorporar un Almacén de Datos que contiene los datos que van a ser utilizados para crear cuadros de mando.
- Desarrollar los procesos de extracción, transformación y carga que lleva los datos desde la base de datos intermedia al Almacén de Datos.
- Verificar la igualdad de los datos de la base de datos intermedia con respecto a los del Almacén.
- Comprobar que el proceso de desarrollo de las consultas satisface los requerimientos de la Gerencia de Mercadeo.

- Generar las consultas y reportes basados en los requerimientos descritos, y los cuadros de mando asociados.

Capítulo II – Marco Conceptual

En este capítulo se detallan conceptos relevantes y necesarios para cumplir con los objetivos planteados en este trabajo. Definiciones como la de Sistema de Información y sus tipos, Bases de Datos, Almacén de Datos, Bodega de Datos, Modelo Dimensional, Inteligencia de negocios entre otros son desarrolladas a continuación.

También se desarrolla la definición de la metodología a utilizar para la implementación del sistema de inteligencia de negocios.

2.1 Sistemas de Información

Según Saroka (2002), un sistema de información es un conjunto de recursos humanos, materiales, financieros, tecnológicos, normativos y metodológicos, organizados para apoyar a los involucrados en la toma de decisiones de una organización, y así obtener información necesaria para desarrollar sus respectivas funciones.

Por otra parte, Laudon K. & Laudon J. (2004), definen a los sistemas de información como un conjunto de componentes interrelacionados que recogen (u obtienen), procesan, almacenan y distribuyen información para ayudar en la toma de decisiones y el control de una organización, tal como se muestra en la Figura 2.1

Tomando en cuenta las definiciones anteriores, se puede armar que un sistema de información es un conjunto de procesos que se encargan de recolectar, procesar, almacenar y distribuir información específica para actividades de dirección y control, apoyando en parte a la toma de decisiones necesarias para el desarrollo de las actividades y funciones de negocios de una organización.

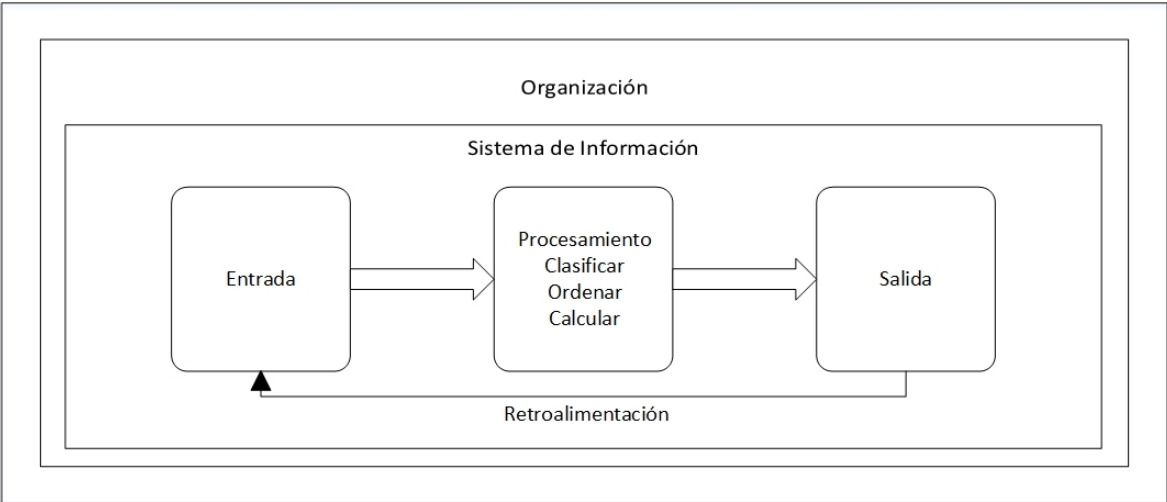


Figura 2.1: Sistema de información. (Laudon K. & Laudon J., 2004)

2.1.1 Tipos de Sistemas de información

Según Laudon K. & Laudon J. (2004), los sistemas de información se dividen en cuanto a los niveles de la organización. En la Figura 2.2, se pueden observar los tipos de sistemas de información, según los niveles de una organización.

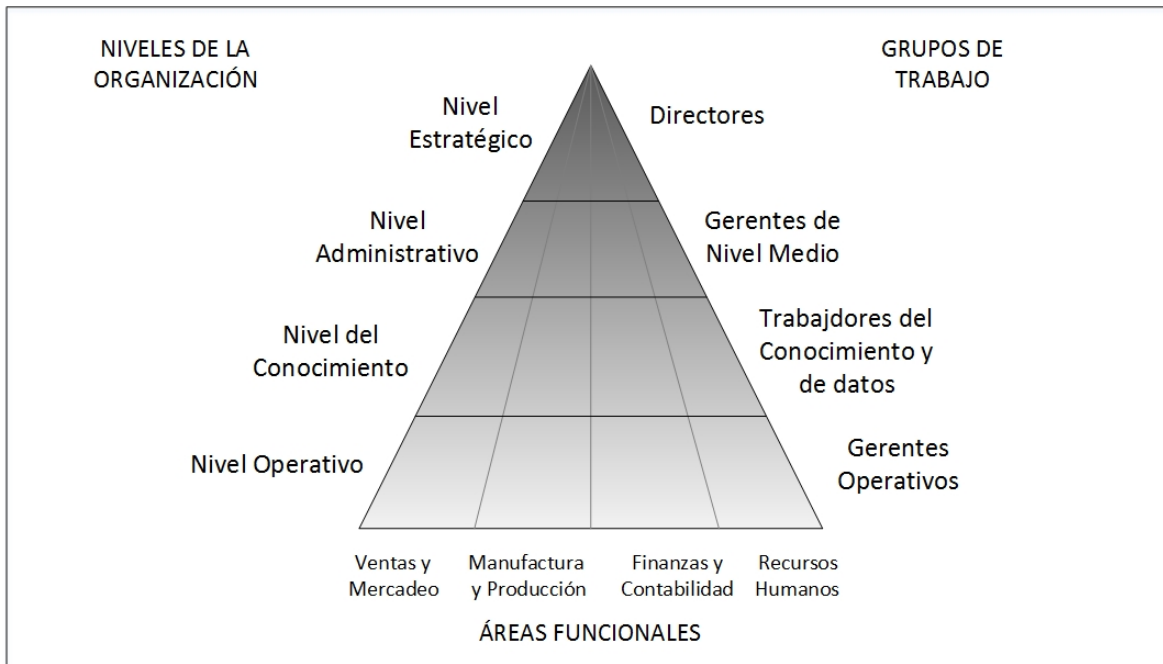


Figura 2.2: Tipos de Sistema de información. (Laudon K. & Laudon J., 2004)

A continuación, se describen cada uno de los niveles:

- **Sistemas a Nivel Operativo:** Son aquellos que apoyan a los gerentes operativos en el seguimiento de las actividades y transacciones cotidianas y elementales de la organización. Este nivel cuenta con sistemas de procesamiento de transacciones (TPS, por sus siglas en inglés), que efectúan y registran transacciones diarias de los negocios.
- **Sistemas a Nivel de Conocimiento:** Son aquellos que apoyan la integración de nuevos conocimientos del negocio, ayudando a controlar el flujo de trabajo dentro del negocio. Este nivel cuenta con los Sistemas de Trabajo de Conocimiento (KWS, por sus siglas en inglés), que apoyan a trabajadores profesionales, generándoles nuevos conocimientos que contribuyen a la organización; y los Sistemas de Ocená que ayudan, principalmente, a los trabajadores de datos, quienes usan la información para analizarla y transformar datos.
- **Sistemas a Nivel Administrativo:** Son aquellos que apoyan a las actividades de supervisión, control y toma de decisiones. En este nivel se encuentran los Sistemas de Información Gerencial (MIS, por sus siglas en inglés), quienes se encargan de dar soporte a las tareas organizacionales, proporcionando informes en períodos de tiempo, que ayuden a los gerentes a la toma de decisiones, control y planeación a nivel administrativo.

Adicionalmente, se encuentran los Sistemas de Apoyo a la Toma de Decisiones (DSS, por sus siglas en inglés), quienes tienen la tarea de ayudar a los gerentes a tomar decisiones exclusivas, cambiantes y no fáciles de anticipar.

- **Sistemas a Nivel Estratégico:** Son los que apoyan a las actividades de planificación de las organizaciones a largo plazo, y, además, ayudan a los directores a enfrentar y resolver aspectos estratégicos; dando así, los cambios necesarios del entorno externo de la organización. En este nivel se encuentran los Sistemas de Apoyo a Ejecutivos (EES, por sus siglas en inglés), quienes dan soporte a la toma de decisiones no rutinarias que requieren juicio, evaluación y comprensión, es decir, apoyan a la toma de decisiones mediante gráficos y comunicaciones avanzadas.

2.2 Base de Datos

Para Oppel (2009), una base de datos identificada normalmente como DB (por sus siglas en inglés) no es más que una colección de datos interrelacionados que son manejados por una sola unidad. Es decir, se considera un sitio físico y lógico donde se almacenan datos, pero estos datos están relacionados entre sí, para darle un sentido al almacenamiento que estamos haciendo.

Una base de datos implica, además de estructuras y modelos, un software que soporte los mismos. Esto quiere decir que existen distintas maneras de modelar y de crear una base de datos y que en estas bases de datos intervienen distintos elementos como los datos en sí, objetos de bases de datos, archivos, entre otros. Pero las bases de datos dependen de un software el cual es el que soportara lógicamente esta estructura, a este software se le conoce como Sistema Manejador de Bases de Datos (DBMS, por sus siglas en inglés) (Oppel, 2009).

Un sistema manejador de bases de datos provee herramientas y servicios que permiten ver, manipular y manejar las bases de datos. Dichos servicios permiten la manipulación de los datos que se encuentran en la base de datos física, el manejo multiusuario de las bases de datos, el manejo de las transacciones que se realizan en la base de datos, dar soporte de comandos que permiten editar y ver datos, el respaldo de las bases de datos y mecanismos de seguridad para prever que solo usuarios autorizados vean y manipulen datos. En el mercado existen distintos sistemas manejadores de bases de datos tales como Oracle, MySQL, Access, Microsoft SQL Server, PostgreSQL, entre los más destacados.

Es importante resaltar que las bases de datos cuentan con una capa física, que son los componentes en los cuales se almacenan físicamente los datos y cuentan con una capa lógica, que es el modelo en el cual se basa para almacenar y relacionar lógicamente los datos.

2.2.1 Modelos de Base de Datos

Modelo jerárquico: Es un modelo en el que los registros se convierten en nodos, y a través de las relaciones entre nodos, se forma una jerarquía, lo cual quiere decir que lógicamente un nodo estará encima de otros, siendo este el padre de los nodos sobre los cuales este, y siendo hijo del nodo que este encima de él. La conexión o relación entre los nodos se hace vía apuntadores, y son relaciones uno a muchos, lo cual quiere decir que un nodo puede tener muchos hijos, pero solo puede tener un padre. En la Figura 2.3, se puede apreciar un ejemplo de dicho modelo.

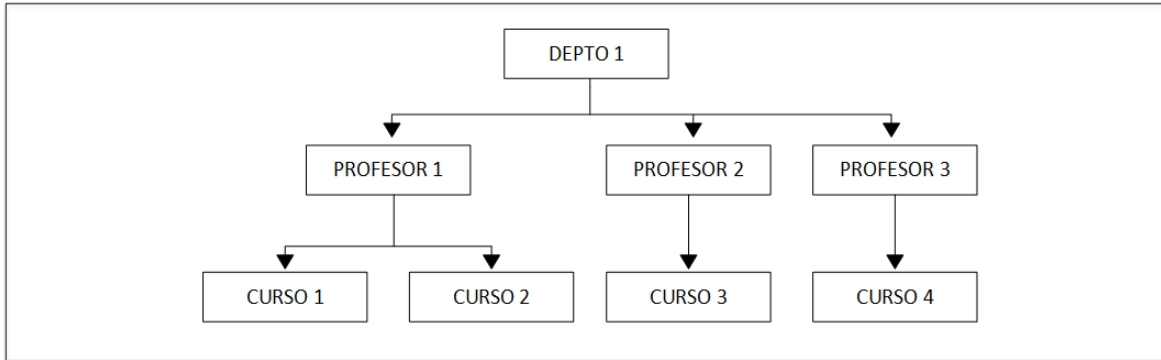


Figura 2.3: Modelo Jerárquico. (Moraga, 2001)

Modelo red: Este modelo se diferencia del anterior en que cambia de ser una relación uno a muchos a ser una relación muchos a muchos, lo cual significa un cambio muy importante ya que esto permitiría navegar con más rapidez entre los nodos y relacionarlos de una mejor manera entre sí. Sin embargo, este modelo sigue constituyendo una jerarquía, lo cual quiere decir que aún existe esa figura de padre-hijo entre los nodos. En la Figura 2.4, se muestra un ejemplo de este modelo.

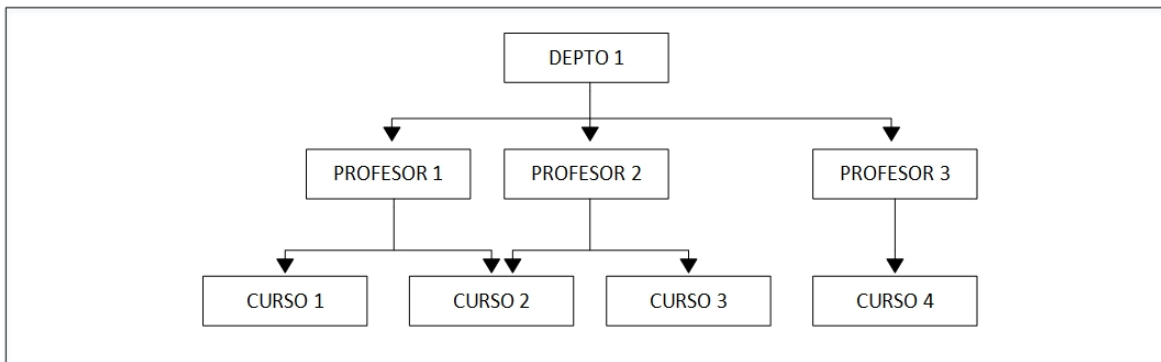


Figura 2.4: Modelo Red. (Moraga, 2001)

Modelo relacional: Es probablemente el modelo más utilizado en la actualidad, el modelo relacional se impuso a los dos anteriores por ser un sistema mucho más flexible y generalizado. Permite relaciones de uno a uno, uno a muchos y muchos a muchos, y el elemento a relacionar se denomina entidad, por esto es común escuchar que se nombra a los modelos que representan base de datos relacionales como modelos entidad-relación. Adicionalmente, se apoya en claves e índices para relacionar y poder viajar de una tabla a otra y conocer sus relaciones, en este caso se utiliza una clave primaria para reconocer al registro una clave foránea que deberá ser una clave primaria de otro registro. Si la relación es de muchos a muchos, se creará una estructura llamada relación, la cual servirá como una especie de puente o conexión entre ambas entidades. En caso de ser relaciones de uno a muchos la clave primaria del uno estará contenida en la entidad muchos, y en el caso uno a uno entonces la clave puede estar en cualquiera de las dos entidades. La Figura 2.5 muestra un ejemplo de este tipo de modelo.

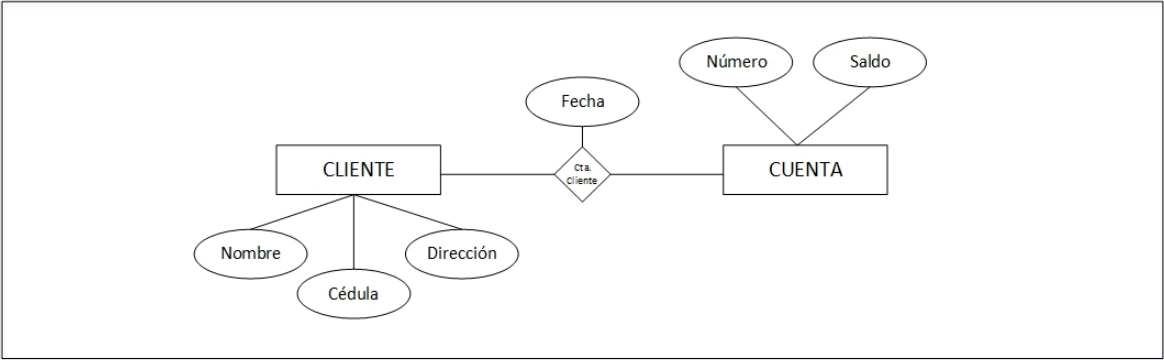


Figura 2.5: Modelo Relacional. (Moraga, 2001)

Modelo objeto-relacional: Este modelo explota las virtudes del modelo relacional y el orientado a objeto. Dicho modelo surge principalmente por la necesidad de manejos de datos complejos como se mencionan anteriormente, pero a su vez brinda las facilidades de consulta que brindan los modelos de base de datos relacionales.

Modelo orientado a objetos: Este modelo se basa (como su propio nombre lo dice) en objetos, estos objetos agrupan datos y funciones asociadas con lo que debe ser un objeto real en el computador. En la Figura 2.6 se puede apreciar un ejemplo de dicho modelo.

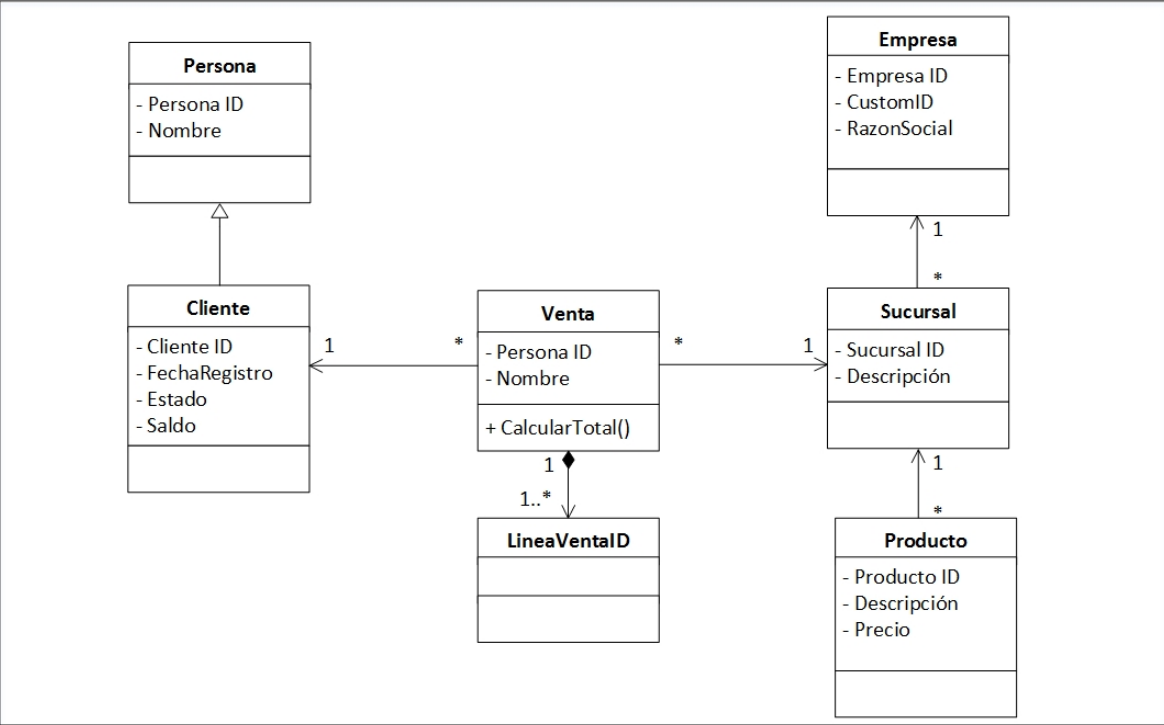


Figura 2.6: Modelo Orientado a Objetos. (Moraga, 2001)

2.2.2 Tipos de Base de Datos según su propósito

Además de tener un propósito general, el cual es el almacenar los datos, las bases de datos pueden tener distintas funciones y pueden servir para distintos propósitos, por ello se cuentan con distintas denominaciones (Oppel, 2009). Se contemplan solo los tipos de bases de datos que podrían ser útiles para el propósito de la investigación:

- **Almacén de datos:** Es considerado un sistema que recupera y consolida datos periódicos desde sistemas fuentes a sistemas dimensionales o almacena datos normalizados (Rainardi, 2008). Los almacenes de datos son utilizados regularmente para sistemas de Inteligencia de Negocios, por su capacidad de almacenar los datos en forma dimensional y de además cargar datos normalizados.
- **Base de datos transaccionales:** es una base de datos regular pero cuyo fin es realizar conexiones rápidas, es decir, es una base de datos usada frecuentemente para transacciones. Estas bases de datos cuentan regularmente con grandes volúmenes de datos, pero también con gran capacidad de conexión para que las transacciones se realicen con la rapidez que ameriten.
- **Base de datos distribuida:** son un conjunto de bases de datos en distintos espacios físicos y que se comunican entre sí a través de una red de computadores. Las bases de datos distribuidas trabajan en conjunto dos áreas importantes de la computación, las bases de datos y las redes, ya que almacenan datos por medio de distintas bases de datos y utilizan los protocolos de comunicación de redes para la comunicación entre ellas.

2.2.3 Sistemas Manejadores de Base de Datos

Un sistema manejador de base de datos, es un software que permite la interacción entre un usuario y distintas bases de datos. Los sistemas manejadores de bases de datos proveen distintas interfaces y funcionalidades que permiten al usuario interactuar de forma natural con las bases de datos. A continuación, en la Tabla 2.1, se muestra un cuadro comparativo entre cuatro sistemas manejadores de base de datos, de los más conocidos.

	PostgreSQL	Oracle	MySQL	FastObjects
Multiplataforma	Si	Si	Si	Si
Multiusuario	Si	Si	Si	Si
Modelo(s)	Relacional, Objeto- Relacional	Relacional, Objeto- Relacional	Objeto- Relacional	O.O.
Licencia	Libre	Propietaria	Propietaria	Libre

Tabla 2.1: Cuadro comparativo de Sistemas Manejadores de Bases de Datos.

2.3 Almacén de Datos

Según Kimball (1997), un almacén de datos, no es más que una copia de datos transaccionales, específicamente estructurados, utilizados para consulta y análisis.

Por su parte Inmon (1998), define un almacén de datos como un conjunto de datos Integrados orientados a materia, variantes con el tiempo y no transitorios, los cuales soportan el proceso de toma de decisiones de una administración.

Partiendo de estas definiciones, de dos de los pioneros en el área de almacenes de datos, se puede armar que ambas se centran en los datos per se. Por lo tanto, podemos definir un almacén de datos, de forma más generalizada, como un repositorio centralizado de datos, que almacena gran cantidad de información, utilizada para apoyar en la toma de decisiones de una organización. Generalmente, estos almacenes son caracterizados por ser la fuente que alimenta los procesos analíticos dentro de una organización.

Adicionalmente, los almacenes de datos se encargan de guardar información histórica y permiten principalmente consultas, debido a que por lo general están conectados a herramientas que extraen la información para mostrarla a gerentes o analistas para la toma de decisiones.

2.3.1 Arquitectura de un Almacén de Datos

A continuación, en la Figura 2.6, se observa la arquitectura de un almacén de datos:

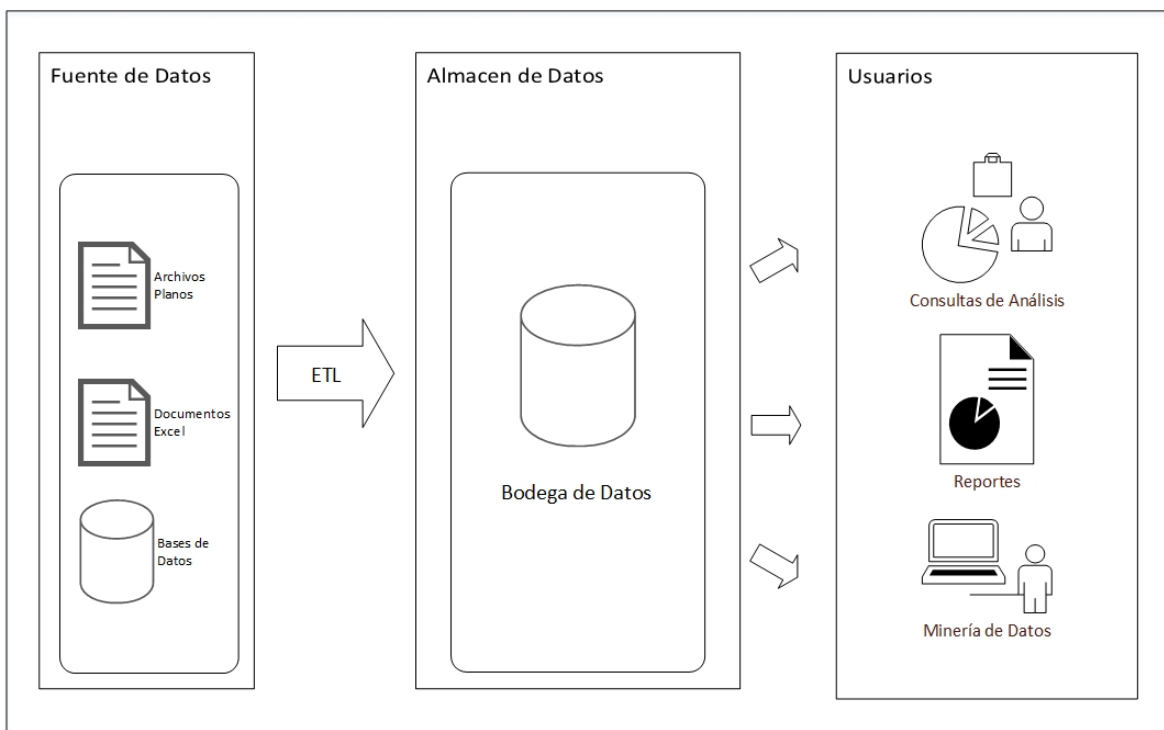


Figura 2.7: Arquitectura de un almacén de Datos. (EcuRed.CU, 2012)

2.3.2 Características

- **Orientado a Temas:** Ya que la información es clasificada en base a los aspectos más importantes de la organización. Principalmente, se busca centrar la información en las actividades básicas, y no en procesos, de dicha organización, ya sea ventas, producción, distribución, etc. Así, se facilita el acceso y entendimiento de los datos por parte de los

usuarios. La diferencia entre en actividades y procesos se ubica en el nivel de detalle de los datos.

- **Variante en el tiempo:** Esto se debe a que los datos del almacén son relativos a un periodo de tiempo. Las variaciones que se produzcan en los datos, a lo largo del tiempo, deben quedar registradas periódicamente. De esta forma, se permite generar reportes que reflejen el comportamiento de dichas variaciones.
- **No Volátil:** Ya que la información que se encuentra en el almacén, no puede ser modificada, ni eliminada, solamente se permite agregar. En otras palabras, los datos cargados en el almacén solo podrán servir de lectura. Por lo tanto, se puede armar que los almacenes de datos permiten dos operaciones: carga inicial de los datos y acceso a los mismos. Así, se elimina la necesidad de tener mecanismos de control de concurrencia y recuperación de datos.
- **Integrado:** Debido a que los datos del almacén deben ser integrados en estructuras consistentes, eliminando así las inconsistencias existentes en las fuentes de datos. Esta característica permite reunir todos los datos, provenientes de las diferentes fuentes, y agruparlos en un modelo generalizado, único y coherente. Adicionalmente, estos datos deben ser transformados antes de incorporarlos al almacén, asegurando así la consistencia de los mismos.

2.4 Bodega de Datos

Inmon (2000), define bodega de datos como una estructura departamental de suministro de datos desde el Almacén de Datos donde los datos están desmoralizados, basándose en la información que necesite un departamento. En una Bodega de Datos se almacena un subconjunto de datos, con el propósito de ayudar a un área específica dentro de un negocio para tomar mejores decisiones.

La Bodega de Datos, suele ser cargada con datos de solo unas pocas fuentes. Las fuentes pueden ser sistemas internos, Almacenes de Datos, o de datos externos.

2.4.1 Características

La Bodega de Datos cumple con ciertas características que son esenciales para poder llevar a cabo su construcción e implementarlo en una solución de Inteligencia de Negocios, estas características fundamentales son (Inmon, 2000) (Kimball & Caserta, 2008):

- Trata con usuarios limitados.
- Trabaja con un área específica de una empresa
- Tiene un propósito u objetivo específico.
- Tiene una función de apoyo con respecto al Almacén de Datos.
- Tiene una mayor rapidez de consulta.
- Facilidad para la historización de los datos.
- Validación directa de los datos.

2.5 Diferencias entre Bodega y Almacén de Datos

Un Almacén de Datos, a diferencia de una Bodega de Datos, se ocupa de varios temas y es controlado e implementado por una unidad central organizacional. Se entiende como central al Almacén de Datos de la empresa. Por lo general un Almacén de Datos reúne información de múltiples sistemas (Julie & Steinbart, 2000).

Ninguna de las definiciones básicas de una Bodega de Datos limita el tamaño como la complejidad del apoyo y la toma de decisiones que puede contener. Sin embargo, las Bodegas de Datos son más pequeñas y mucho menos complejas que los Almacenes de Datos, por lo que en general son más fáciles de construir y mantener. En la Tabla 2.2 se presenta una comparación detallada entre ellos (Inmon, 2000), (Kimball & Margy, 2002).

	Almacén de Datos	Bodega de Datos
Alcance	Corporativo	Línea de Negocios
Tema	Temas	Único Tema
Orígenes de Datos	Muchos	Pocos
Tamaño Relativo	100GB-TR+	menos de 100GB
Tiempo de Implementación	Meses, Años	Meses

Tabla 2.2: Cuadro comparativo entre Almacén de Datos y Bodega de Datos

2.6 Extracción, Transformación y Carga (ETL)

1. Según el Portal de ETL y Data Warehousing (ETLTools.org, 2009), la extracción, transformación y carga es un proceso usado en las áreas de base de datos, especialmente en los almacenes de datos, con el fin de migrar de un punto a otro. Este proceso consta de tres fases, que se combinan en una sola herramienta, facilitando el transporte de los datos, desde la fuente hasta el almacén. Las tres fases de este proceso se describen a continuación:
2. Extracción: Es la primera fase del proceso ETL, en esta fase se obtienen los datos provenientes de las diferentes fuentes (sean externas o internas) que serán utilizados por el almacén de datos. Esta fase busca estandarizar los datos, de tal forma que los mismos sean entendibles en la fase de transformación. Además, se incluye un filtrado de datos para eliminar la redundancia o los datos de poco interés.
3. Transformación: En esta fase se aplican una serie de reglas y funciones a los datos extraídos y que serán cargados posteriormente en el almacén. En algunos casos, se aplican a la una serie de transformaciones, como, por ejemplo: ordenación de los datos, unión de los datos desde diferentes fuentes, validación, etc.
4. Carga: Representa la última fase del proceso ETL, y esta se encarga de guardar los datos transformados en el almacén. Debido a que en esta fase existe una interacción directa con la base de datos, al momento de la inserción de datos, generalmente se activan los disparadores, o los llamados triggers, y las restricciones existentes.

2.7 Modelo Dimensional

El modelo dimensional es una técnica de diseño para Almacenes de datos (a nivel lógico) que pretende representar los hechos del negocio. Este modelo esta optimizado para llevar a cabo consultas con un alto rendimiento (Verástegui, 2007).

Todo diseño de un Almacén de Datos debe comenzar con un modelo dimensional, debido a que se identifica cuáles son los hechos que se desean medir y desde cuales perspectivas se desean agrupar los mismos. Por tal motivo, es necesario definir y tener claros los conceptos básicos concernientes al modelo dimensional, como lo son: dimensión, hecho, granularidad, jerárquicos, los tipos de tablas involucradas como tabla de dimensiones y tablas de hechos, que están conformados por dos (2) tipos de esquemas los cuales son copo de nieve y esquema estrella.

2.7.1 Dimensión

Para Kimball & Caserta (2008), una dimensión es una entidad independiente en el modelo dimensional que sirve como un punto de entrada o como un mecanismo de reordenamiento y fraccionamiento de las medidas sumariadas, localizadas en la tabla de hechos del modelo. Son criterios puntuales y conocidos por el área de negocio, para la presentación de los datos al usuario final, como, por ejemplo: producto, zonas, tiempo, entre otros. Las dimensiones son la información que define a cada uno de los registros de la tabla de hechos.

2.7.2 Tabla Dimensión

Según Kimball & Caserta (2008), la tabla de dimensión es una tabla en el modelo dimensional con una clave primaria simple y columnas de atributos descriptivos.

Cuando se crea un modelo dimensional, se requiere una tabla de hechos y se crean varias tablas de dimensiones (su forma se muestra en la Figura 2.8), estas tablas de dimensiones acompañan a la tabla de hecho para definir los parámetros de los cuales dependen los hechos registrados en ella, para eso se tiene que entender que una tabla de dimensión son ciertos elementos que contiene atributos (campos), que se utilizan para agrupar o restringir los datos almacenados en una tabla de hechos que son consultadas en un entorno de Bodega de Datos o Almacén de Datos.



Dimension_Producto	
idProducto	INTEGER
Nombre	VARCHAR(60)
Categoria	VARCHAR(60)
Subcategoria	VARCHAR(60)
Marca	VARCHAR(50)
Peso	FLOAT
Altura	INTEGER
Anchura	INTEGER
Profundidad	INTEGER

Figura 2.8: Ejemplo: Tabla Dimensión. (Darmawikarta, 2007)

2.7.3 Hecho (Medida)

Kimball & Caserta (2008), definen un hecho como una medida del desempeño empresarial, usualmente son valores numéricos y aditivos que son almacenados en una tabla de hecho.

Generalmente los hechos que se almacenan son números enteros (ejemplo: cantidad vendida de un producto, cantidad solicitada de un producto) o números reales (ejemplo: costo de producción de un producto, monto obtenido por venta).

2.7.4 Tabla de Hechos

En el modelo dimensional, la tabla de hechos es la tabla principal con las mediciones de rendimiento numéricas que son caracterizadas por una clave compuesta, las cuales cada elemento es una clave foránea que corresponde a una tabla de dimensiones (Kimball & Caserta, 2008).

Se entiende como tabla de hechos, la tabla principal en un modelo numérico o dimensional, donde las mediciones de una empresa se almacenan. Cada medida se toma de la intersección de las dimensiones que la definen, estas dimensiones estarán alrededor de la tabla de hechos directamente relacionados.

Una fila de la tabla de hechos corresponde a una medida. La medida es una fila de la tabla de hechos. Todas las medidas deben estar basadas en un mismo nivel de granularidad (grano) (Kimball & Caserta, 2008).

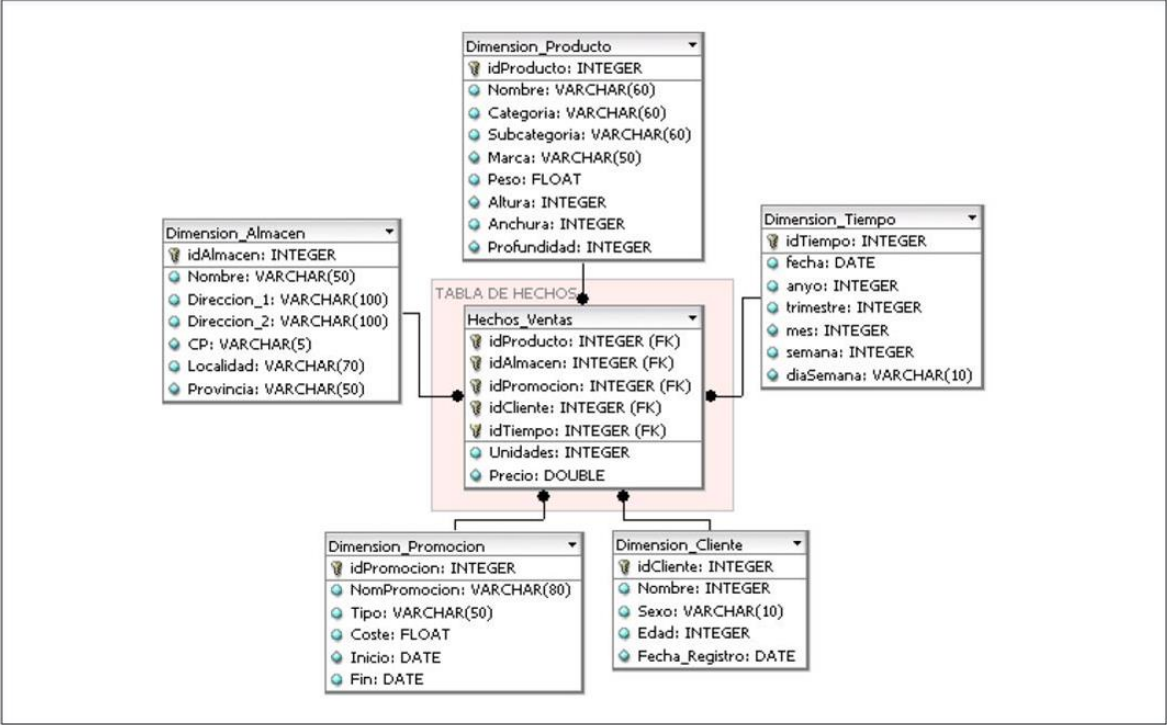


Figura 2.9: Ejemplo: Tabla de Hechos. (Darmawikarta, 2007)

Los hechos más útiles en una tabla de hechos son tanto los numéricos como los aditivos.

La tabla de hechos representa esa relación de muchos a muchos, entre las distintas dimensiones que se relacionan. En la Figura 2.9 se observa un ejemplo donde, una tabla de hechos Ventas está representado por un esquema de tipo estrella y las tablas de dimensión que están alrededor de la tabla de hechos que son las tablas de producto, tiempo, almacén, promoción y cliente. En la tabla de hechos se almacenan dos variables, la de unidades en venta y precio de venta del producto o los productos.

- **Esquema de Estrella:** Es la representación genérica de un modelo dimensional en una base de datos relacional, en la cual una tabla de hechos con una clave compuesta es unida a un número de tablas de dimensiones, cada una con una clave primaria simple (Kimball & Caserta, 2008). Consiste en estructurar la información en procesos, vistas y métricas a modo de estrella. En la tabla de hechos se encuentran los atributos destinados al hecho que constituye el proceso de negocio a medir, es decir, sus métricas. Mientras, en las tablas de dimensión, los atributos se destinan a elementos de nivel (que representan los distintos niveles de las jerarquías de dimensión) y a atributos de dimensión (encargados de la descripción de estos elementos de nivel).
- **Esquema de Copo de Nieve:** Es una dimensión normalizada en el cual una tabla de dimensión simple es descompuesta en una estructura de árbol con muchos niveles (Kimball & Margy, 2002). Un esquema copo de nieve es más complejo que un esquema estrella, ya que algunas dimensiones están compuestas por más de una tabla de datos o tablas de dimensión. Para normalizar las tablas y reducir el espacio de almacenamiento, eliminando la redundancia de datos. La única desventaja que presentan es la creación de otras tablas de dimensión y relaciones entre las tablas. Se tiene un impacto sobre el rendimiento del Almacén de Datos.

2.7.5 Granularidad (Grano)

La granularidad es el nivel de detalle que posee cada registro de una tabla de hechos (Kimball & Caserta, 2008). La granularidad representa el nivel de detalle al que se desea almacenar la información sobre el Almacén de Datos que se esté analizando. Por ejemplo, los datos referentes a ventas o compras realizadas por una empresa, pueden registrarse día a día, en cambio, los datos pertinentes a pagos de sueldos o cuotas de socios, podrán almacenarse a nivel de mes.

2.7.6 Jerarquía

La jerarquía es una serie de relaciones en cascada de uno a muchos (Kimball & Caserta, 2008). Una dimensión debe contener al menos una jerarquía, la cual puede tener varios niveles. Por ejemplo, se puede tener en una jerarquía que organiza a los clientes por ubicación geográfica, en la que se puede incluir los niveles Región, País, Ciudad y el nivel cliente.

Después de conocer cuál será la granularidad, se procede a realizar la jerarquía. Esta jerarquía corresponde con las tablas de dimensión que se definen como los niveles de asociación que se tienen de los datos.

2.7.7 Agregación

La agregación es una medida, es decir un dato contable (Williams, 2007). Por ejemplo, si se quiere tener el total del salario de una empresa, se realiza un proceso de agregación, definiendo como operación la suma de los sueldos de los empleados, logrando que cuando se reciba más información en el atributo sueldo, se muestre suma rizada. Otro ejemplo, para una dimensión Producto se puede incluir un valor llamado “Total Productos”, que será padre de todos los demás niveles de jerarquía y que contendrá el acumulado de todos ellos, logrando que, al introducir más productos, se sume y se actualice el resultado.

2.8 Inteligencia de Negocios

La Inteligencia de Negocios se basa en procesos, tecnologías y herramientas necesarias para transformar datos en información, información en conocimiento y conocimiento en planes de negocios rentable (Loshin, 2003). Es el conjunto de productos y servicios que tiene como objetivo analizar a los usuarios finales, acceder, manipular y estudiar, de una manera rápida y sencilla, la información para la toma de decisiones del negocio.

También se puede definir como un conjunto de metodologías, herramientas, aplicaciones o tecnologías que sirven para reunir, depurar y transformar datos de los sistemas transaccionales e información no estructurada en información estructurada, logrando así realizar una exploración y análisis de datos, proporcionando soporte a la toma de decisiones. La Inteligencia de Negocios abarca un conjunto de estrategias y herramientas enfocadas a la administración y creación de conocimiento mediante el análisis de datos existentes (Hamel & Prahalad, 1994).

2.8.1 Características

La Inteligencia de Negocios posee características como (Hamel & Prahalad, 1994):

- Fácil acceso a la información: permite el acceso de los usuarios finales a los datos, con independencia de la procedencia de estos.
- Apoyo a la toma de decisiones: permite que los usuarios finales tengan acceso a las herramientas de análisis, las cuales permiten seleccionar y manipular los datos de la empresa.
- Orientación al usuario final: busca independencia entre los conocimientos técnicos de los usuarios y la capacidad para utilizar las herramientas.

2.8.2 Beneficios

La Inteligencia de Negocios, permite que las compañías puedan generar el mayor valor de las líneas existentes del negocio en estudio, creando nuevas oportunidades y reduciendo los ciclos de desarrollo de un producto, agilizando operaciones.

La Inteligencia de Negocios ayuda a resolver ciertas inquietudes que se puedan presentar, como (Martínez, 2002):

- ¿Cuáles son los mejores proveedores o cuales son los clientes que proporcionan mayor ganancia, y cómo hacer para atraerlos aún más?
- ¿Cuáles son los candidatos a clientes?
- ¿Cuál es la proporción de los gastos con respecto a las ventas?

Lo que se busca es encontrar información, que no solamente contestes preguntas de lo que pasó o lo que está sucediendo en el negocio, sino que también construya modelos que sirvan para predecir futuros eventos.

También se pueden obtener otros beneficios como (Guarente, 2002):

- Controlar los costos, ya que ofrece una solución que permite manejar fácilmente las inversiones monetarias que se generan en los distintos departamentos de la organización.
- Mejorar la colaboración y la calidad de las decisiones, facilitando el acceso a la información en todos los niveles de la organización de manera rápida y contable.
- Orientar las soluciones tecnológicas hacia el usuario, ya que reduce los tiempos de aprendizaje mediante el uso de herramientas.
- Proporcionar una profunda visión del negocio a través de un sistema integrado de usos: cuadros de mando, tableros de instrumentos, informes, migración de datos y almacenamiento analítico.
- Proveer asistencia a los ejecutivos para planear y pronosticar el trabajo, presentando una descripción común de los procesos del negocio de una compañía

2.8.3 Tipos de Soluciones

La Inteligencia de Negocios muestra resultados de acuerdo a las necesidades de los distintos niveles jerárquicos de la organización.

La Figura 2.10 presenta una pirámide con la ubicación de los sistemas de Inteligencia de Negocios con respecto a los sistemas de información y los niveles jerárquicos dentro de la organización, los cuales serán detallados a continuación (Inteligencia de Negocios, 2001):

1. Nivel estratégico, en el cual la Inteligencia de Negocios permite que la parte directiva de la empresa pueda analizar y seguir día a día las tendencias, patrones, metas y objetivos estratégicos de la empresa. Un ejemplo de ello, lo constituye el cuadro de mando integral, cuadros de mando, entre otros.
2. Nivel táctico, en el cual la Inteligencia de Negocios permite a los analistas de datos y a la gerencia media de una empresa, utilizar herramientas de análisis y consultas con el fin de obtener acceso a la información sin intervención de terceros.

- 3. Nivel operativo, en el cual la Inteligencia de Negocios permite a los empleados de la empresa recibir de forma oportuna, exacta y adecuada la información operativa, basándose en herramientas de trabajo como reportes, hojas de cálculo, manteniendo siempre un formato cuya información se actualiza cada cierto tiempo.

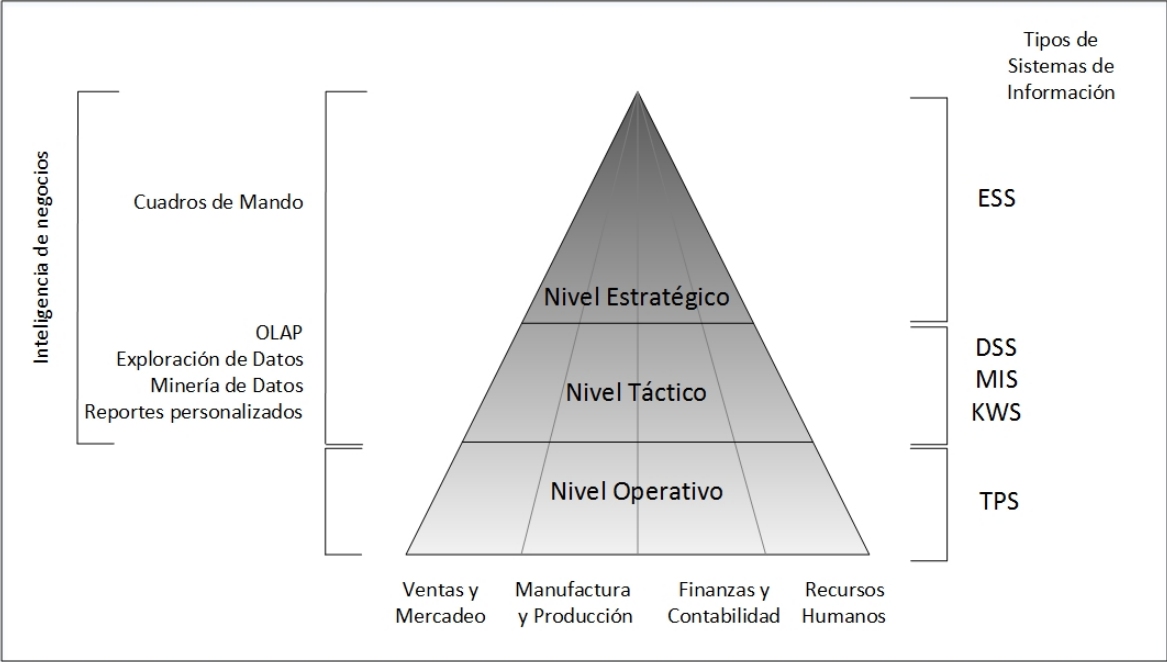


Figura 2.10: Sistemas de IN dentro de los tipos Sistemas de Información (Inteligencia de Negocios, 2001).

2.8.4 Arquitectura de una Solución

Una solución de Inteligencia de Negocios se caracteriza por presentar 3 bloques importantes: las fuentes de datos, los componentes y las herramientas para presentar los datos resumidos o en detalle (NASE, 2010). En la Figura 2.11, se diferencian los tres bloques que contiene una arquitectura de una solución de Inteligencia de Negocios:

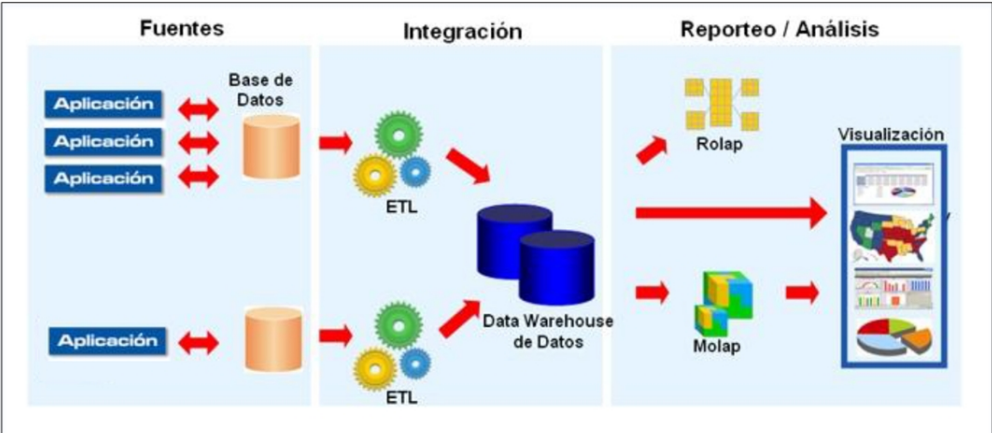


Figura 2.11: Arquitectura de la solución de Inteligencia de Negocios (NASE, 2010)

En el primer bloque se tienen las fuentes de datos, que están representadas por aquellos sistemas que contienen datos que pueden ser explotados potencialmente por los componentes de una solución de Inteligencia de Negocios. Generalmente, estos datos suelen presentarse en sistemas transaccionales, archivos planos, hojas de cálculos o bodegas de datos de diferentes departamentos de una organización.

El segundo bloque contiene los componentes que integran los datos obtenidos de las fuentes, ya sean internas o externas. Estos datos pasan por un proceso en el cual se filtran, se estandarizan y se cargan en almacenes o bodegas. Estos almacenes son los componentes de una solución BI, ya que se encargan de garantizar la integridad de los datos contenidos en ellos.

Por último, se tiene el tercer bloque, donde se encuentran las herramientas que se utilizan para representar los datos contenidos en el almacén o bodega de datos. Estos datos pueden representarse en forma de consultas, reportes, análisis, cuadros de mando, tablas dinámicas, etc.

Estas herramientas son indispensables para mostrar los datos relevantes, resumidos o en detalle, de una organización o empresa. Asimismo, se garantiza facilidad al momento de realizar los análisis y reportes, con el fin de agilizar el proceso de toma de decisiones.

Por lo tanto, se puede decir que toda solución de Inteligencia de Negocios permite agrupar datos de diferentes fuentes en un repositorio centralizado, con el fin de representar los datos más relevantes de manera resumida o detallada.

2.8.5 Herramientas de la Inteligencia de Negocios

Con el surgimiento de aplicaciones de Inteligencia de Negocios, se han creado herramientas especializadas que permiten mostrarle a la alta gerencia los datos más relevantes de una empresa de manera resumida, en forma de indicadores y reportes para la toma de decisiones.

Reportes: es un informe generado por un sistema, que nos presenta de forma estructurada y/o resumida, datos relevantes generados por las aplicaciones (ver Figura 2.12), de tal manera que se pueda realizar la toma de decisiones (SIPEC, 2011).

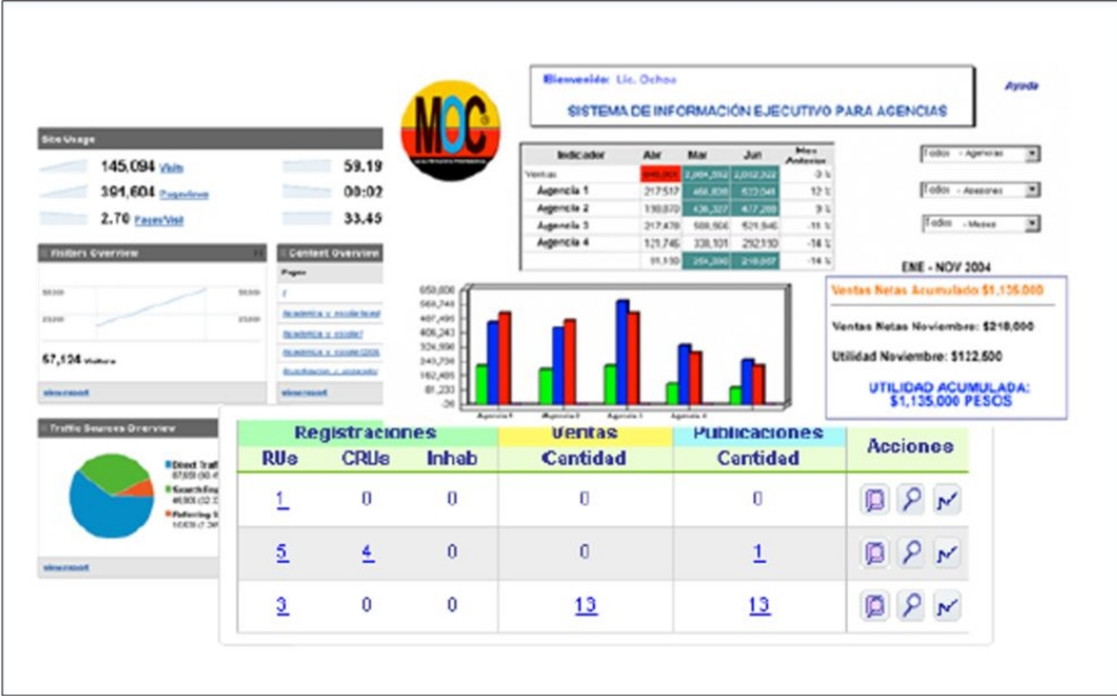


Figura 2.12: Ejemplos de Reportes (SIPPEC, 2011).

Indicador: es una medida cuantitativa o la observación cualitativa que permite identificar cambios en el tiempo (Figura 2.13), y cuyo propósito es determinar el buen funcionamiento de un sistema o negocio. Pudiendo así identificar los problemas llevando a cabo las medidas para solucionarlo (Vélez, 2000).

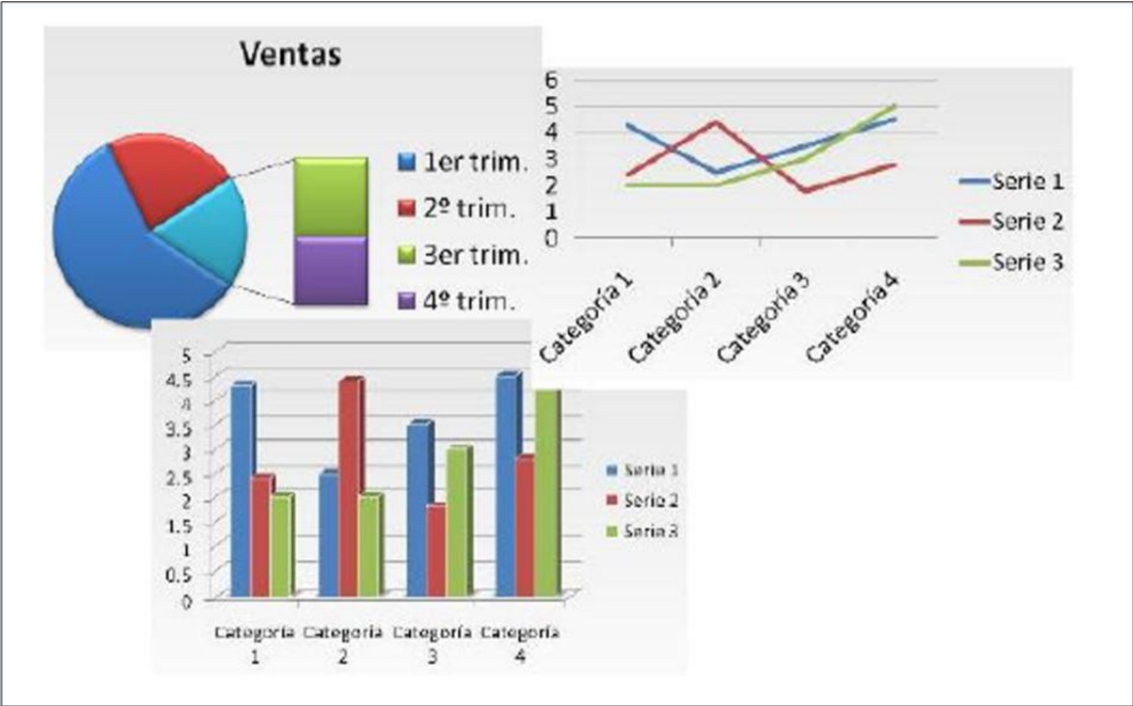


Figura 2.13: Ejemplos de Indicadores (Vélez, 2000).

2.9 Desarrollo de sistemas

Se tiene más de siete décadas desarrollando sistemas sobre computadoras programables de propósito específico y general, y para cada uno de ellos el proceso de desarrollo es, y probablemente será, cada vez más grande, costoso y complejo. Todo ello debido a la gran extensibilidad y adaptabilidad que ganan las computadoras con los años. Esto genera la necesidad de tener metodologías para el desarrollo de dichos sistemas que logren controlar con rigurosidad y a lo largo de cada una de sus fases, quien está haciendo que, cuando y como.

Según Elliott (2004), no fue sino hasta 1960 cuando surge la primera metodología para el desarrollo de software llamada Ciclo de Vida del Desarrollo de Sistemas (Systems Development Life Cycle, SDLC) también conocida como Ciclo de Vida del Desarrollo de Software, el cual está compuesto por cuatro fases fundamentales (Tegarden, Dennis & Wixom, 2009):

- La planificación, como proceso fundamental para entender el por qué se debe construir el sistema y como hará el equipo de desarrollo para construirlo. En un análisis de factibilidad se examinan los siguientes puntos clave: ¿Podemos construirlo?, ¿Proveerá valor al negocio?, ¿Será usado por el cliente? Una vez aprobada su construcción, se informa a gestión de proyecto para que realice el plan de trabajo y consiga a los miembros del equipo necesario para el desarrollo.
- El análisis, resuelve puntos fundamentales como: ¿Quién usará el sistema?, ¿Que hará el sistema?, ¿Dónde y cuándo será usado? Así mismo se debe investigar si ya existe algún otro sistema que se esté utilizando, que mejoras se pueden introducir al negocio, y desarrollar los conceptos del nuevo sistema.
- El diseño, fase en la que se decide cómo va a operar el sistema en términos de hardware, software, infraestructura en la red, interfaz de usuario, reportes de sistema, base de datos, archivos, etc. La información recopilada deriva en lo que se conoce como las especificaciones del sistema que son otorgadas al equipo de desarrollo para el inicio de la siguiente fase. Una vez que se termina el diseño, se revisa el análisis de factibilidad y la planificación previamente realizada para someter a votación la continuidad o no del desarrollo del sistema.
- La implementación donde el sistema es construido (o comprado si es que se trata de adquirir un paquete adicional del sistema que actualmente se está usando) instalado y mantenido. Es usualmente la más larga y costosa de todas las fases.

En la actualidad se cuenta con nueva metodología para sobrellevar los distintos tipos de problemas a la hora de desarrollar grandes y complejos sistemas informáticos. Tegarden (2009) nos asegura que podemos clasificar estas metodologías en tres grupos:

- Las metodologías de diseño estructurado, consideradas como las primeras en ser utilizadas en el desarrollo de software, provienen directamente del SDLC. Entre ellas tenemos, por ejemplo, el Desarrollo en Cascada y el Desarrollo Paralelo, que según el propio Tegarden, han sido predominantes a partir de los años 80, en sustitución de las metodologías con poca disciplina o personalizada (ad hoc).
- Las metodologías para el desarrollo rápido de aplicaciones (RapidApplication Development, RAD), emergieron en los años 90, tratando de resolver las debilidades de

las metodologías estructuradas ajustando las fases del Ciclo de Vida del Desarrollo de Sistemas con el fin de colocar rápidamente fragmentos del sistema desarrollado en manos del cliente, permitiéndole así tener una idea más clara del sistema para poder sugerir mejoras que realmente se adecuen a sus necesidades. La mayoría de las metodologías basadas en RAD sugieren que el analista utilice técnicas especiales y herramientas de desarrollo que aceleren las fases de análisis, diseño e implementación, tales como las CASE (Computer Aided Software Engineering), JAD (Joint Application Design) y los lenguajes de programación de cuarta generación.

- Las metodologías para el desarrollo ágil, se enfocan en simplificar y agilizar el Ciclo de Vida de Desarrollo de Sistemas eliminando parte del modelado y la documentación excesiva, así como el tiempo invertido en ambas actividades. Su énfasis se encuentra en el desarrollo simple e iterativo. Sus orígenes se remontan a los años 90, pero no fue sino hasta febrero del 2001 cuando un grupo de desarrolladores practicantes de esta metodología, decidieron crear el Manifiesto para el Desarrollo ágil de Software, en el cual se definen los estatutos y principios por el cual se debe regir cualquier metodología de desarrollo que se desee llamar ágil (Beck, K.,2001).

2.9.1 Metodologías para el desarrollo de sistemas de Inteligencia de Negocios

No cabe duda que las metodologías ya descritas son válidas para todos los desarrollos de sistemas. Sin embargo, al momento de desarrollar un Sistema de Inteligencia de Negocio se debe considerar en primera instancia los almacenes de datos, ya que son el pilar fundamental de la arquitectura. Esto lleva a Inmon, W. H (2002) a afirmar que, para desarrollar este tipo de sistema, se debe utilizar una metodología de desarrollo centrada en los datos y no centrada en los requerimientos como lo es el Ciclo de Vida del Desarrollo de Sistemas. Esto traza una línea difusa entre el desarrollo de un almacén de datos y el desarrollo de un Sistema de Inteligencia de Negocios, lo que implica que el hablar de uno implica al otro y viceversa.

Actualmente se cuenta con cuatro metodologías reconocidas para el desarrollo de los Almacenes de Datos / Sistemas de Inteligencia de Negocio, a saber: ascendente, descendente, híbrida y federada. Estas se pueden apreciar de forma resumida en la Tabla 2.3. Cabe destacar que, dada la particularidad de las organizaciones y los requerimientos de sus negocios, es posible que las soluciones que se desean desarrollar no utilicen ninguna de las metodologías mencionadas, sin embargo, son un excelente patrón para iniciar cualquier proyecto en dicha área.

Metodologías	Descendentes	Ascendentes	Híbridas	Federadas
Énfasis	Almacén de datos	Mercado de datos	Almacén de datos y Mercado de datos	Integración de entornos heterogéneos
Diseño	Modelo normalizado basado en el mercado de empresa	Modelo dimensional de datos, usa esquema estrella	Modelos locales, con uno o más esquemas estrellas	Se construye a través de la compartición de dimensiones, hechos, reglas a través de la organización
Arquitectura	Compuesto de varios niveles de áreas de interés y mercado de datos dependientes	área de interés y mercado de datos	Modelo empresarial normalizado de alto nivel; existen mercado de datos inicial	Realidad del cambio en organizaciones y sistemas
Datos	Almacén de datos:	Contiene datos	Carga mercado de datos con datos atómicos y	Uso de cualquier Significado posible

	Datos a nivel atómicos. Mercado de datos: Datos sumariados	atómicos y sumariados	sumariados sin utilizar el área temporal para los mismos	para integrar las necesidades del negocio
--	--	--------------------------	--	---

Tabla 2.3: Metodologías para el desarrollo de Sistemas de IN (Eckerson, s.f.)

2.9.2 Kimball Lifecycle

Publicada por primera vez en 1998 bajo el nombre The Business Dimensional Lifecycle, nace la primera metodología para el desarrollo de Sistemas de Inteligencia de Negocio de tipo ascendente, donde el almacén de datos empresarial es construido a partir de otros almacenes de datos especializados, que pertenecen a cada área funcional del negocio (Bodegas de datos).

Esta nueva aproximación de desarrollo tiene la gran ventaja de poder agregar nuevos mercados de datos al almacén empresarial según como se vayan incorporando nuevas áreas de negocio a la organización. Esto último hace que la metodología Kimball Lifecycle sea una de las metodologías preferidas para el desarrollo de sistemas de inteligencia de negocio.

La metodología se centra en el Modelado dimensional de los datos, donde Kimball, R (2011), proponen la construcción de un almacén de datos basándose en dos elementos claves: los hechos, que son los eventos generados de un proceso del negocio de la organización y las dimensiones, que son los datos necesarios para contextualizar el evento en el momento que ocurre un hecho. Dicho almacén de datos propuesto se caracteriza por tener dimensiones compartidas, también llamadas dimensiones conformadas, que son capaces de contextualizar a los distintos eventos derivados de las operaciones del negocio en una sola consulta analítica.

Para el desarrollo de un Sistema de Inteligencia de Negocios bajo la metodología Kimball Lifecycle se sugiere ejecutar 13 grupos de actividades, los cuales se pueden apreciar en la Figura 2.14 y que se detallan a continuación.

- Planificación del proyecto. Define la identidad del proyecto. Establece el alcance, los recursos necesarios, así como la definición de un plan de ejecución como de comunicación entre todos los participantes del proyecto.
- La gestión de proyecto, que se enfoca en el seguimiento y control de las actividades del mismo con el fin de mantenerlo dentro de los límites del alcance planteado.
- Definir requerimientos del negocio para la correcta recaudación de las especificaciones de una necesidad u oportunidad del negocio. En este sentido se hace indispensable que dichas especificaciones se documenten con claridad, sin ambigüedades y de forma consistente. Dicho así, los requerimientos del negocio son fundamentales para el buen entendimiento de las directrices que ha de tomar el proyecto a desarrollar.
- El Diseño técnico de la arquitectura enfocándose en la integración de tecnologías necesarias para el desarrollo del proyecto ofreciendo un marco conceptual de trabajo que permita a esta establecer una visión general de la arquitectura tomando en cuenta principalmente tres factores: requerimientos del negocio, tecnología actualmente disponible y las directrices de las estrategias técnicas planificadas.

- La selección de productos e instalación. Basándose en la arquitectura técnica, es posible hacer una lista de los componentes que el sistema ha de necesitar: El equipo físico, los DBMS, las herramientas para los procesos de migración de datos entre otros, han de ser evaluados, seleccionados, instalados y probados debidamente para asegurar una integración total del sistema a desarrollar.
- El Modelado dimensional, apoya las necesidades analíticas del negocio, todo ello mediante una aproximación distinta en el diseño de los datos que difiere de los sistemas transaccionales. Este proceso de modelado implica el uso de herramientas como la Matriz de Bus Empresarial y su Arquitectura asociada, en donde la segunda provee una aproximación incremental para la construcción de un sistema de inteligencia de negocio empresarial, descomponiendo los procesos de planificación en pedazos más manejables para enfocarse en los procesos de negocios, mientras que la primera es una herramienta esencial para el diseño de esta arquitectura, y que se presenta generalmente como un arreglo rectangular donde las celdas marcan la relación entre las filas (procesos) y las columnas (dimensiones), todo lo cual se muestra en la Tabla 2.4.

	DIM 1	DIM 2	DIM 3
Proceso 1	X	X	X
Proceso 2	X	X	
Proceso 3	X		X

Tabla 2.4: Ejemplo de la Matriz de Bus Empresarial

- El diseño físico en donde se busca la correcta definición de todas las estructuras necesarias que conformaran la base del sistema de inteligencia de negocio. Aquí se han de considerar aspectos como: la seguridad de los datos, estrategias de rendimiento, indexación, partición de tablas y agregación de las mismas.
- El diseño y desarrollo de los procesos de migración de datos que se enfoca en la migración de los datos operacionales (origen) hacia el almacén de datos (datos). Aquí se han de considerar aspectos como la integración, limpieza y estandarización de datos proveniente desde múltiples fuentes.
- El diseño del Sistema de Inteligencia de Negocios donde se idéntica las aplicaciones candidatas sobre las cuales se construirá el sistema, tomando en cuenta factores como la interfaz para la navegación de los datos y la capacidad de poder responder a las necesidades de sus usuarios.
- El desarrollo del Sistema de Inteligencia de Negocios centrado en la configuración del modelo de datos, las herramientas que dan soporte a la infraestructura y la construcción y validación del Sistema de Inteligencia de Negocios.
- El despliegue, como punto de encuentro de los tres caminos ya recorridos por la metodología: tecnología, datos y aplicaciones. Sumando y cohesionando todos los esfuerzos para la correcta puesta en marcha del sistema. Con el sistema en funcionamiento se hace necesario mantener un constante monitoreo sobre el comportamiento del mismo, para así garantizar su óptimo rendimiento. Este proceso de Mantenimiento debe considerar aspectos como: mejoras en el rendimiento, mantenimiento de los índices en la base de datos, respaldo de los datos y seguimiento de uso (monitoreo activo).

- Extensión. Todo sistema deberá tener la capacidad de expandirse acorde a las necesidades de la organización. Los sistemas de inteligencia de negocio no son lo excepción, por lo que deben incorporar cada vez con mayor ciencia todos aquellos nuevos elementos que agreguen valor al negocio.

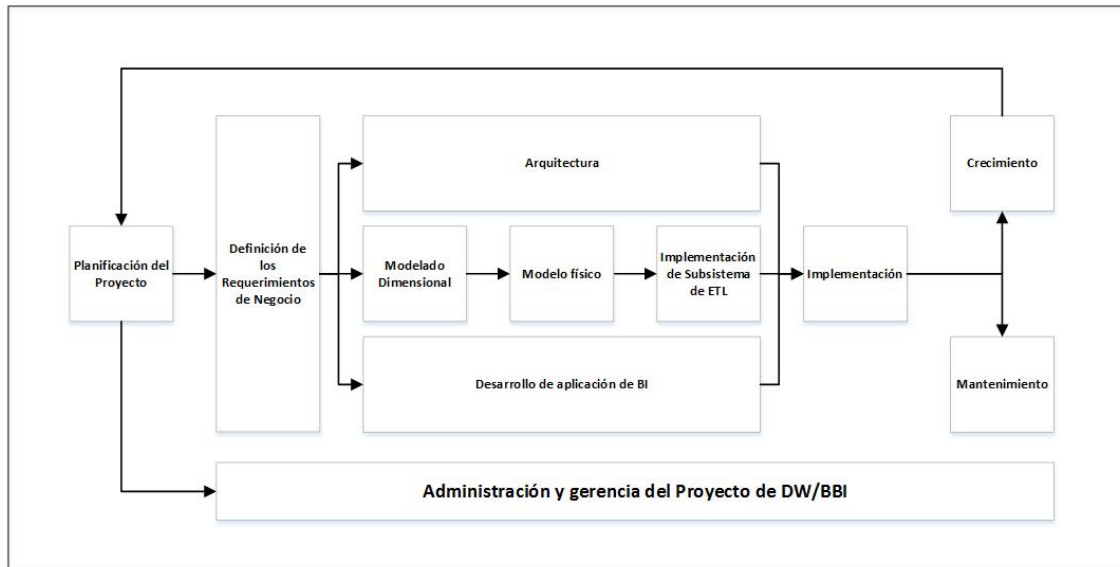


Figura 2.14: Diagrama de la metodología de Kimball Lifecycle (Kimball, R., 2011).

2.10 Metodología

La metodología que fue seleccionada para el desarrollo del Sistema de Inteligencia de Negocio fue la de Kimball Lifecycle, esta metodología gobernará los procesos metodológicos, aplicativos y a los efectos del desarrollo del presente TEG, debido a que se considera la más adecuada para el correspondiente caso de estudio.

2.11 Fases de la Investigación

La aplicación de cualquier metodología de investigación, especialmente en el área de las ciencias puede y debe ser considerada como un proyecto. En este sentido, se hace necesario decirlo, planificarlo y manejarlo. No se considera dentro del alcance de esta tesis profundizar en este tipo de aspectos, que además son propios del área de la Gerencia de Proyectos (Project Management, PM), sin embargo y a los efectos de aplicar esta metodología en trabajos de investigación se recomienda encarecidamente estudiar la Guía de los Fundamentos para la Dirección de Proyectos (Project Management Body Of Knowledge Guide, PMBOK) cuyo cuerpo de conocimiento proviene de las revisiones periódicas del gremio perteneciente al Instituto de Manejo de Proyectos (Project Management Institute, PMI), que viene a ser además uno de los estándares globalmente reconocido en el manejo de proyectos a nivel internacional.

Una vez se cuenta con estas técnicas, y con las tecnologías asociadas a ellas, se procede entonces a desarrollar las fases fundamentales para la conceptualización, creación y puesta en marcha del sistema de inteligencia de negocio.

De acuerdo con esto, se propone una segmentación de la anteriormente citada metodología de Kimball, dividida en 4 fases, según se aprecia en la Figura 2.15.

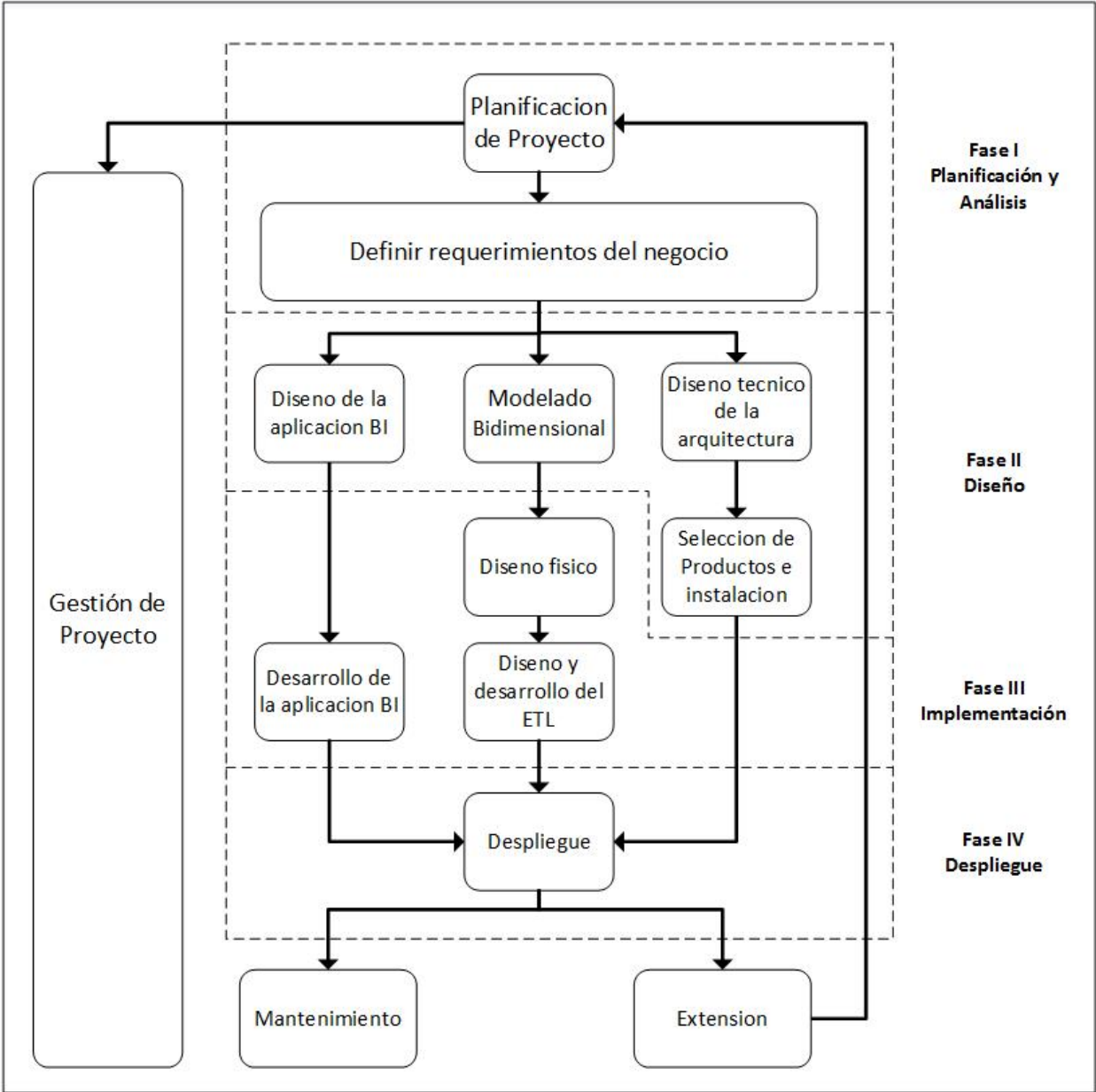


Figura 2.15: Fases del desarrollo utilizando la metodología Kimball Lifecycle

Se resume a continuación el concepto general de cada una de las fases, y se propone además que los próximos trabajos de investigación que tomen este como referencia se avoquen a interpretar de la manera más alineada posible con su proyecto cada una de las mismas.

2.11.1 Fase I: Planificación y Análisis

En cualquier proyecto y sobretodo en Sistemas de Inteligencia de Negocios la relación entre el desarrollador y el responsable por parte de la empresa de que se logren los objetivos del proyecto debe ser muy estrecha. Y esto es así desde la definición del proyecto, hasta la planificación del mismo. Este proceso de definición permite establecer los objetivos, acotar el alcance, reconocer las limitaciones iniciales o más evidentes, determinar las acciones o metas primarias, y crear el esquema preliminar de desarrollo; pasando por el caso de negocio y justificando el proyecto en cada una de sus concepciones.

Una vez el proyecto este enmarcado y definido, el próximo paso consiste en su planificación. Que generalmente, y de acuerdo con la bibliográfica, al menos contempla: la determinación de la Identidad del proyecto, el personal necesario, el plan del proyecto, y el plan de comunicaciones. Por último, y en consonancia con las acotaciones hechas más arriba, se procede a la Gestión, Conducción, Monitoreo, y Consolidación del proceso de PM de acuerdo a los estándares del PMBOK.

Toda vez que el Proyecto esté definido, configurado y en capacidad de gestionarse, se procede a realizar la captación de requerimientos, que en general se engloba en una primera aproximación obtenida en el tramo inicial de esta fase. Luego de que dicha aproximación se revise y se le hagan los ajustes pertinentes, se configurarían las entrevistas.

Es de suma importancia prestar especial atención al proceso de Entrevistas, pues de él se derivan todos los datos e información relevante a tomar en cuenta a los efectos de satisfacer expectativas, siempre con alcances realistas y ajustados a los acuerdos logrados entre el equipo desarrollador y la empresa. En cuanto a las entrevistas, se debe evaluar, por ejemplo: la modalidad cara-cara vs. automatizada, que se quiere y por qué, cual es el tipo de ambiente analítico en que se desenvolverá el conglomerado de herramientas DW/BI, cuales son los roles de los usuarios dentro de la empresa y en qué forma utilizaran la herramienta a los efectos de apoyar la toma de decisiones, que tipo de entrevistas se adecuan mejor al negocio, Funcionales o de Tecnología, desarrollar los cuestionarios, configurar la entrevistas, seleccionar el personal entrevistado, además de las muchas otras a las que hubiere lugar.

Finalmente, los resultados de estas entrevistas deben ser contrastados con su respectivo feedback, analizando los hallazgos, documentándolos y publicándolos como derivables de la FASE I.

2.11.2 Fase II: Diseño

Solo cuando la FASE I ha sido declarada como finalizada por el Responsable del Proyecto, es entonces cuando se puede comenzar con la FASE II: Diseño. Es importante resaltar que, en la metodología escogida para este TEG, específicamente a nivel de Diseño- Implementación (FASE II y FASE III en este arreglo) se pueden identificar tres caminos: el de la aplicación BI, el tecnológico, y el de los datos.

En el camino de la aplicación BI, concretamente en la FASE II, se conceptualizarán y se decidirán las plantillas de presentación de los datos, el alcance de la información, el nivel de profundidad, las variables espaciales de reporte a nivel grafico (pantalla, diagramas, etc.), además de temas de seguridad relacionados con el acceso, por ejemplo, de reportes. Todo ello tomando siempre en cuenta la creación de estándares de desarrollo y el metadato asociada; lo que terminara

derivando en la creación de formatos para la organización que evolucionara dentro de la misma y con el tiempo.

Luego de esto, vendrán las primeras especificaciones técnicas detalladas para el desarrollo de la aplicación BI (en la FASE III), contando siempre con la verificación de los usuarios finales.

A los efectos del camino tecnológico, explicaremos los tres bloques principales que componen la Figura 2.16.

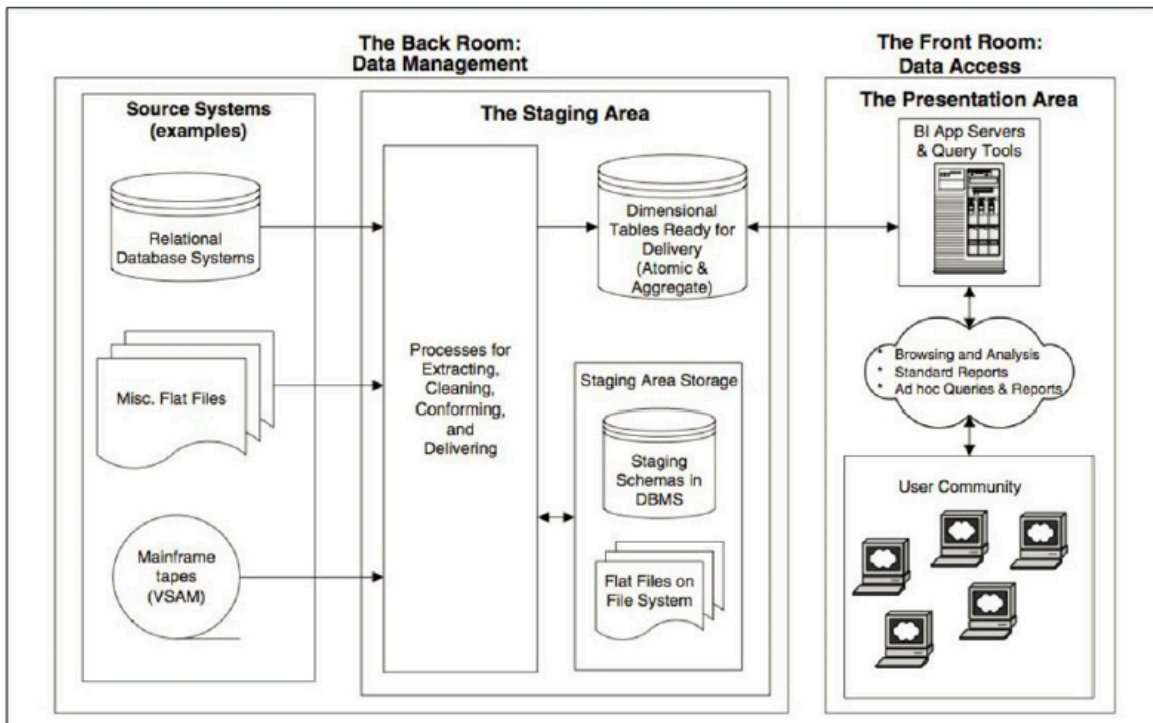


Figura 2.16: Back Room y Front Room de un DW. (Kimball & Caserta, 2004)

El sistema DW/BI puede ser diagramado tomando en cuenta dos puntos de vista, el Back Room, y el Front Room. El primero relacionado con los procesos de Diseño- Desarrollo-Operación del lado de la gobernación de los datos, y el segundo del lado de los usuarios finales.

Teniendo en cuenta esto se puede apreciar el flujo de información, los componentes que gestionan ese flujo, el almacén y la forma en que los usuarios tienen acceso a ella.

De acuerdo con esto, es evidente que el enfoque cambia de ser teórico alrededor de los componentes y servicios de arquitectura hacia los asuntos prácticos de como diseñar realmente la misma. Kimball et al. (2008) nos advierte que El plan de arquitectura es la traducción técnica de los requerimientos del negocio. De acuerdo con esto las propuestas tienen que ser capaces de satisfacer las necesidades de información de la empresa, donde la arquitectura y la selección de productos asociada a ella realce su importancia debido a su in usencia en el mapa de infraestructura, el cual resulta de la culminación de la fase de diseño.

El tercer camino, el de los datos, viene representado específicamente por el Modelado Dimensional del Almacén de Datos, donde se utiliza la técnica de diseño dimensional de cuatro pasos propuesta por Kimball y Ross (2013):

1. Selección del proceso del negocio a modelar. El criterio de selección de los procesos a modelar responde únicamente a las iniciativas estratégicas del negocio. Dichos procesos son los que realizan las actividades operacionales de la organización. Generan o capturan los eventos que luego son traducidos en hechos en las tablas de hechos. Dichas tablas siempre están asociadas al resultado de un proceso único, donde cada proceso de negocio a modelar, debe ser registrado en la matriz empresarial del almacén de datos.
2. Declaración del grano de dicho proceso. El grano representa exactamente lo que se debe guardar en tabla de hechos. Debe ser declarado antes de las dimensiones o tabla de hechos porque que las dimensiones o los hechos deben ser consistente con el grano declarado. Este es uno de los pasos más importantes en el desarrollo dimensional y ha de ser realizado con cuidado.
3. Identificar las dimensiones. Las dimensiones contextualizan los eventos ocurridos en los procesos del negocio. Contienen atributos descriptibles que pueden ser utilizados por los sistemas de inteligencia de negocios para responder a las siguientes preguntas: quien, que, donde, cuando, por qué y cómo.
4. Idéntica los hechos. Los hechos son medidas realizadas sobre un proceso del negocio. Usualmente son numéricos. Un registro en la tabla de hechos tiene una relación 1 a 1 con la medida de un evento descrita en el grano del proceso.

Como consideración general y de acuerdo con Kimball, R. (2008) las aplicaciones BI son el puente entre la organización y la ventaja táctica sobre la información dentro de los sistemas DW/BI. En la mayoría de los casos existe un período de espera significativo entre la finalización de la captación de los requerimientos y el desarrollo de la solución BI, bien sea completa o integrada, ya que es necesario tener el modelado dimensional y al menos uno o varios subconjuntos de datos, donde se tenga la Matriz de Bus Empresarial completamente identificadas, así como detallada su arquitectura. Para construir soluciones robustas lo recomendable es que el equipo de diseño tenga a su disposición personal dedicado y específico por área de requerimiento tanto empresarial como de desarrollo, que comprendan el negocio y que sean conscientes del efecto que los aspectos de idiosincrasia-prejuicio-opinión generan sobre el equipo de trabajo.

2.11.3 Fase III: Implementación

De acuerdo con la FASE II, existen tres caminos, pero a los efectos de la implementación solo quedan dos: el de la aplicación BI y el de los datos. El tecnológico, ha culminado completamente en la Fase de Diseño.

Desarrollo de la Aplicación BI

Siguiendo el camino de la aplicación BI, se procede entonces a preparar el desarrollo o configuración del Sistema de Inteligencia de Negocios, se instalan las herramientas o suits de herramientas, se prueban y verifican las estrategias de interfaz, y se configuran los parámetros de seguridad de los usuarios

Diseño Físico

A nivel metodológico se entenderá que el Diseño Físico es realmente el proceso de modelar, desarrollar y construir el modelo físico de los datos atendiendo sus aspectos tanto lógicos (modelado de datos) como físicos (disposición física de datos). En este sentido es importante señalar que aun cuando la metodología identifica este proceso como físico, lo preciso en este caso sería distinguir y contemplar la transversalidad del proceso a través de los tres niveles: el externo, el conceptual y el interno, de manera que siempre se tenga en cuenta estos factores con respecto a los datos: el modelado, como se van consultar, como se van almacenar, como se van a acceder y por qué se almacenan así.

Teniendo el modelo dimensional en mano, se necesita centrar la atención en el diseño y creación de las estructuras de datos necesarias para el soporte de los mismos. Hay que tener en cuenta que en el ámbito de los almacenes de datos un modelo dimensional aplicado en un DBMS relacional se le conoce como esquema estrella, mientras que si lo fuera en un Servidor OLAP se le ubicará dentro de los cubos de información.

Justo antes de que los datos puedan ser consultados desde la herramienta analítica, se necesita venir el modelo dimensional ya planteado dentro de servidor OLAP, el cual servirá de traductor entre el almacén de datos y las consultas multidimensionales realizadas por los usuarios sobre los datos existentes.

Diseño y desarrollo de los procesos de extracción, transformación y carga

Lograr una correcta construcción del Sistema de Extracción, Transformación y Carga (Extract, Transform, Load, ETL) generalmente redundante en una tarea desafiante, debido a los factores y variables de entorno que hay que tomar en cuenta, entre los cuales están por ejemplo el volumen y la complejidad de los datos.

Sin embargo, y según la metodología seleccionada, el proceso de Diseño y Desarrollo de ETL se puede resumir en 4 áreas claves, las cuales gobiernan dicho proceso y que son: extracción de datos, adecuación de los mismos (limpieza y enriquecimiento), entrega de la data para su posterior consulta, y manejo de del entorno ETL. El considerar todos estos factores garantiza en gran medida el éxito.

De hecho, Kimball, R. (2008) resalta que el diseño y desarrollo de ETL es uno de los mayores retos del proyecto DW/BI. No importa cuán bien fueron platicadas, diagramadas y ejecutadas las entrevistas o los análisis, la calidad de los datos siempre y en cualquier caso puede representar un problema. De hecho, estos problemas pueden llegar a ser tan graves que se vean afectados incluso el alcance y los objetivos del proyecto.

En resumen, los procesos ETL más que extracción, transformación y carga, son un conglomerado de tareas complejas e importantes, que representan con frecuencia los obstáculos a vencer y de mayor envergadura dentro de los Proyecto DW/BI.

El proceso general de migración de datos ocurre de la siguiente forma (Ver Figura 2.17):

1. Inicio: fase en la cual se prepara el ambiente de ejecución de los procesos de ETL. La misma realiza las siguientes actividades: verifica la terminación exitosa de la ejecución anterior con el fin de saber si debe proceder a una nueva ejecución o no. Limpia los

archivos temporales de las corridas pasadas y define las rutas de acceso en el ambiente de los ETLs.

2. Punto de control: verifica si la ejecución actual es la continuación de una corrida anterior. De ser así salta directamente a la fase de carga de los datos del almacén.
3. Extracción + Transformación: los datos son extraídos, limpiados, validados, transformados y guardados en archivos listos para su carga.
4. Carga: fase encargada de insertar los datos en el almacén de datos.
5. Cierre: verifica la culminación de los procesos de ETL. En caso de ser exitosa, archiva la bitácora (también conocido como el archivo log) de ejecución en conjunto con los datos extrados en el proceso. Caso contrario deja intacto el área de trabajo y comunica al personal operativo la ocurrencia de una falla en la migración de datos.

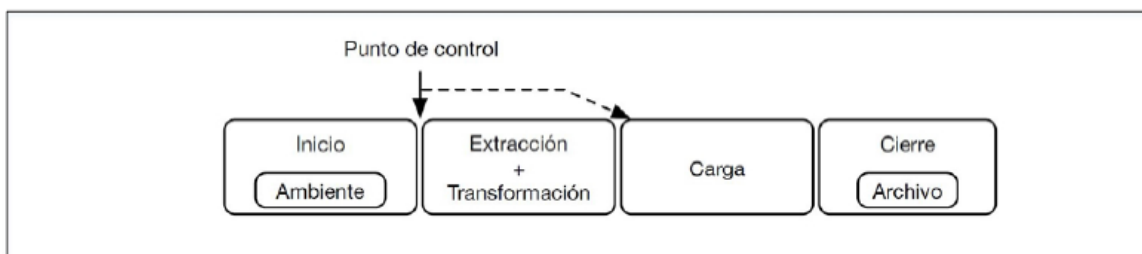


Figura 2.17: Fases del proceso de extracción transformación y carga

2.11.4 Fase IV: Despliegue

La planificación y coordinación detallada de los esfuerzos de desarrollo, son los factores críticos a considerar en la FASE de Despliegue. En donde se prueban, de diferentes maneras, las aplicaciones y los componentes del desarrollo DW/BI, incluyendo: pruebas de aceptación, de calidad de datos, de procesos operacionales, de rendimiento, de utilidad, de despliegue, y de capacidad en entornos de usuario

Una vez superadas las pruebas con éxito, se procede al despliegue en sí mismo. Más allá de esto la FASE IV se concentra en poner en operación la solución DW/BI así como el adiestramiento y acompañamiento del usuario.

Finalmente, la solución implementada y puesta en producción debe centrarse en gestionar y mantener tanto el Back Room como el Front Room, manejar eficiente y continuamente las variables de entorno, acelerar su integración como herramienta de utilidad y uso cotidiano entre los usuarios del negocio, y prepararse para su crecimiento, evolución, y retro compatibilidad, aprovechando las oportunidades de expansión y fusionándose cada vez más con el modelo y los procesos de negocio.

Capítulo III - Marco Aplicativo

En el presente capítulo describiremos única y específicamente las actividades realizadas durante el desarrollo de la solución planteada, las mismas están enmarcadas dentro de la metodología Kimball Lifecycle, el cual funge como marco referencial para el Sistema de Inteligencia de Negocios desarrollado en este T.E.G.

3.1 Fase I: Planificación y Análisis

Se estableció, en primera instancia, una reunión entre el personal de la organización y el equipo desarrollador donde se establecieron los criterios generales que regirían el desarrollo del proyecto, entre los cuales estuvieron: alcance, metodología, situación actual del sistema y canales de comunicación a ser usados.

Se conformó un equipo de desarrollo que constaba de 3 personas adicionales a los dos desarrolladores del presente trabajo de investigación, todas por parte de la organización, de las cuales 2 son del área de tecnología y 1 de la gerencia mercadeo.

EDT	Tarea	Duración Pronosticada	Inicio Pronosticado	Fin Pronosticado
#	Nombre de la actividad	Días	Fecha inicio	Fecha Fin

Tabla 3.1: Plantilla para el registro de actividades a realizar

Una vez conformado el equipo y con el Plan de Comunicación en operación, se propuso la creación de la lista de tareas a desarrollar, basada en las estimaciones de las reuniones preliminares. En este sentido, se configuró un grupo de actividades que permitiera, en principio, el logro de los objetivos del proyecto. Esta lista de actividades se muestra según plantilla Tabla 3.1 en la Tabla 3.2

EDT	Tarea	Duración Pronosticada	Inicio Pronosticado	Fin Pronosticado
1	Implementación de un ambiente de BI	96.65 días	jue 20/6/13	mar 5/11/13
1.1	Inicio del Proyecto	2 hrs	jue 20/6/13	jue 20/6/13
1.2	Hito: Reunión de Arranque del Proyecto	8 hrs	jue 20/6/13	jue 20/6/13
1.3	Especificación de Requerimientos	27.5 días	vie 21/6/13	jue 1/8/13
1.3.1	Recolección de los reporte actuales	4 días	vie 21/6/13	jue 27/6/13
1.3.1.1	Reunión con el Gte. Financiero	2 días	vie 21/6/13	mar 25/6/13
1.3.1.2	Reunión con la Gcia. De Mercadeo	2 días	mié 26/6/13	jue 27/6/13
1.3.2	Recolección de las necesidades de la gerencia	10 días	mié 26/6/13	mié 10/7/13
1.3.2.1	Reunión con la Gcia. De Finanzas	10 días	mié 26/6/13	mié 10/7/13
1.3.2.2	Reunión con la Gcia. De Mercadeo	1 día	mié 3/7/13	mié 3/7/13
1.3.3	Plataforma Tecnológica	6 días	mié 3/7/13	jue 11/7/13
1.3.3.1	Evaluación de la plataforma tecnológica	8 hrs	mié 3/7/13	mié 3/7/13
1.3.3.2	Elaborar la Hoja de Configuración	8 hrs	jue 11/7/13	jue 11/7/13
1.3.3.3	Hito: Hoja de Configuración completa	0 hrs	jue 11/7/13	jue 11/7/13
1.3.4	Mesa de trabajo para evaluación de requerimientos	14.5 días	vie 12/7/13	jue 1/8/13
1.3.4.1	Revisar información recopilada	1 día	vie 12/7/13	vie 12/7/13
1.3.4.2	Reunión #1 de validación con Experto	2 hrs	mié 17/7/13	mié 17/7/13

Tabla 3.2: Extracto plantilla para el registro de actividades a realizar

A los efectos de lograr una retroalimentación constante, fluida, y oportuna, se dispuso de un *Plan de Comunicación* basado en video conferencias y mensajería a distancia, con una frecuencia de comunicación de 3 a 5 veces por semana y de alrededor 6 horas de duración, todo esto debido a la complejidad del proyecto. Durante estas jornadas comunicacionales se realizaron entrevistas presenciales por video conferencia donde se registraron las minutas de reunión. Dado la gran disponibilidad de tiempo del personal de la empresa (uno asignado a tiempo completo y dos con cita) no fue necesaria la diagramación de entrevistas, sin embargo, era muy común preparar o elaborar cuestionarios para usarlos como puntos de cuenta y guía durante las mismas, las cuales derivaban posteriormente en minutas que contenían sus respectivos acuerdos y compromisos.

3.1.1 Requerimientos Capturados

En este sentido se muestran a continuación los requerimientos de información analítica capturados en función de los reportes requeridos por la organización.

Productividad por centro de llamadas

- Formato: tabla
- Descripción: medición por cada Call Center, la cantidad de llamadas recibidas, atendidas y perdidas
- Medida: número de llamadas (número entero)
- Filtro:
 - Período de tiempo (Y): por minuto, por hora, diario, semanal y mensual
 - Tipo de llamadas (Z): recibidas, atendidas y perdidas
 - Call Center (X): nombre del Call Center
- Diseño: formato del reporte en forma de tabla, los gráficos se definirán en el momento de la elaboración del reporte

Total de ventas generadas por código-corto en un espacio de tiempo determinado

- Formato: tabla
- Descripción: medición del total de bolívares fuertes facturados por el lado de WTFE por shortcode, durante un periodo de tiempo
- Medida: moneda (BsF)
- Filtro:
 - Período de tiempo (Y): diario, semanal, mensual y anual
 - Operadoras (Z): Digitel, Movistar y Movilnet
 - Shortcode (X)
- Diseño: formato del reporte en forma de tabla, los gráficos se definirán en el momento de la elaboración del reporte

Top 1000 de clientes en consumos por espacio de tiempo

- Formato: tabla
- Descripción: identificación de los mejores 1000 clientes, los cuales se medirán sumando el total en BsF consumidos mediante el consumo de mensajes SMS y el total de minutos

de llamadas 0900, incluyendo datos del consumo por asociado, como también se podrá filtrar el total de consumo en BsF o el mismo total en cantidades

- Medida: moneda (BsF)
- Filtro:
 - Período de tiempo (Y): mes, trimestre, año
 - Cliente
 - BSF - Totalconsumo (Z)
- Diseño: formato del reporte en forma de tabla, los gráficos se definirán en el momento de la elaboración del reporte

Duración promedio de las llamadas cortas y largas a los 0900

- Formato: tabla
- Descripción: Este indicador muestra el promedio de minutos consumidos en cada línea 0900, dividido en llamadas cortas que son aquellas que su duración es menor a 3 minutos y la llamada larga las cuales tienen una duración mayor a 3 minutos
- Medida: minutos (0900)
- Filtro:
 - Período de tiempo (Y): semanal y mensual
 - Numero 0900 (X)
 - Tipo de llamadas (Z): Cortas o largas
- Diseño: formato del reporte en forma de tabla, los gráficos se definirán en el momento de la elaboración del reporte

Comparativa de metas mensuales por unidad de negocios vs. Facturación mensual por unidad de negocios

- Formato: grafico de barras
- Descripción: comparación del presupuesto estimado para las unidades de negocios (SMS, 0900 y pégate 503) con el facturado por unidad de negocios (SMS- 0900) en bolívares fuertes
- Medida: moneda (BsF)
- Filtro:
 - Unidad de negocio (X): SMS, Mensajería 0900
- Diseño: formato del reporte en forma de tabla, los gráficos se definirán en el momento de la elaboración del reporte

Cantidad de clientes activos y desafiliaciones mensuales en Infoline

- Formato: grafico de barras
- Descripción: cantidad de clientes activos en el servicio Infoline y la cantidad de desafiliaciones
- Medida: mensual
- Filtro:
 - Período de tiempo (Y): mensual
 - Status (Z): activo, inactivo

- Diseño: formato del reporte en forma de tabla, los gráficos se definirán en el momento de la elaboración del reporte

Clientes nuevos para la empresa

- Formato: tabla
- Descripción: medir los números (ANI) que empiecen a usar los servicios ofrecidos por la empresa (WTFE)
- Medida: número de los clientes (ANI)
- Filtro:
 - Período de tiempo (Y): día, semana, mes y año
- Diseño: formato del reporte en forma de tabla, los gráficos se definirán en el momento de la elaboración del reporte

Números (ANI) en mensajes por el 530 (pégate)

- Formato: tabla
- Descripción: Monitorear los números que empiezan a mandar mensajes a través del 530 (pégate) y los números (ANI) que dejaron de enviar SMS por las operadoras y lo empezaron a hacer mediante el servicio 530.
- Medida: número de clientes (ANI)
- Filtro:
 - Período de tiempo (Y): día, semana, mes, año
 - Números ANI (X): Clientes nuevos y clientes que dejaron de enviar SMS por las operadoras
- Diseño: formato del reporte en forma de tabla, los gráficos se definirán en el momento de la elaboración del reporte

Contenido consumido por número de teléfono (ANI)

- Formato: grafico de barras
- Descripción: Monitorear el número de mensajes y el número de minutos que consume un cliente por asociados, filtrados por etiqueta (palabra clave) y por rubros.
- Medida: Mensajes y minutos
- Filtro:
 - Período de tiempo (Y): mes, año
 - Clientes (X)
 - Asociado (Z)
- Diseño: formato del reporte en forma de tabla, los gráficos se definirán en el momento de la elaboración del reporte

Inversión publicitaria invertida por asociado

- Formato: tabla
- Descripción: Inversión publicitaria invertida por asociado en donde un asociado está vinculado a un ejecutivo de ventas

- Medida: BsF (moneda)
- Filtro:
 - Período de tiempo (Y): mes, año
 - Asociado (X)
 - Ejecutivo de Ventas (Z)
- Diseño: formato del reporte en forma de tabla, los gráficos se definirán en el momento de la elaboración del reporte

Desafiliaciones de clientes para la empresa

- Formato: tabla
- Descripción: medir los números (ANI) que dejan de usar los servicios ofrecidos por la empresa (WTFE)
- Medida: número de los clientes (ANI)
- Filtro:
 - Período de Tiempo (X): Mes y año
- Diseño: formato del reporte en forma de tabla, los gráficos se definirán en el momento de la elaboración del reporte

3.2 Fase II: Diseño

3.2.1 Diseño de la aplicación de Inteligencia de Negocio

Dado a que el proyecto estuvo centrado especialmente en las capacidades analíticas del sistema, la configuración de las opciones de visualización de los Reportes Analíticos, así como su formato no fue considerada como fundamental para este desarrollo. Asimismo, y de acuerdo con el balance costo-beneficio considerado, la incorporación de los aspectos de seguridad tampoco fueron incluidos, por lo cual este grupo de actividades se vio reducido al mínimo esfuerzo necesario para decidir el curso de acción del desarrollo del Sistema de Inteligencia de Negocios, es decir, basarse única y exclusivamente en el *Camino Tecnológico* y el de los *Datos*.

3.2.2 Diseño Técnico de la Arquitectura

Para el desarrollo del sistema se conformó una arquitectura clásica de tres niveles acorde con la forma en que los datos serían manipulados. Dicha arquitectura puede ser observada en la Figura 3.1 y se describe de la siguiente manera:

- El *primer nivel*, se encargó de la manipulación cruda de los datos, los cuales una vez extraídos de los sistemas operacionales fueron preparados para su disposición y posterior consumo.
- El *segundo nivel*, correspondiente a los servidores que apoyan al análisis de los datos, el almacén de datos administrado por el DBMS bajo un esquema estrella y el servidor OLAP el cual provee los análisis multidimensionales sobre los datos dispuestos en el almacén.
- El *tercer nivel* presenta a los usuarios del sistema, una interfaz de consulta de los datos que son ofrecidos por el segundo nivel. En dicha interfaz se realizan las consultas relevantes en forma de reportes analíticos.

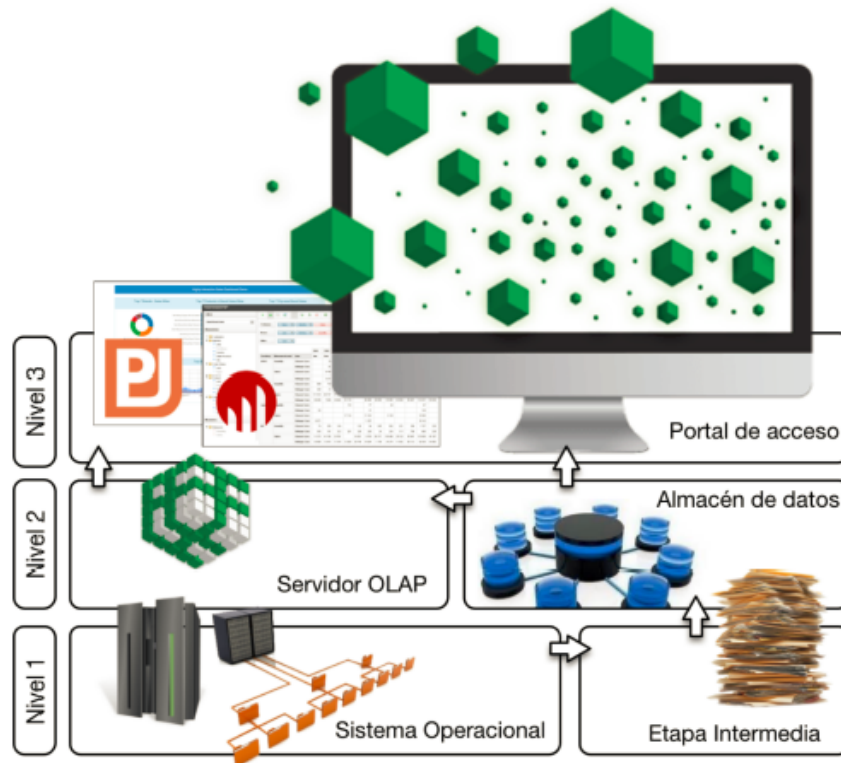


Figura 3.1: Arquitectura conceptual del Sistema de Inteligencia de Negocios

3.2.3 Selección de productos e instalación

A los efectos de la selección de software se lista a continuación cada elemento contenido de la solución desarrollada, así como las razones de dicha selección:

- El módulo Pentaho Data Integration de la Suit Comunitaria de Pentaho para Inteligencia de Negocios (PentahoBI Suite ComunnityEdition, PBICE), ya que la versión licenciada de Oracle Warehouse Builder 11G(OWB) con la que contaba la organización presentaba fallos a los efectos de los procesos de ETL, Oracle Warehous Builder (OWB) para la creación de los esquemas del servidor OLAP y para lo relacionado con las consultas multidimensionales de los datos, así como el servidor BI conocido como Oracle Bussines Intelligence (OBIEE) que sirve como portal de acceso.
- Oracle Business Intelligence Publisher 11G, como herramienta de consulta y para la generación de reportes analíticos.
- Oracle data Base 11G R2, como DBMS para el resguardo del almacén de datos, así como registros temporales de los procesos de ETL
- Oracle Linux 5, como sistema operativo (Operating System, OS) seleccionado como el ambiente de ejecución del Oracle Warehouse builder

- Windows Server 2008 R2, como sistema operativo seleccionado como ambiente de la ejecución del OBIEE
- VMware Vsphere 5.0, como hipervisor y herramienta de virtualización base de los ambientes anteriormente mencionados.

Es importante señalar que los SO fueron requisito obligatorio en el uso de la configuración de la solución BI dado que el ente decisor poseía un cocimiento operativo basado en estas tecnologías, por lo que era más que deseable aprovechar la experiencia del personal en el manejo y mantenimiento de herramientas ya conocidas.

Por otra parte, la selección del VMware Vsphere, le da a la solución un nivel de respaldo adicional a los configurados en la base de datos ya que se hizo uso del mismo para el completo respaldo de los ambientes instalados.

3.3 Modelado dimensional

De acuerdo con la metodología, y una vez descartados gran parte de los reportes por falta de datos en el origen, se detectaron 6 Dimensiones y unas 5 Tablas de Hechos. Ambas se relacionan en la Tabla 3.3 que representa en definitiva la *Matriz del Bus Empresarial*. Asimismo, cabe señalar que las Dimensiones Degeneradas no deben ser contempladas en la referida matriz.

	FACT_SMS Generados	FACT_0900	FACT_INFOLINE	FACT CALL_CENTER	FACT PUBLICIDAD
DIM_DATE	✓	✓	✓	✓	✓
DIM_COMPANIA	✓	✓	✓	✓	✓
DIM_ASOCIADO	✓	✓	✓	✓	✓
DIM_SERVICIO	✓	✓	✓	✓	✓
DIM_CALL_CENTER				✓	
DIM_GERENTE					✓

Tabla 3.3: Matriz del Bus Empresarial

3.3.1 Dimensiones

A continuación, se especifica en detalle las Dimensiones (por atributos y jerarquía) de la Matriz de Bus Empresarial:

DIM_DATE: Es la encargada de manejar todas las fechas de los reportes, descrita en la Tabla 3.4

Nombre del Atributo	Contenido	
	Descripción	Tipo
DATE_ID	Clave primaria de la dimensión	Entero
DATE_DATE	Fecha Completa	Date
DATE_YEAR_NUMBER	Año Calendario	Entero

DATE_DAY_OF_YEAR_NUMBER	Día del año	Entero
DATE_DAY_OF_MONTH_NUMBER	Día del mes	Entero
DATE_DAY_OF_WEEK_NUMBER	Día de la semana	Entero
DATE_DAY_OF_WEEK_NAME	Nombre del día de la semana	Cadena
DATE_DAY_OF_WEEK_NAME_SHORT	Diminutivo del día de la semana	Cadena
DATE_DAY_OF_WEEKEND_INDICATOR	Indica si el día es fin de semana o no	Carácter
DATE_MONTH_NUMBER	Número del Mes	Entero
DATE_MONTH_NAME	Nombre del mes	Cadena
DATE_MONTH_NAME_SHORT	Diminutivo del nombre del mes	Cadena
DATE_WEEK_OF_YEAR_NUMBER	Número de la semana del año	Entero
DATE_WEEK_OF_YEAR_NAME	Nombre de la semana del año	Cadena
DATE_WEEK_OF_MONTH_NUMBER	Número de la semana del mes	Entero
DATE_WEEK_OF_MONTH_NAME	Nombre de la semana del mes	Cadena
DATE_QUARTER_NUMBER	Número del Trimestre	Entero
DATE_QUARTER_NAME	Nombre del trimestre	Cadena
DATE_QUARTER_YEAR_NAME	Nombre del trimestre + número del año	Cadena
DATE_SEMESTER_NUMBER	Número del semestre	Entero
DATE_SEMESTER_NAME	Nombre del semestre	Cadena
DATE_SEMESTER_YEAR_NAME	Nombre de semestre del año + número del año	Cadena

Tabla 3.4: Dimensión Date

DIM_COMPANIA: En la dimensión compañía descrita en la Tabla 3.5 se guardarán los registros de todas las empresas que interactúen con el proceso.

Nombre del Atributo	Contenido	
	Descripción	Tipo
COMPANIA_ID	Clave primaria de la dimensión	Entero
COMPANIA_NUMBER	Número de la compañía	Entero
COMPANIA_NOMBRE	Nombre de la compañía	Cadena
COMPANIA_TIPO	Tipo de la compañía	Cadena

Tabla 3.5: Dimensión Compañía

DIM_ASOCIADO: La dimensión asociada descrita en la Tabla 3.6 se encargará de registrar todos los clientes o asociados de la empresa.

Nombre del Atributo	Contenido	
	Descripción	Tipo
ASOCIADO_ID	Clave primaria de la dimensión	Entero
ASOCIADO_NUMBER	Número telefónico del asociado	Entero
ASOCIADO_NOMBRE	Nombre del asociado	Cadena
ASOCIADO_RIF	Rif del asociado	Cadena
ASOCIADO_TELEFONO	Teléfono del asociado	Entero
GERENTE_ID	La clave foránea del gerente encargado del asociado	Entero

Tabla 3.6: Dimensión Asociado

DIM_SERVICIO: La dimensión servicio descrita en la Tabla 3.7 se encargará de registrar todos los tipos de servicio que puede tener un asociado.

Nombre del Atributo	Contenido	
	Descripción	Tipo
SERVICIO_ID	Clave primaria del servicio	Entero
SERVICIO_NUMBER	Código del servicio	Entero
SERVICIO_NOMBRE	Nombre del servicio	Cadena

SERVICIO_DESCRIPCION	Descripción del servicio	Cadena
----------------------	--------------------------	--------

Tabla 3.7: Dimensión Servicio

DIM_CALL_CENTER: La dimensión descrita en la Tabla 3.8 call center es la encargada de registrar los call center de la organización

Nombre del Atributo	Contenido	
	Descripción	Tipo
CALL_CENTER_ID	Clave primaria de la dimensión	Entero
CALL_CENTER_NUMBER	El número del Call center	Entero
CALL_CENTER_CODIGO	El código asociado al Call center	Cadena
CALL_CENTER_NOMBRE	El nombre del call center	Cadena

Tabla 3.8: Dimensión Call Center

DIM GERENTE: La dimensión gerente descrita en la Tabla 3.9 se encarga de registrar los gerentes de cuentas que llevan asociados en la empresa.

Nombre del Atributo	Contenido	
	Descripción	Tipo
GERENTE_ID	Clave primaria de la dimensión	Entero
GERENTE_CEDULA	Cedula del gerente de cuenta	Entero
GERENTE_NOMBRE	Nombre del gerente de cuenta	Cadena

Tabla 3.9: Dimensión Gerente

3.3.2 Hechos

A continuación, se especifica en detalle los Hechos (por atributos y jerarquía) de la Matriz de Bus Empresarial:

FACT_SMS_GENERADOS: La tabla de hecho descrita en la Tabla 3.9 se encarga de registrar las transacciones asociadas a los mensajes SMS enviados por cada cliente.

Nombre del Atributo	Contenido	
	Descripción	Tipo
FACT_SMS_ID	Clave primaria de la tabla de hechos	Entero
DIM_DATE	Clave foránea de la dimensión DATE	Entero
DIM_COMPANIA	Clave foránea de la dimensión COMPANIA	Cadena
DIM_SERVICIO	Clave foránea de la dimensión SERVICIO	Cadena
DIM_ASOCIADO	Clave foránea de la dimensión ASOCIADO	Entero
SHORTCODE_NUMBER	registrará el short-code al cual fue enviado el SMS	Entero
CLIENTE_NUMERO	Este atributo registra el número del cliente que envió el SMS	Entero
PRECIO	Este atributo registra el precio del SMS enviado	Decimal

Tabla 3.10: Tabla de hechos SMS generados

FACT_0900: La tabla de hecho descrita en la Tabla 3.9 se encarga de registrar las transacciones asociadas a las llamadas realizadas a los números 0900.

Nombre del Atributo	Contenido	
	Descripción	Tipo
FACT_0900_ID	Clave primaria de la tabla	Entero
DIM_DATE	Clave foránea de la dimensión DATE	Entero
DIM_COMPANIA	Clave foránea de la dimensión COMPANIA	Cadena

DIM_SERVICIO	Clave foránea de la dimensión SERVICIO	Cadena
DIM_ASOCIADO	Clave foránea de la dimensión ASOCIADO	Entero
NUMERO_0900	Este atributo representa el numero asignado al asociado que recibe la llamada	Entero
CLIENTE_NUMERO	Este atributo registra el número del cliente que realizo la llamada	Entero
PRECIO	En este atributo se registra el precio por minuto de la llamada	Real
DURACION_LLAMADA	Este atributo registra la duración de la llamada redondeada a minutos es entero	Real
MINUTOS_CONSUMIDOS	Este atributo representa la duración real de la llamada	Real

Tabla 3.11: Tabla de hechos 0900

FACT_INFOLINE: La tabla de hecho descrita en la Tabla 3.9 se encarga de registrar las transacciones asociadas a los contratos de servicios infoline asumidos por los clientes.

Nombre del Atributo	Contenido	
	Descripción	Tipo
FACT_INFOLINE_ID	Clave primaria de la tabla	Entero
DIM_DATE	Clave foránea de la dimensión DATE	Cadena
DIM_COMPANIA	Clave foránea de la dimensión COMPANIA	Cadena
DIM_SERVICIO	Clave foránea de la dimensión SERVICIO	Cadena
DIM_ASOCIADO	Clave foránea de la dimensión ASOCIADO	Entero
INFOLINE_CONTRATO	Número del contrato de infoline entero	Entero
PRECIO	Precio del contrato	Real

Tabla 3.12: Tabla de hechos Infoline

FACT_CALL_CENTER: La tabla de hecho descrita en la Tabla 3.9 se encarga de registrar las transacciones asociadas a las llamadas recibidas por los diferentes centros de contactos ofrecidos por la organización.

Nombre del Atributo	Contenido	
	Descripción	Tipo
FACT_CALL_CENTER_ID	Clave primaria de la tabla	Entero
DIM_DATE	Clave foránea de la dimensión DATE	Cadena
DIM_SERVICIO	Clave foránea de la dimensión SERVICIO	Cadena
DIM_ASOCIADO	Clave foránea de la dimensión ASOCIADO	Entero
CALL_NUMBER_SERVICE	Número del call center al cual llaman los clientes	Entero
ESTATUS	Estatus de la llamada recibida	Cadena
DURCION_LLAMADA	Registra la duración de la llamada	Real

Tabla 3.13: Tabla de hechos Call Center

FACT_PUBLICIDAD: La tabla de hecho descrita en la Tabla 3.9 se encarga de registrar las transacciones asociadas a los montos de inversiones realizadas en las campañas de los asociados

Nombre del Atributo	Contenido	
	Descripción	Tipo
FACT_PUBLICIDAD_ID	Clave primaria de la tabla	Entero
DIM_DATE	Clave foránea de la dimensión DATE	Cadena
DIM_COMPANIA	Clave foránea de la dimensión COMPANIA	Cadena
DIM_SERVICIO	Clave foránea de la dimensión SERVICIO	Cadena
DIM_ASOCIADO	Clave foránea de la dimensión ASOCIADO	Entero
PUBLICIDAD_MONTO	Monto invertido en publicidad	Real

PUBLICIDAD_PORCENTAJE_COMPANI	Porcentaje invertido por la empresa WTFE	Real
PUBLICIDAD_PORCENTAJE_ASOCIADO	Porcentaje invertido por el asociado	Real

Tabla 3.14: Tabla de hechos Publicidad

3.4 Fase III: Implementación

3.4.1 Desarrollo de la aplicación BI

Se configuró la solución BI instalando la Herramienta Pentaho Data Integrator en el servidor de desarrollo, así como el restante paquete de software seleccionado. De acuerdo con esto, se personalizaron solo aquellas características que no estuvieran predefinidas en la configuración empresarial por defecto.

Este proceso duró aproximadamente 3 días, y contó con la asistencia de personal técnico de la organización especialmente asignado para tal fin.

Con respecto a la configuración general de la Solución de Inteligencia de Negocios implementada, es importante señalar que el PDI fue configurado con 1 GB de memoria RAM, mientras que al Pentaho BI Server le fueron asignados 2 GB. Asimismo, en cuanto ubicación y estructura se optó por utilizar el Estándar Jerárquico del Sistema de Archivos (Filesystem Hierarchy Standard, FHS) que se presenta a continuación en la Tabla 3.15

Rutas	Descripción
PentahoBI/	Jerarquía Primaria, Raíz o root de la solución instalada
PentahoBI/bin	Contenedor de aplicaciones
PentahoBI/bin/scripts	Colección de instrucciones en lenguaje bash que facilitan el uso de herramientas y despliegue de la solución instalada
PentahoBI/bin/pdi	Contiene al PDI
PentahoBI/bin/psw	Contiene al PSW
PentahoBI/bin/pbi-server	Contiene al Pentaho BI Server
PentahoBI/etc	Guarda las configuraciones de la solución instalada
PentahoBI/lib	Contenedor los conectores JDBC necesarios para la adecuada conexión con las fuentes de datos así como el repositorio de los ETLs
PentahoBI/tmp	Contenedor temporal de archivos
PentahoBI/var	Contenedor de archivos dinámicos necesarios durante la ejecución del sistema
PentahoBI/var/lock	Contenedor de archivos de seguimiento de recursos en uso
PentahoBI/var/logs	Contenedor de archivos de registro y auditoría del sistema
PentahoBI/var/run	Contenedor del identificador del proceso que se está ejecutando
PentahoBI/tmp	Contenedor de archivos temporales o de registro que deben perdurar durante el reinicio del Sistema de Inteligencia de Negocio

Tabla 3.15: Estructura de archivos del Sistema de Inteligencia de Negocios

3.5 Diseño físico

Teniendo en cuenta lo estipulado en la sección 2.11.3, y atendiendo específicamente el apartado **Diseño Físico**, se presentan a continuación los cinco esquemas estrellas generados a partir de la aplicación de la metodología que forman parte de la conformación del estándar organizacional de la solución implementada, y que van desde la Figura 3.2 y hasta la Figura 3.6

3.5.1 SMS Generados

Modelo estrella simplificado del cubo SMS:

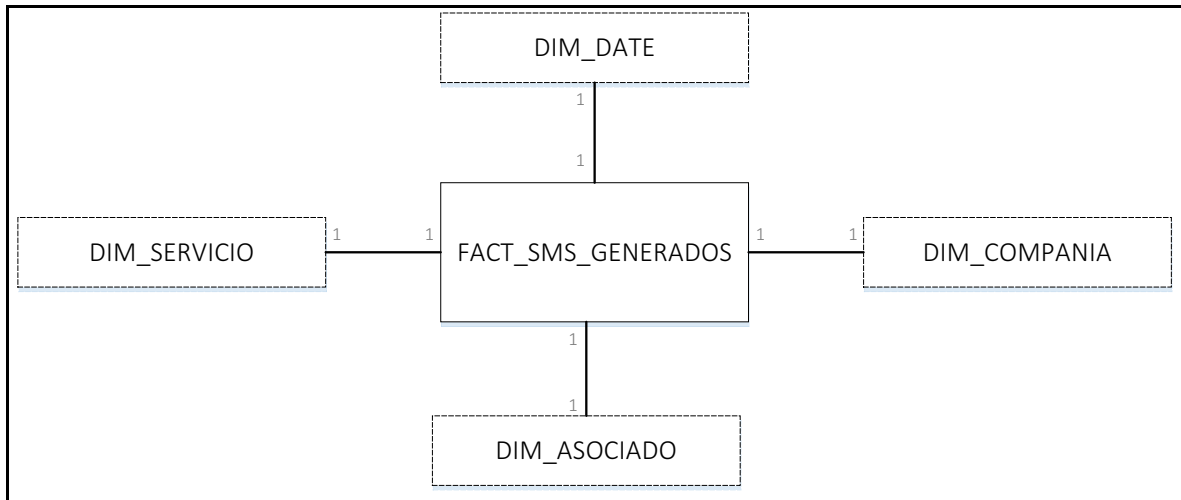


Figura 3.2: SMS Generados

3.5.2 Servicios 0900

Modelo estrella simplificado del cubo 0900:

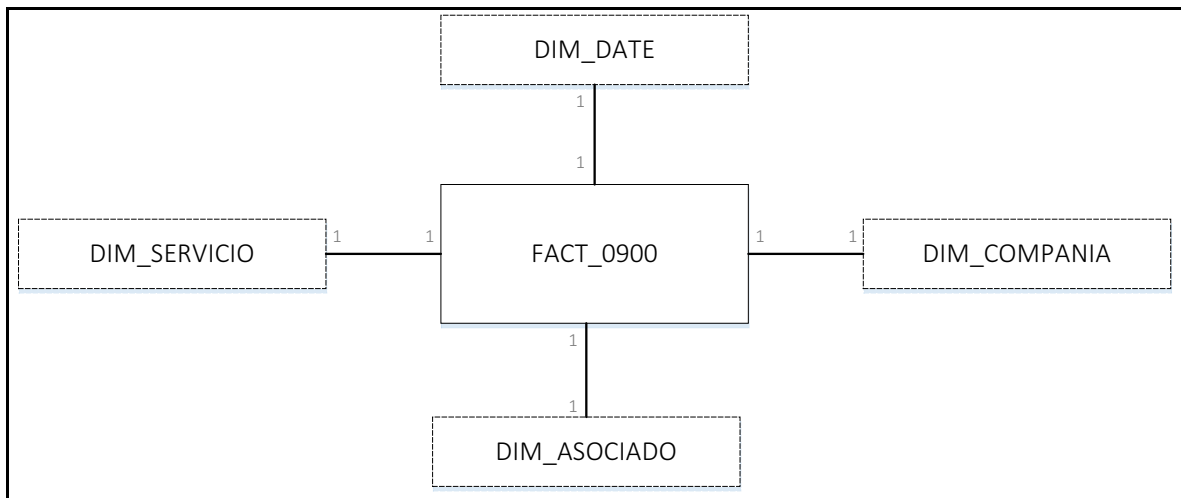


Figura 3.3: Servicios 0900

3.5.3 Infoline

Modelo estrella simplificado del cubo Infoline:

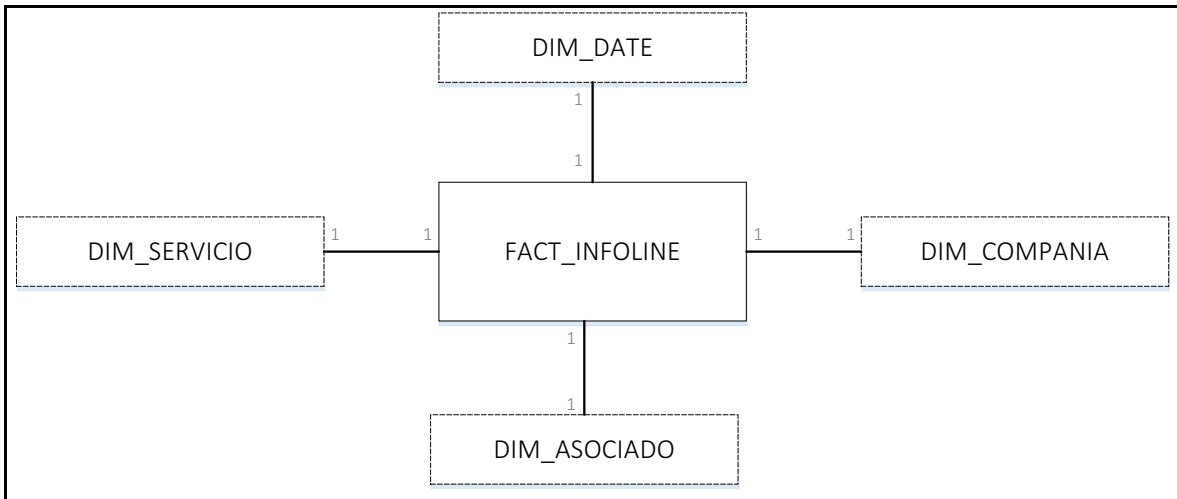


Figura 3.4: Infoline

3.5.4 Call center

Modelo estrella simplificado del cubo Call Center:

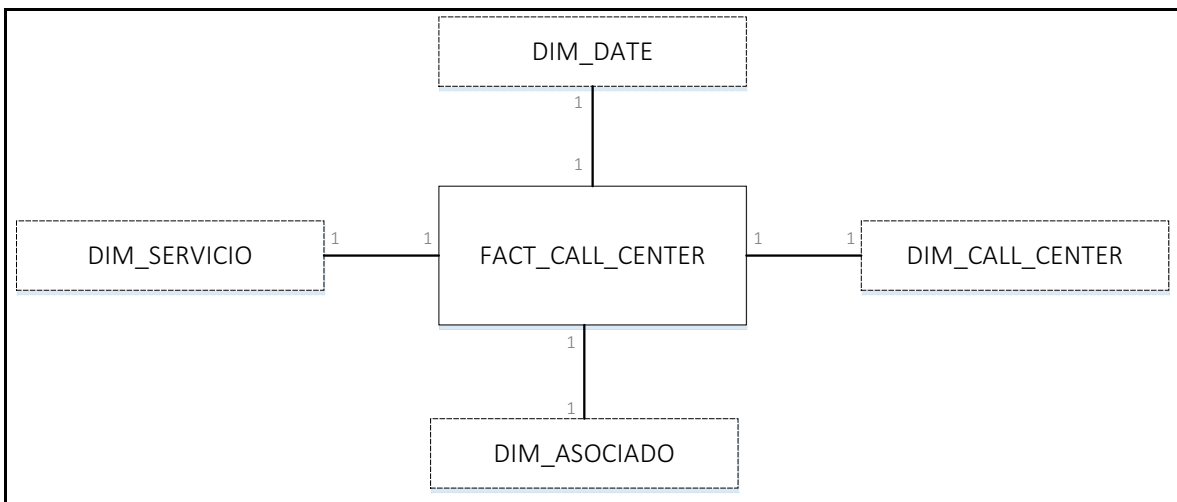


Figura 3.5: Call Center

3.5.5 Publicidad

Modelo estrella simplificado del cubo de publicidad:

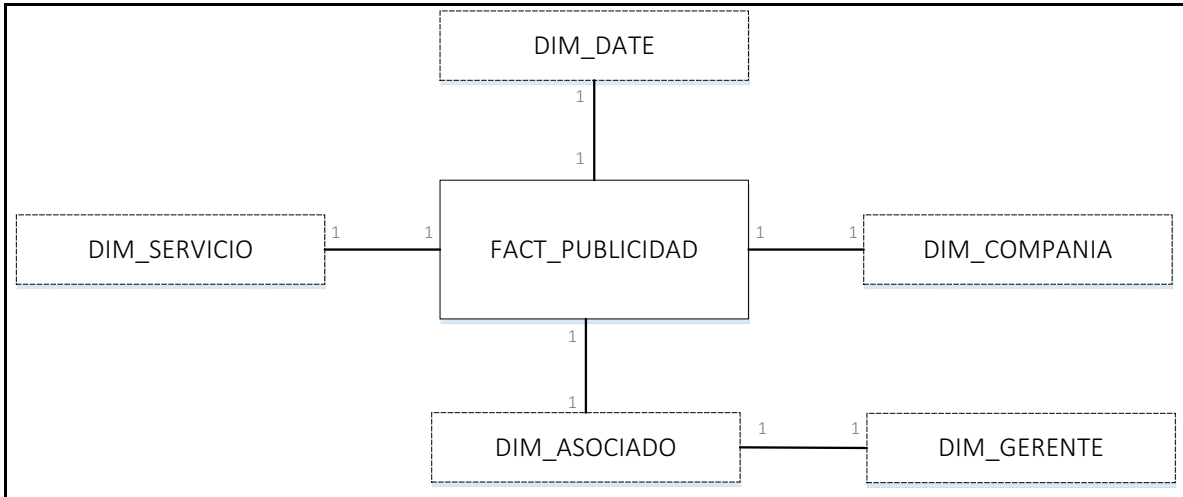


Figura 3.6: Publicidad

Para el desarrollo de los cubos OLAP se utilizó la herramienta de administración de OBIEE, dado a que facilita tanto la creación de los cubos OLAP en un entorno gráfico como el despliegue de dichos cubos en el servidor de Inteligencia de Negocios según se muestra en la Figura 3.7

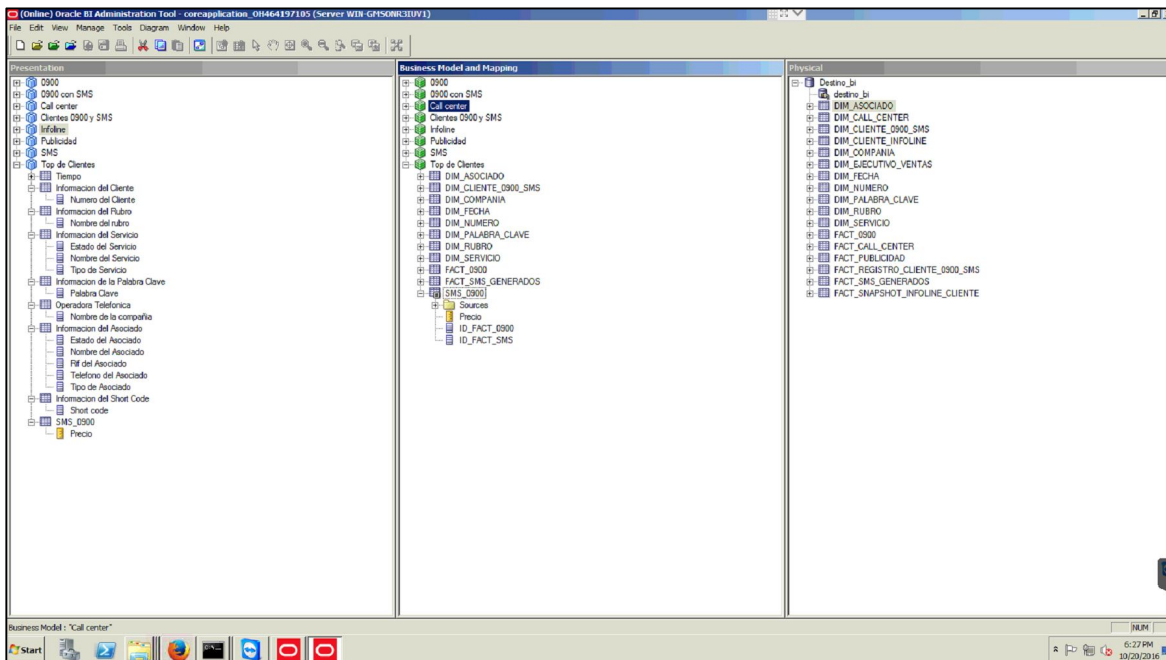


Figura 3.7: Oracle BI Administrator tool

3.6 Diseño y desarrollo de los procesos de ETL

Teniendo el almacén de datos listo para su llenado, debimos enfocarnos en la migración de los datos desde las distintas fuentes hacia el almacén de datos, para ello se diseñaron y construyeron los ETL, bajo las siguientes condiciones:

- Se transformaron los datos obtenidos de las fuentes de cada área de negocio en un solo proceso.
- Una vez transformados los datos se guardaron en un área temporal para ser cargados al almacén de los datos.
- Se cargaron consecutivamente los datos para cada una de las áreas de negocio procesadas.
- Se identificaron debidamente los datos guardados en el área temporal, con el número de la corrida del proceso de ETL y se archivaron para su posterior uso.

3.6.1 Origen y destino de los datos

La organización estudiada, almacena la totalidad de sus datos en múltiples bases de datos administradas por varias instancias de PostgreSQL, que además están distribuidas a lo largo de múltiples servidores.

3.6.2 Procesos de extracción transformación y carga

Para el desarrollo de este Sistema de Inteligencia de Negocios, se crearon 32 procesos de ETL, que pueden ser agrupados según sus funciones: Orquestador, Carga, Extractor + Transformador, Generador, Pruebas y Control. El resumen de los ETL desarrollados se encuentra en la Tabla 3.16

	Proceso	Tipo	Comentario
1	ACTUALIZAR-ALMACEN-DIARIO.KJB	Orquestador	
2	PARTE1-EXTRACION-TRANSFORMACION-DIARIA.KJB	Orquestador	
3	DIM_ASOCIADO.EXTRACCION.KTR	Extractor + Transformador	
4	DIM_CLIENTE_0900.EXTRACCION.KTR	Extractor + Transformador	
5	DIM_COMPANIA.EXTRACCION.KTR	Extractor + Transformador	
6	DIM_NUMERO.EXTRACCION.KTR	Extractor + Transformador	
7	DIM_PALABRA_CLAVE.EXTRACCION.KTR	Extractor + Transformador	
8	DIM_RUBRO.EXTRACCION.KTR	Extractor + Transformador	
9	DIM_SERVICIO.EXTRACCION.KTR	Extractor + Transformador	
10	DIM_CALL_CENTER.EXTRACCION.KTR	Extractor + Transformador	
11	DIM_ASOCIADO.CARGA.KTR	Carga	
12	DIM_CLIENTE_0900.CARGA.KTR	Carga	
13	DIM_COMPANIA.CARGA.KTR	Carga	
14	DIM_NUMERO.CARGA.KTR	Carga	
15	DIM_PALABRA_CLAVE.CARGA.KTR	Carga	
16	DIM_RUBRO.CARGA.KTR	Carga	
17	DIM_SERVICIO.CARGA.KTR	Carga	
18	DIM_CALL_CENTER.CARGA.KTR	Carga	

	Proceso	Tipo	Comentario
19	PARTE2-CARGA-DIARIA.KJB	Orquestador	
20	FACT_0900.EXTRACCION.KTR	Extractor + Transformador	
21	FACT_SMS_GENERADOS.EXTRACCION.KTR	Extractor + Transformador	
22	FACT_CALL_CENTER.EXTRACCION.KTR	Extractor + Transformador	
23	FACT_0900.CARGA.KTR	Carga	
24	FACT_SMS_GENERADOS.CARGA.KTR	Carga	
25	FACT_CALL_CENTER.CARGA.KTR	Carga	
26	FACT_REGISTRO_CLIENTES_0900_SMS.KBJ	Orquestador	
27	FACT_RC900SMS-CALCULAR_DIAS_A_PROCESAR.KTR	Carga	
28	FACT_RC900SMS-GENERAR_REGISTROS_DIA_A_PROCESAR	Carga	
29	ACTUALIZAR-ALMACEN-MENSUAL.KJB	Orquestador	
30	PARTE1-EXTRACION-TRANSFORMACION-MENSUAL.KJB	Orquestador	
31	DIM_EJECUTIVO_VENTAS.EXTRACCION.KTR	Extractor + Transformador	
32	FACT_INFOLINE.EXTRACCION.KTR	Extractor + Transformador	
33	DIM_CLIENTE_INFOLINE.CARGA.KTR	Carga	
34	DIM_EJECUTIVO_VENTAS.CARGA.KTR	Carga	
35	FACT_INFOLINE.EXTRACCION.KTR	Extractor + Transformador	
36	DIM_CLIENTE_INFOLINE.CARGA.KTR	Carga	
37	DIM_EJECUTIVO_VENTAS.CARGA.KTR	Carga	
38	PARTE2-CARGA-MENSUAL.KJB	Orquestador	
39	CT_PUBLICIDAD.EXTRACCION.KTR	Extractor + Transformador	
40	FACT_SNAPSHOT_INFOLINE.CARGA.KTR	Carga	
41	DIM_EJECUTIVO_VENTAS.CARGA.KTR	Carga	

Tabla 3.16: Lista de los procesos de extracción, transformación y carga

Los procesos asociados en la Tabla 3.16 y que fueron desarrollados en Pentaho Data Integration, se describen a continuación utilizando el formato de presentación descrito en la Tabla 3.17

Proceso: << NOMBRE DEL PROCESO A DESCRIBIR >>	
Objetivo:	Origen de los datos:
<< Propósito por el cual el proceso fue creado >>	<< Descripción o lista del origen de donde el proceso obtiene los datos a procesar >>
	Destino de los datos
	<< Descripción o lista del destino donde el proceso guarda los datos a procesar >>

Tabla 3.17: Formato de presentación de los procesos ETL

Proceso : ACTUALIZAR-ALMACEN-DIARIO.KJB	
Objetivo:	Origen de los datos:
Punto de entrada para los procesos de extracción transformación y carga referente al manejo de las transacciones diarias. Crea el archivo de control de la ejecución actual. Ejecuta los procesos: parte1-extracion-transformacion-diaria.kjb y parte2-carga-diaria.kjb.	Ninguno
	Destino de los datos:
	Ninguno

Tabla 3.18: Proceso Actualizar Almacén Diario

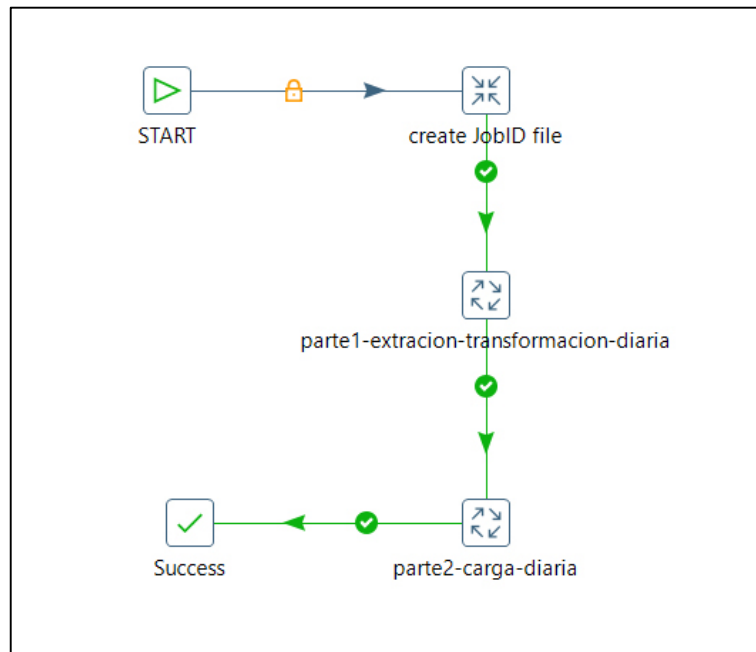


Figura 3.8: Proceso Actualizar Almacén diario

Proceso: PARTE1-EXTRACION-TRANSFORMACION-DIARIA.KJB	
Objetivo:	Origen de los datos:
Punto de entrada para los procesos de extracción transformación y carga referente a las dimensiones involucradas en el manejo de las transacciones diarias.	Ninguno
	Destino de los datos:
	Ninguno

Tabla 3.19: Proceso Parte 1 Extracción Transformación Diaria

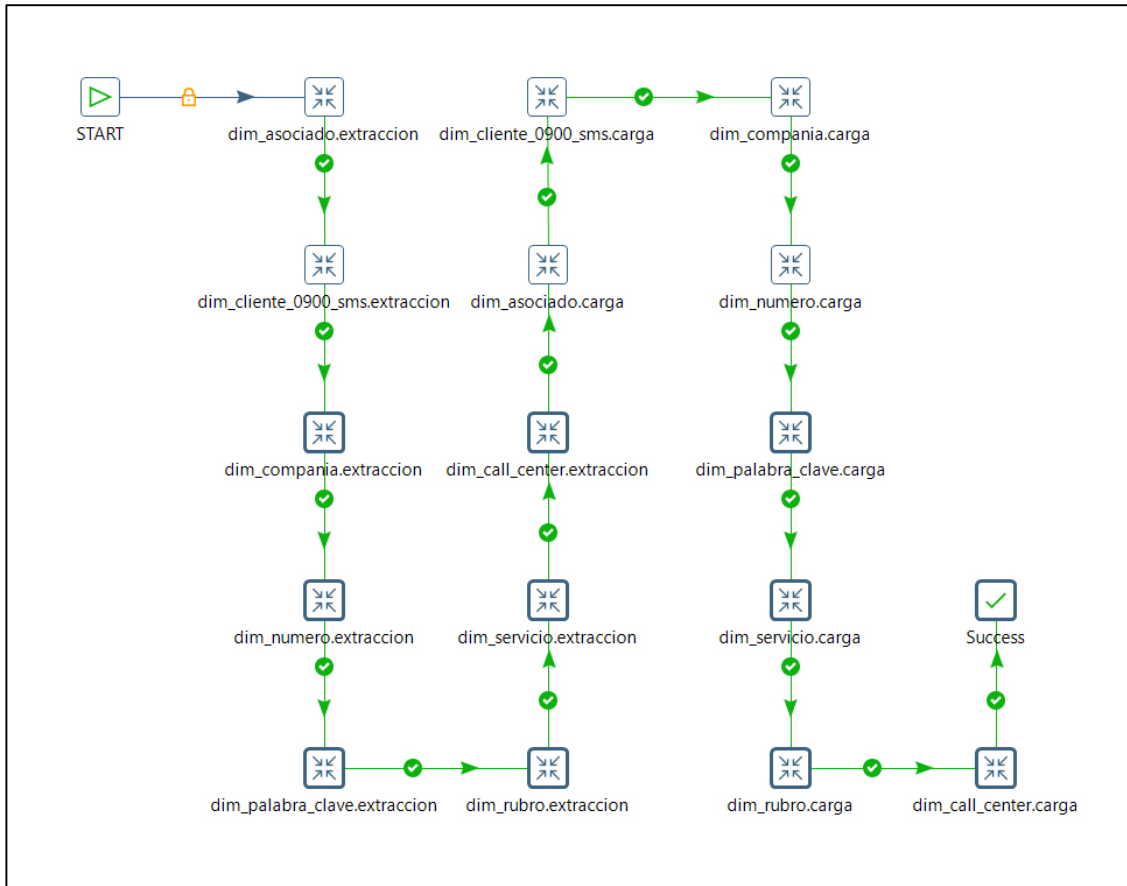


Figura 3.9: Proceso Parte 1 Extracción Transformación Diaria

Proceso: DIM_ASOCIADO.EXTRACCION.KTR	
Objetivo:	Origen de los datos:
Extrae y almacena en archivo la lista de clientes asociados desde el origen de su creación o los modificados desde la última ejecución del proceso	db_facturacion.ad_asociado
	Destino de los datos:
	dim_asociado.extraccion.out

Tabla 3.20: Proceso Dimensión Asociado extracción

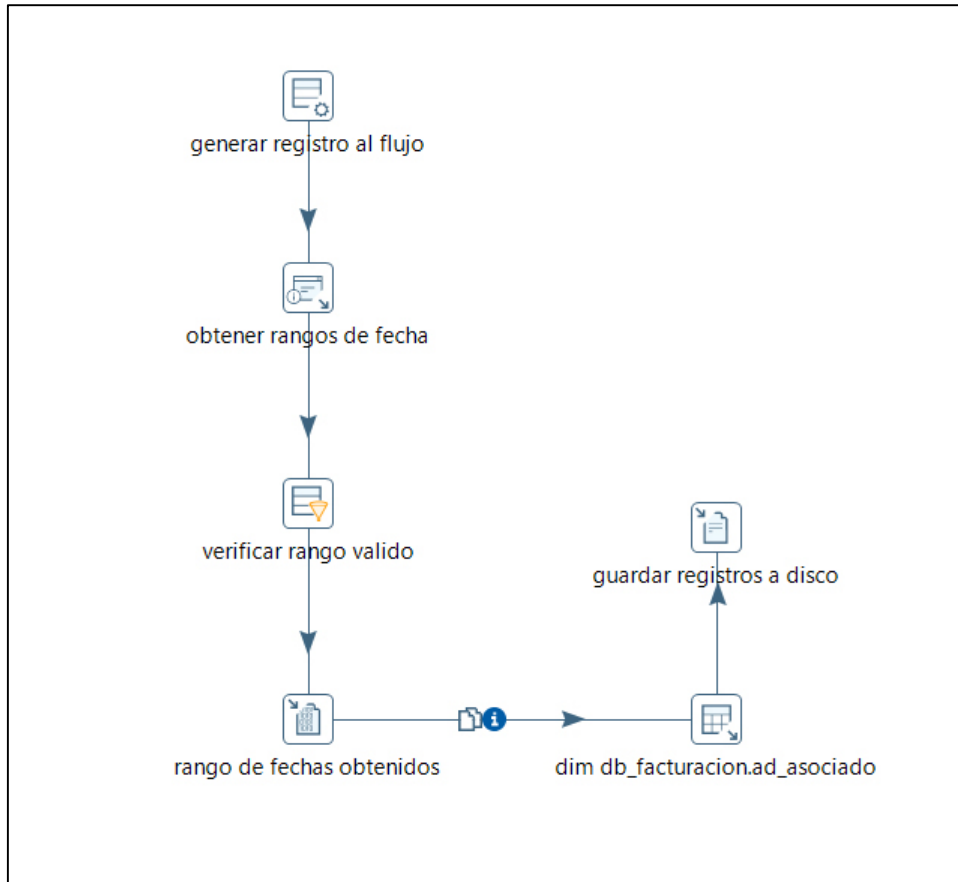


Figura 3.10: Proceso Dimensión Asociado extracción

Proceso: DIM_CLIENTE_0900.EXTRACCION.KTR	
Objetivo: Extrae y almacena en archivo una lista con los ANI desde el origen de su creación o los modificados desde la última ejecución del proceso	Origen de los datos: public.registros_facturables_sms_digitel public.registros_facturables_sms_movilnet public.registros_facturables_sms_telcel ivr.registros_facturables_900_cantv ivr.registros_facturables_900_cantv_prepago ivr.registros_facturables_900_movistar ivr.registros_facturables_900_movilnet ivr.registros_facturables_900_digitel
	Destino de los datos: dim_cliente_0900_sms.extraccion.out

Tabla 3.21: Proceso Dimensión Cliente 0900 extracción

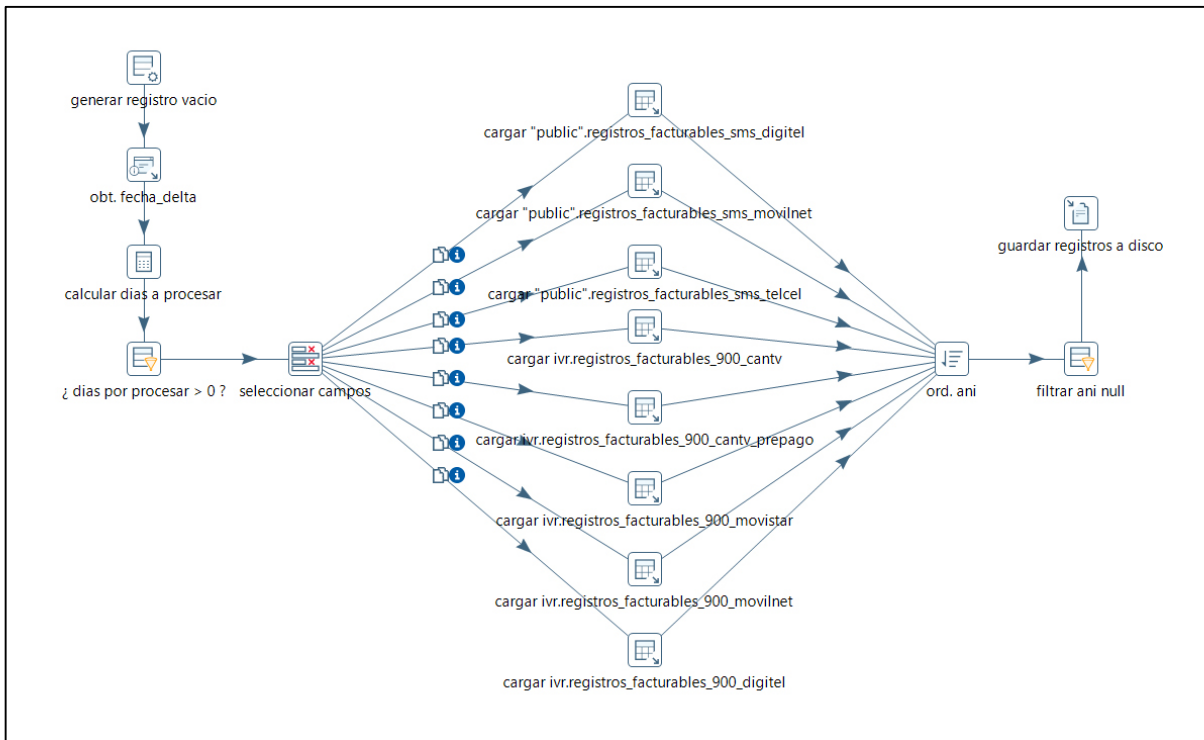


Figura 3.11: Proceso Dimensión Cliente 0900 extracción

Proceso: DIM_COMPANIA.EXTRACCION.KTR	
Objetivo:	Origen de los datos:
Extrae y almacena en archivo una lista con los datos de las Operadoras desde el origen de su creación o los modificados desde la última ejecución del proceso	db_facturacion.ad_operadora
	Destino de los datos:
	dim_compania.extraccion.out

Tabla 3.22: Proceso Dimensión Compañía extracción

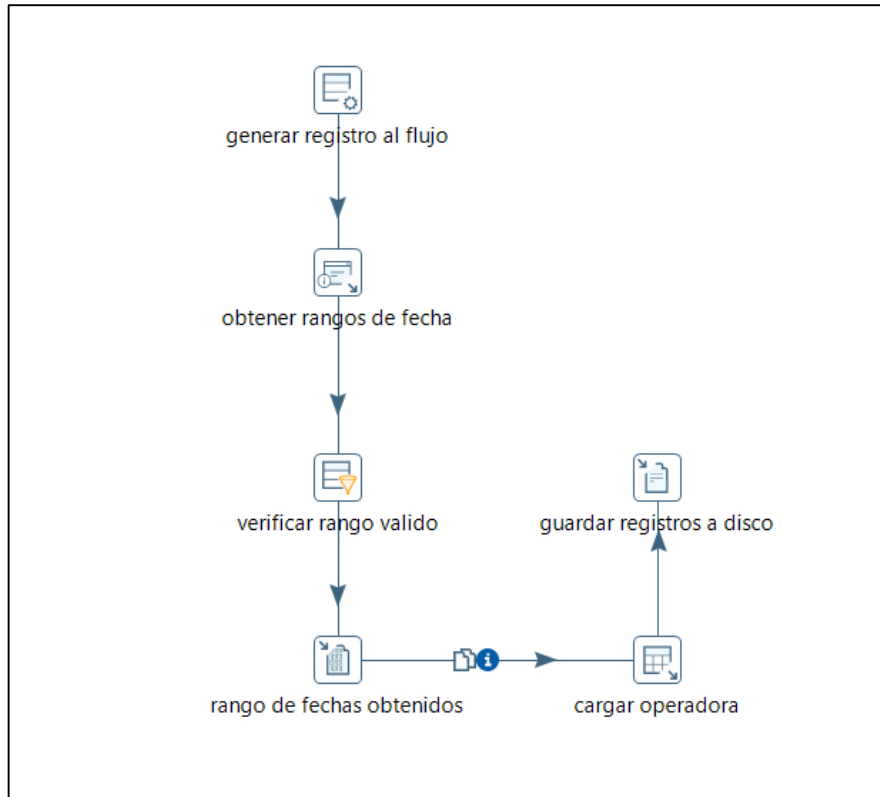


Figura 3.12: Proceso Dimensión Compañía extracción

Proceso: DIM_NUMERO.EXTRACCION.KTR	
Objetivo:	Origen de los datos:
Extrae y almacena en archivo una lista las tarifas asociadas a cada número telefónica desde el origen de su creación o los modificados desde la última ejecución del proceso	db_facturacion.ad_numero_tarifa
	Destino de los datos:
	dim_numero.extraccion.out

Tabla 3.23: Proceso Dimensión Numero

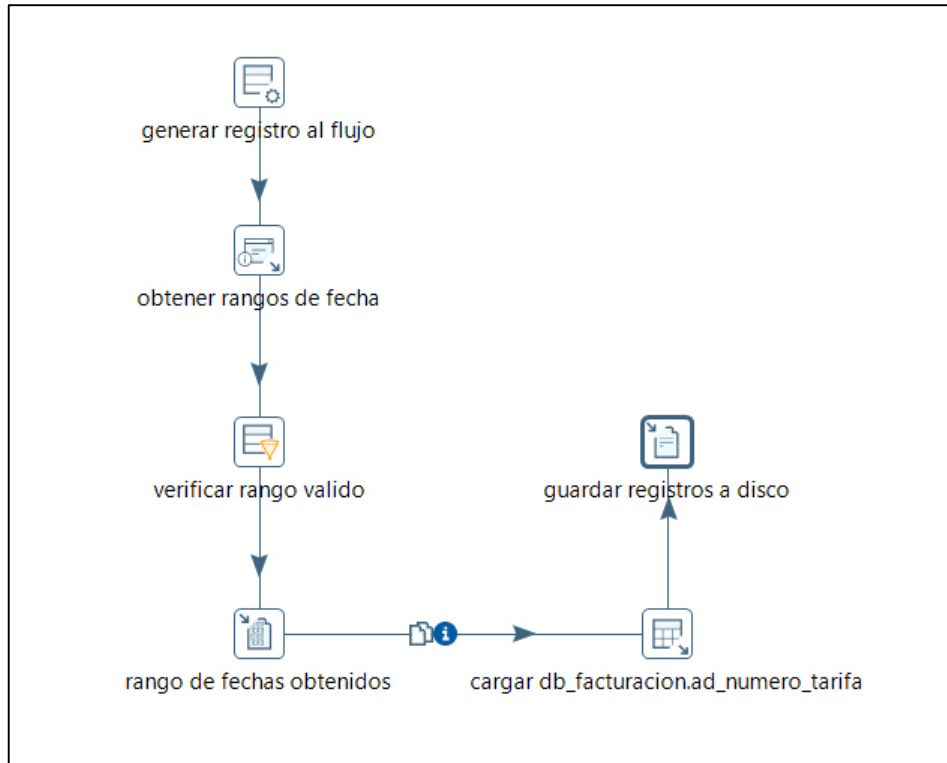


Figura 3.13: Proceso Dimensión Numero

Proceso: DIM_PALABRA_CLAVE.EXTRACCION.KTR	
Objetivo:	Origen de los datos:
Extrae y almacena en archivo una lista de las palabras claves asociadas al servicio SMS desde el origen de su creación o los modificados desde la última ejecución del proceso	public.palabraclave
	Destino de los datos:
	dim_palabra_clave.extraccion.out

Tabla 3.24: Proceso Dimensión Palabra Clave

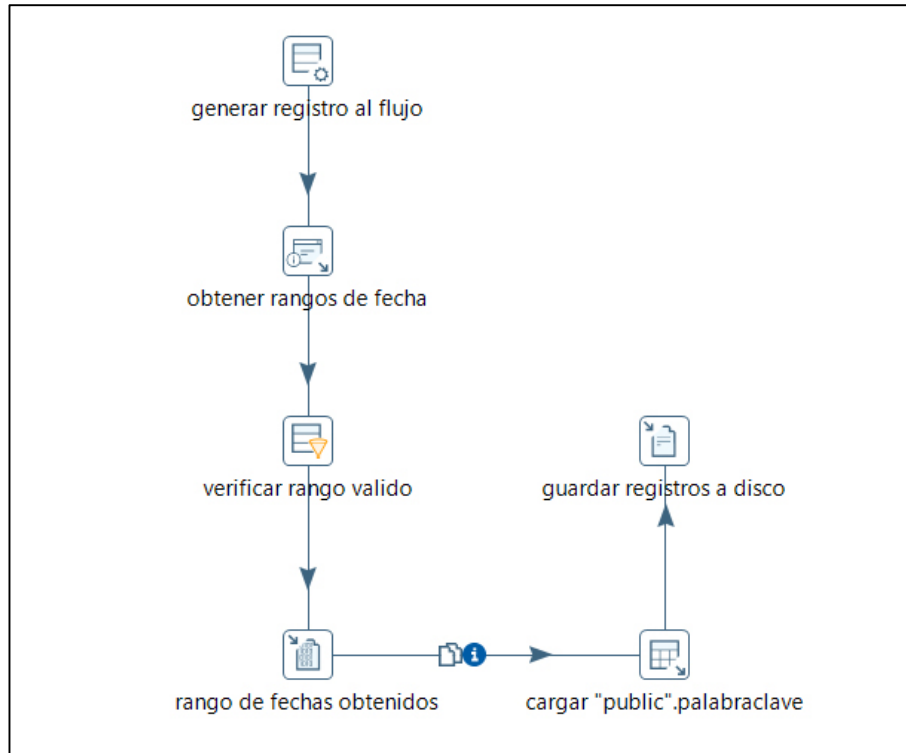


Figura 3.14: Proceso Dimensión Palabra Clave

Proceso: DIM_RUBRO.EXTRACCION.KTR	
Objetivo:	Origen de los datos:
Extrae, trasforma y almacena los rubros existentes desde el origen de su creación o los modificados desde la última ejecución del proceso	db_facturacion.ad_catalogo_rubro public.catalogo_rubro
	Destino de los datos
	dim_rubro.extraccion.out

Tabla 3.25: Proceso Dimensión Rubro

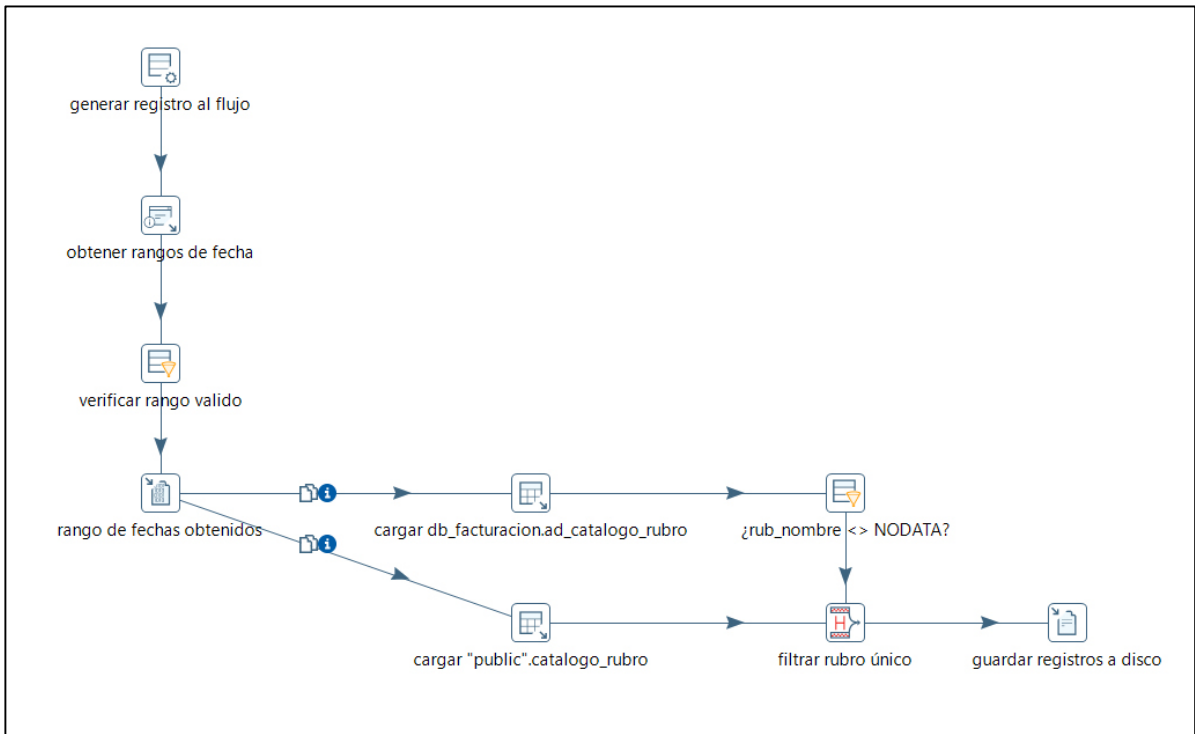


Figura 3.15: Proceso Dimensión Rubro

Proceso: DIM_SERVICIO.EXTRACCION.KTR	
Objetivo:	Origen de los datos:
Extrae, trasforma y almacena en archivo los servicios ofrecidos existentes desde el origen de su creación o los modificados desde la última ejecución del proceso	db_facturacion.ad_servicio public.palabraclave
	Destino de los datos
	dim_servicio.extraccion.out

Tabla 3.26: Proceso Dimensión Servicio

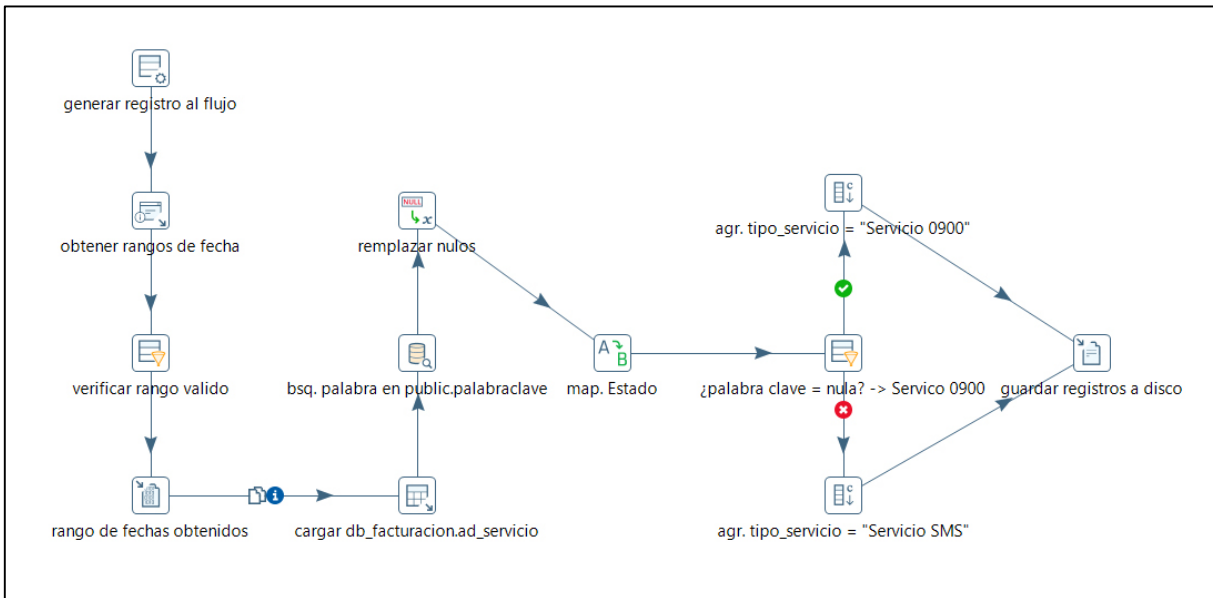


Figura 3.16: Proceso Dimensión Servicio

Proceso: DIM_CALL_CENTER.EXTRACCION.KTR	
Objetivo:	Origen de los datos:
Extrae, trasforma y almacena en archivo los Call Centers existentes desde el origen de su creación o los modificados desde la última ejecución del proceso	db_seguridad.ad_call_center
	Destino de los datos:
	dim_call_center.extraccion.out

Tabla 3.27: Proceso Dimensión Call Center

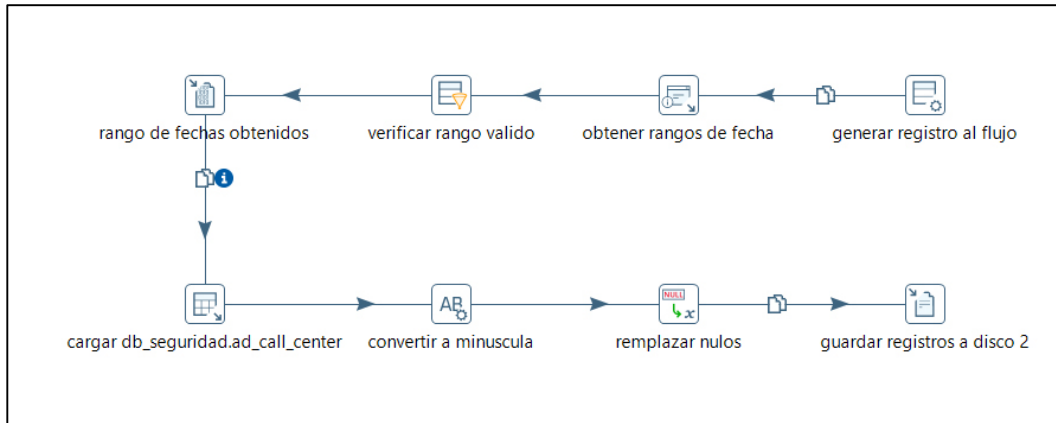


Figura 3.17: Proceso Dimensión Call Center

Proceso: DIM_ASOCIADO.CARGA.KTR	
Objetivo:	Origen de los datos:
Trasforma y carga desde archivo la lista de clientes asociados	dim_asociado.extraccion.out
	Destino de los datos:
	destino_bi.dim_asociado

Tabla 3.28: Dim Asociado Carga

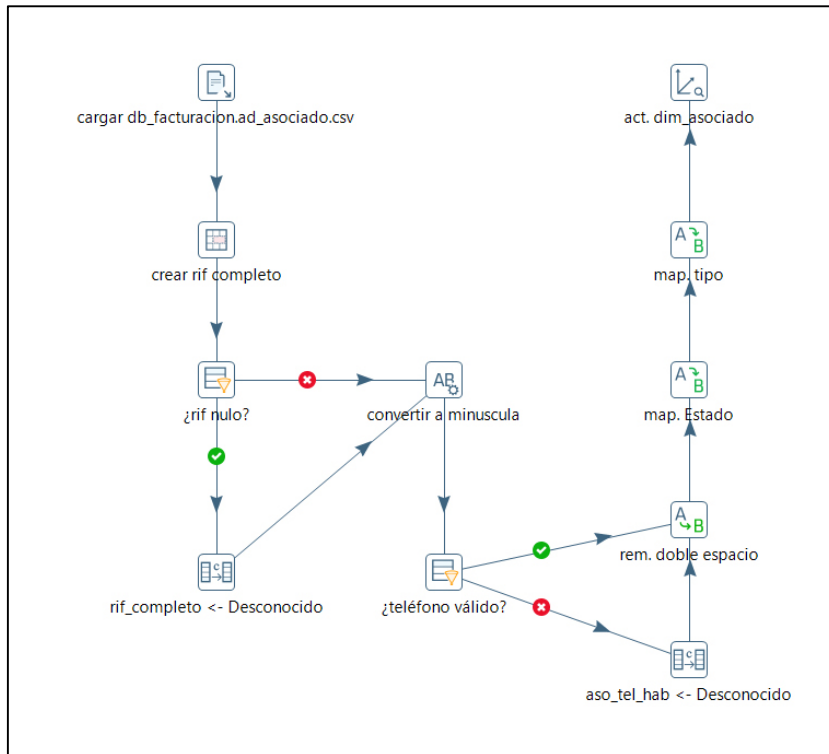


Figura 3.18: Dim Asociado Carga

Proceso: DIM_CLIENTE_0900.CARGA.KTR	
Objetivo:	Origen de los datos:
Carga desde archivo la lista con los ANI obtenidos anteriormente	dim_cliente_0900_sms.extraccion.out
	Destino de los datos:
	destino_bi.dim_cliente_0900_sms

Tabla 3.29: Dim Cliente 0900 Carga

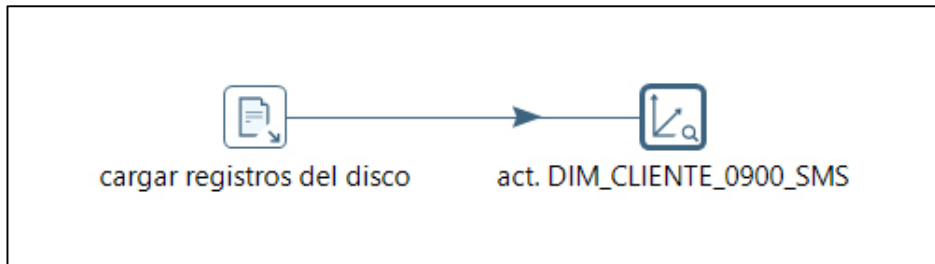


Figura 3.19: Dim Cliente 0900 Carga

Proceso: DIM_COMPANIA.CARGA.KTR	
Objetivo:	Origen de los datos:
Carga desde archivo la lista con las Operadoras obtenidas anteriormente	dim_compania.extraccion.out
	Destino de los datos:
	destino_bi.dim_compania

Tabla 3.30: Dim compañía carga

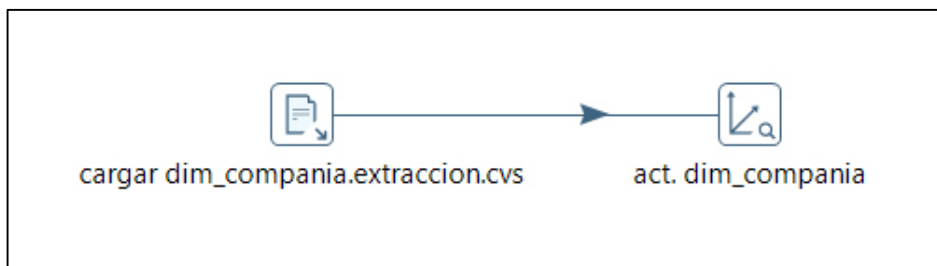


Figura 3.20: Dim compañía carga

Proceso: DIM_NUMERO.CARGA.KTR	
Objetivo:	Origen de los datos:
Carga desde archivo la lista con las tarifas y números asociados obtenidos anteriormente	dim_numero.extraccion.out
	Destino de los datos:
	destino_bi.dim_numero

Tabla 3.31: Dim Numero Carga

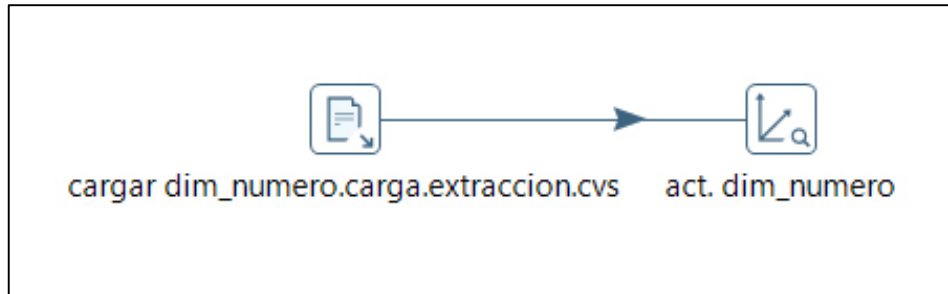


Figura 3.21: Dim Numero Carga

Proceso: DIM_PALABRA_CLAVE.CARGA.KTR	
Objetivo:	Origen de los datos:
Carga desde archivo la lista con las palabras claves para SMS obtenidos en procesos anteriores	dim_palabra_clave.extraccion.out
	Destino de los datos:
	destino_bi.dim_palabraclave

Tabla 3.32: Dim Palabra Clave Carga

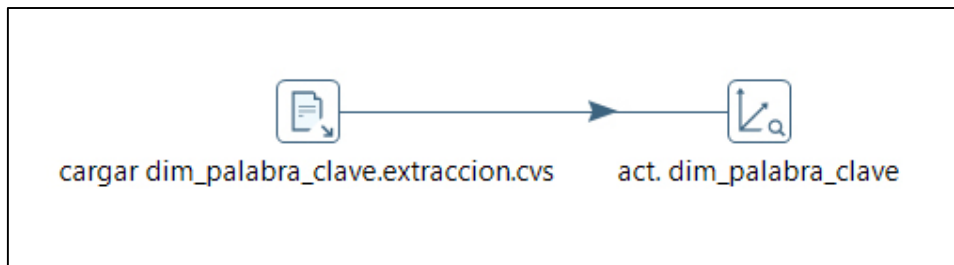


Figura 3.22: Dim Palabra Clave Carga

Proceso: DIM_RUBRO.CARGA.KTR	
Objetivo:	Origen de los datos:
Carga desde archivo la lista con los rubros obtenidos en procesos anteriores	dim_rubro.extraccion.out
	Destino de los datos:
	destino_bi.dim_rubro

Tabla 3.33: Dim rubro

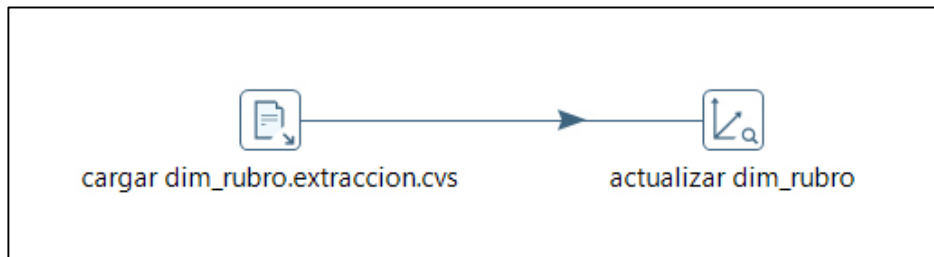


Figura 3.23: Dim rubro

Proceso: DIM_SERVICIO.CARGA.KTR	
Objetivo:	Origen de los datos:
Carga desde archivo la lista con los servicios obtenidos en procesos anteriores	dim_servicio.extraccion.out
	Destino de los datos:
	destino_bi.dim_servicio

Tabla 3.34: Dim Servicio Carga

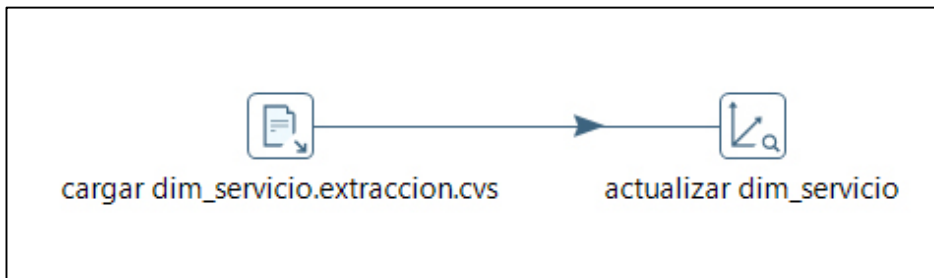


Figura 3.24: Dim Servicio Carga

Proceso: DIM_CALL_CENTER.CARGA.KTR	
Objetivo:	Origen de los datos:
Carga desde archivo la lista con los datos de los Call Center obtenidos en procesos anteriores	dim_call_center.extraccion.out
	Destino de los datos:
	destino_bi.dim_call_center

Tabla 3.35: Dim Call Center Carga

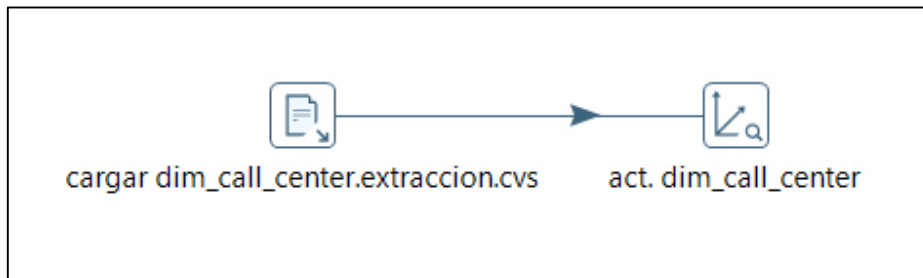


Figura 3.25: Dim Call Center Carga

Proceso: PARTE2-CARGA-DIARIA.KJB	
Objetivo:	Origen de los datos:
Punto de entrada para los procesos de extracción transformación y carga referente a las tablas de hechos (FACTS) involucradas en el manejo de las transacciones diarias.	Ninguno
	Destino de los datos:
	Ninguno

Tabla 3.36: Parte 2 Carga Diaria

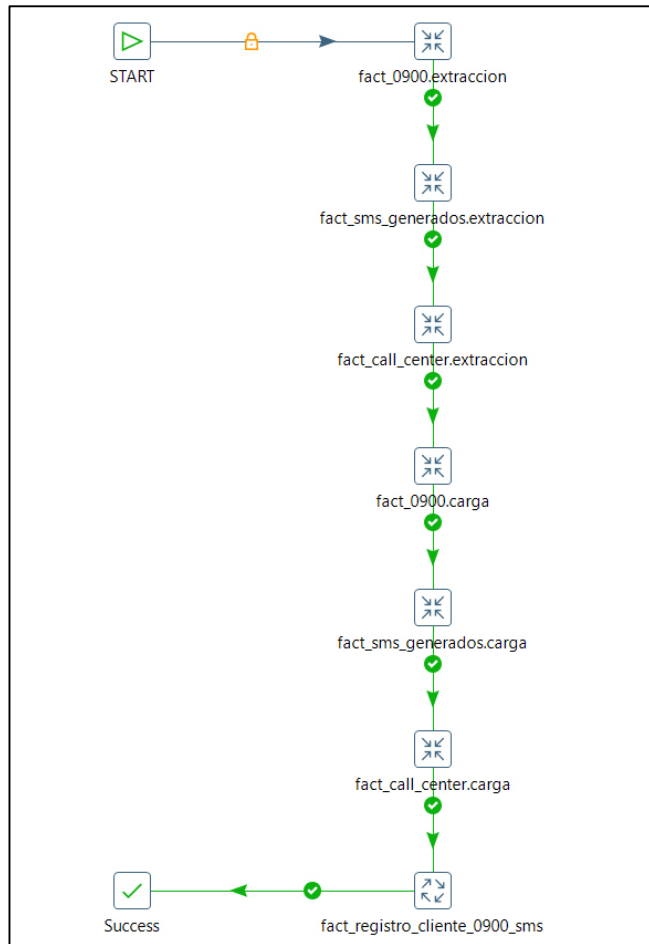


Figura 3.26: Parte 2 Carga Diaria

Proceso: FACT_0900.EXTRACCION.KTR

Objetivo:

Extrae, transforma y almacena en archivo, las transacciones diarias relacionadas con el servicio 0900, desde su creación de origen o los modificados desde la última ejecución del proceso.

Origen de los datos:

ivr.registros_facturables_900_cantv
 ivr.registros_facturables_900_cantv_prepago
 ivr.registros_facturables_900_movistar
 ivr.registros_facturables_900_movilnet
 ivr.registros_facturables_900_digitel

Destino de los datos

fact_0900.extraccion.out

Tabla 3.37: Proceso tabla de hechos 0900

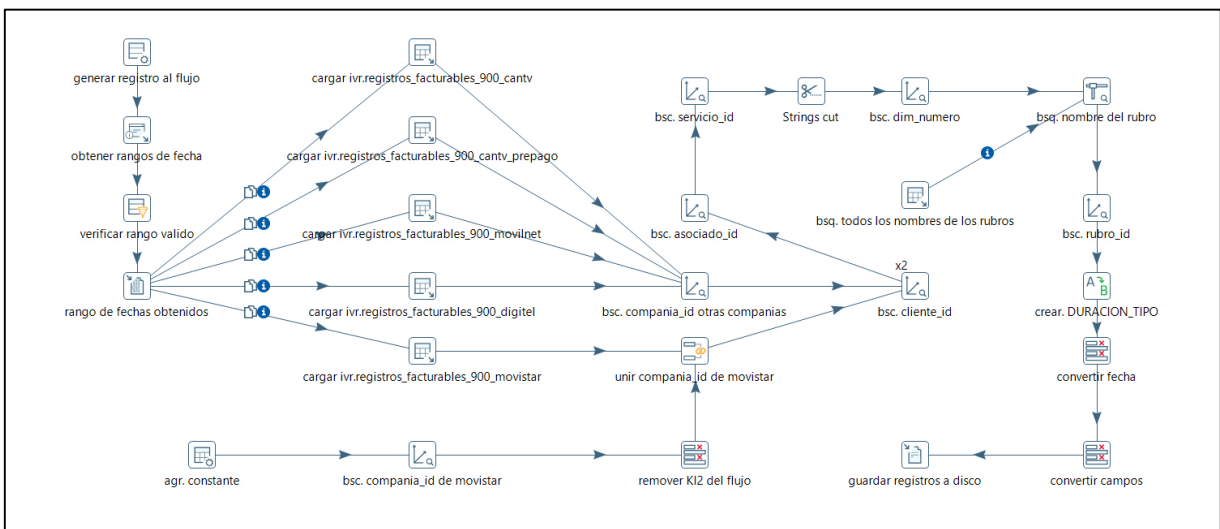


Figura 3.27: Proceso tabla de hechos 0900

Proceso: FACT_SMS_GENERADOS.EXTRACCION.KTR

Objetivo:

Extrae, transforma y almacena en archivo, las transacciones diarias relacionadas con el servicio de SMS generados, desde su creación de origen o los modificados desde la última ejecución del proceso.

Origen de los datos:

public.registros_facturables_sms_digital
public.registros_facturables_sms_movilnet
public.registros_facturables_sms_telcel

Destino de los datos

fact_sms_generados.extraccion.out

Tabla 3.38: Proceso tabla de hechos SMS Generados

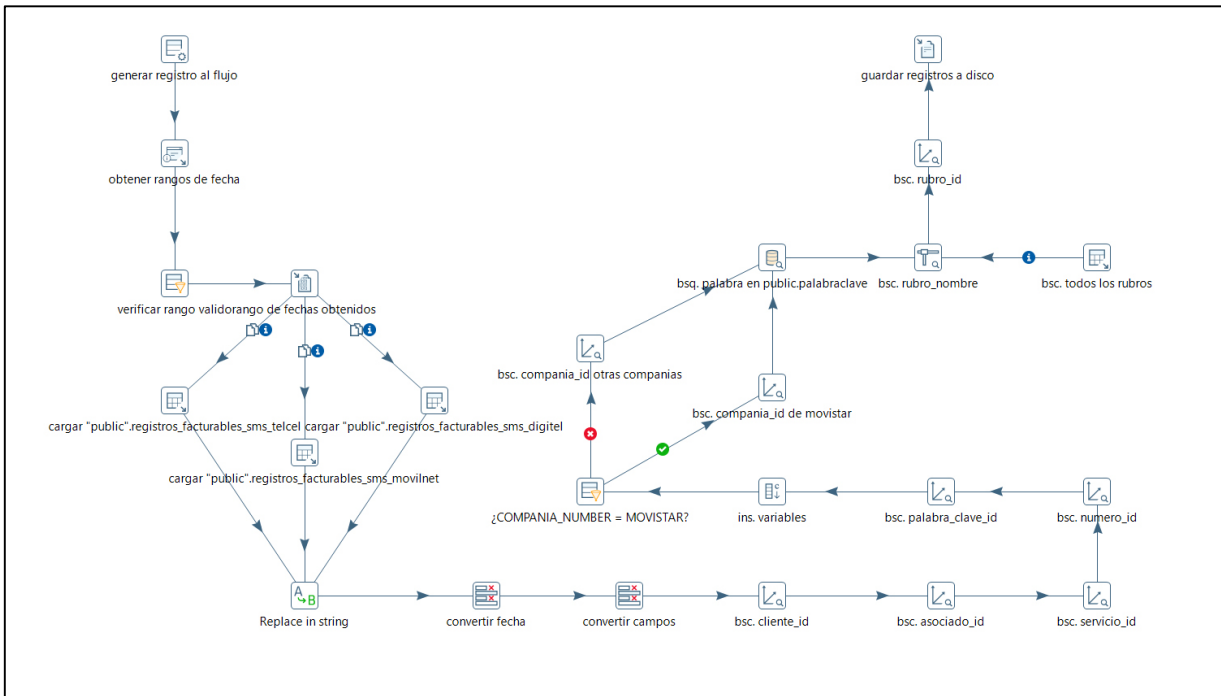


Figura 3.28: Proceso tabla de hechos SMS Generados

Proceso: FACT_CALL_CENTER.EXTRACCION.KTR

Objetivo:

Extrae, transforma y almacena en archivo, los registros de las llamadas relacionadas con los servicios de Call Center, desde su creación de origen o los modificados desde la última ejecución del proceso.

Origen de los datos:

public.cdr_sip6

Destino de los datos

fact_call_center.extraccion.out

Tabla 3.39: Proceso tabla de hechos Call Center

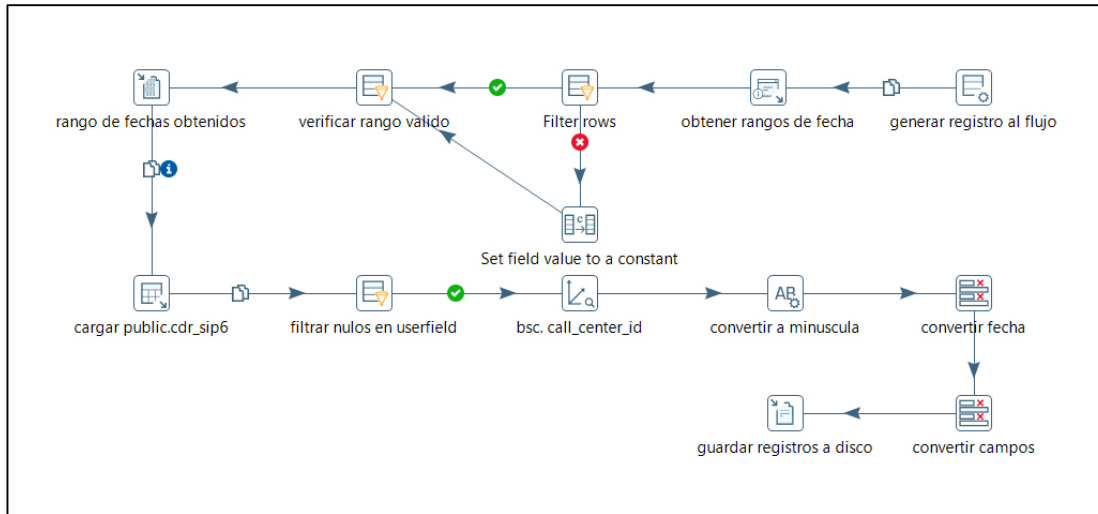


Figura 3.29: Proceso tabla de hechos Call Center

Proceso: FACT_0900.CARGA.KTR

Objetivo:

Extrae de archivo y carga en la base de datos destino Fact_0900 con la información de todas las transacciones y llamadas relacionadas al servicio 0900, desde su creación de origen o los modificados desde la última ejecución del proceso.

Origen de los datos:

fact_call_center.extraccion.out

Destino de los datos

destino_bi.fact_0900

Tabla 3.40: Fact 0900 Carga

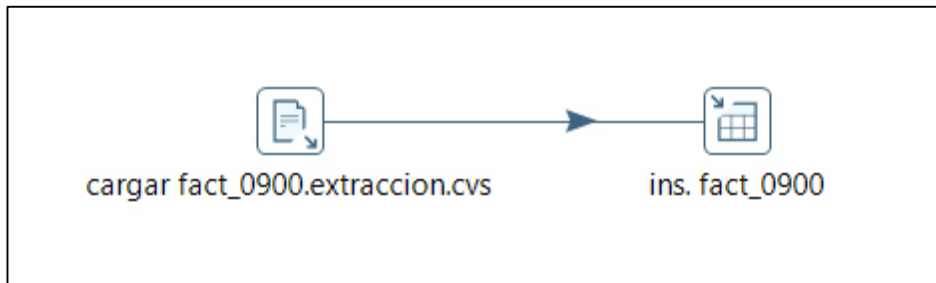


Figura 3.30: Fact 0900 Carga

Proceso: FACT_SMS_GENERADOS.CARGA.KTR

Objetivo:

Extrae de archivo y carga en la base de datos destino sms_generados con la información de todas las transacciones relacionadas al servicio SMS, desde su creación de origen o los modificados desde la última ejecución del proceso.

Origen de los datos:

fact_sms_generados.extraccion.out

Destino de los datos

destino_bi.fact_sms_generados

Tabla 3.41: Fact SMS Generados

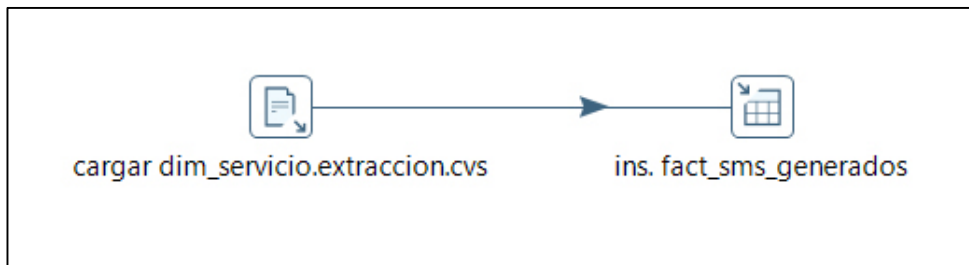


Figura 3.31: Fact SMS Generados

Proceso: FACT_CALL_CENTER.CARGA.KTR

Objetivo:

Extrae de archivo y carga en la base de datos destino call_center con la información de todas las transacciones relacionadas a las llamadas asociadas al servicio de call center, desde su creación de origen o los modificados desde la última ejecución del proceso.

Origen de los datos:

fact_call_center.extraccion.out

Destino de los datos

destino_bi.fact_call_center

Tabla 3.42: Fact Call Center Carga

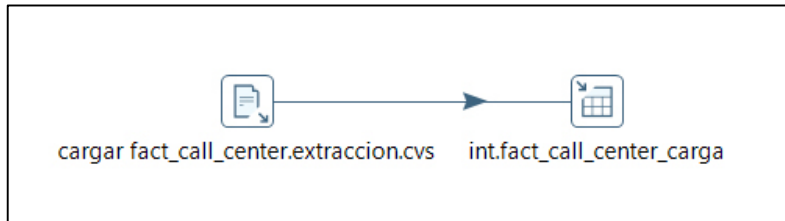


Figura 3.32: Fact Call Center Carga

Proceso: FACT_REGISTRO_CLIENTES_0900_SMS.KBJ

Objetivo:

Proceso encargado de invocar a los procesos de validación de rangos de fechas y días a procesar.

Origen de los datos:

Ninguno

Destino de los datos

Ninguno

Tabla 3.43: Fact Registro Clientes 0900

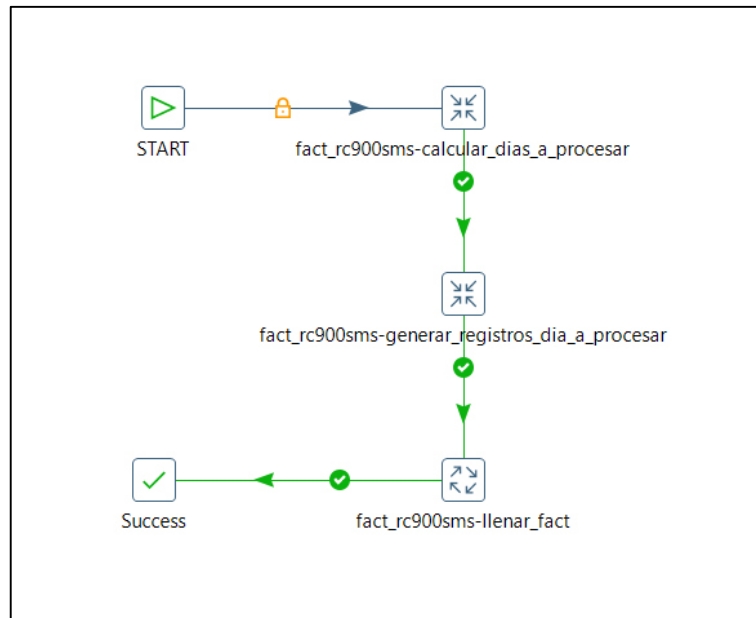


Figura 3.33: Fact Registro Clientes 0900

Proceso: FACT_RC900SMS-CALCULAR_DIAS_A_PROCESAR.KTR	
Objetivo:	Origen de los datos:
Proceso encargado de calcular los día a procesar.	Ninguno
	Destino de los datos:
	Memoria

Tabla 3.44: Fact RC900 SMS Calcular Días a Procesar

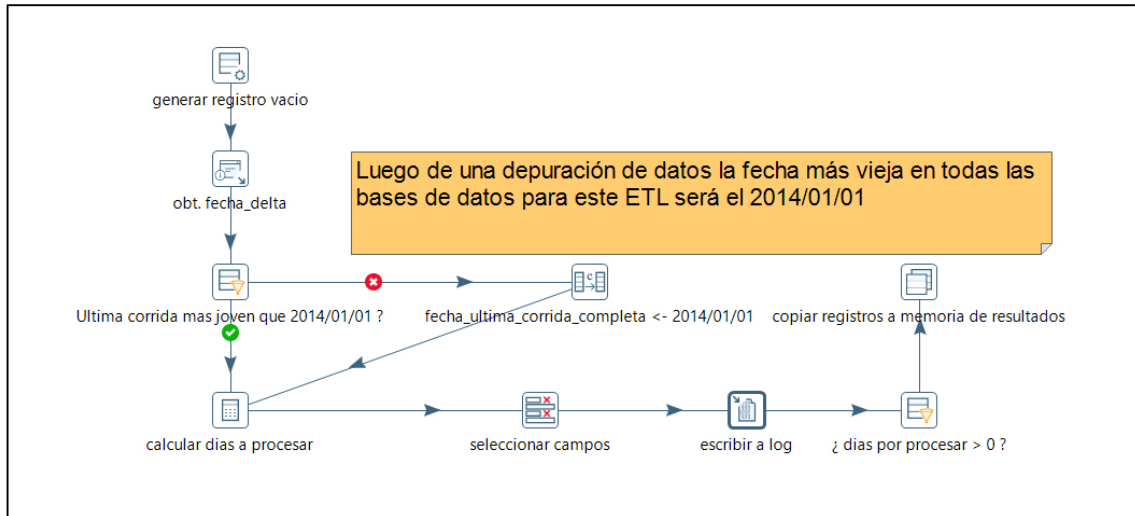


Figura 3.34: Fact RC900 SMS Calcular Días a Procesar

Proceso: FACT_RC900SMS-GENERAR_REGISTROS_DIA_A_PROCESAR

Objetivo:

Proceso encargado de crear los registros a procesar.

Origen de los datos:

Ninguno

Destino de los datos:

Memoria

Tabla 3.45: Fact RC900 SMS Generar Registro Día a Procesar

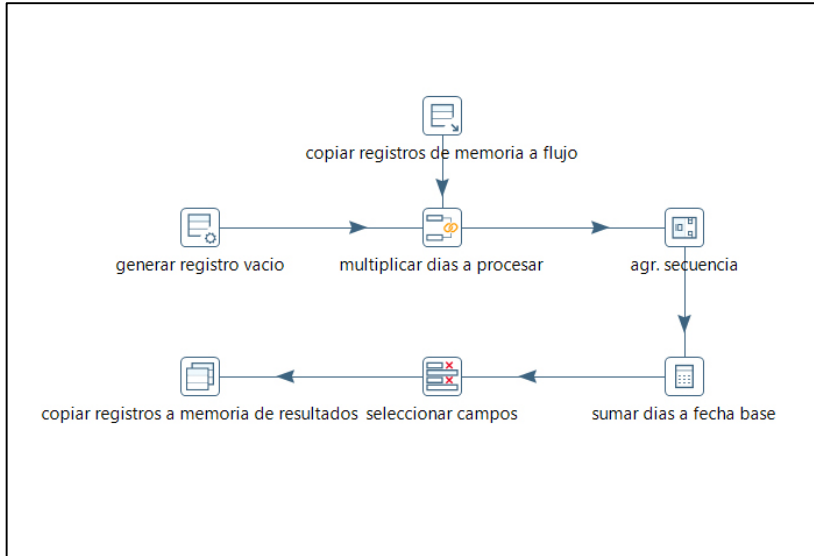


Figura 3.35: Fact RC900 SMS Generar Registro Día a Procesar

Proceso: ACTUALIZAR-ALMACEN-MENSUAL.KJB	
Objetivo:	Origen de los datos:
Carga desde archivo la lista con los ANI obtenidos anteriormente	dim_asociado.extraccion.out
	Destino de los datos:
	destino_bi.dim_asociado

Tabla 3.46: Actualizar Almacén Mensual

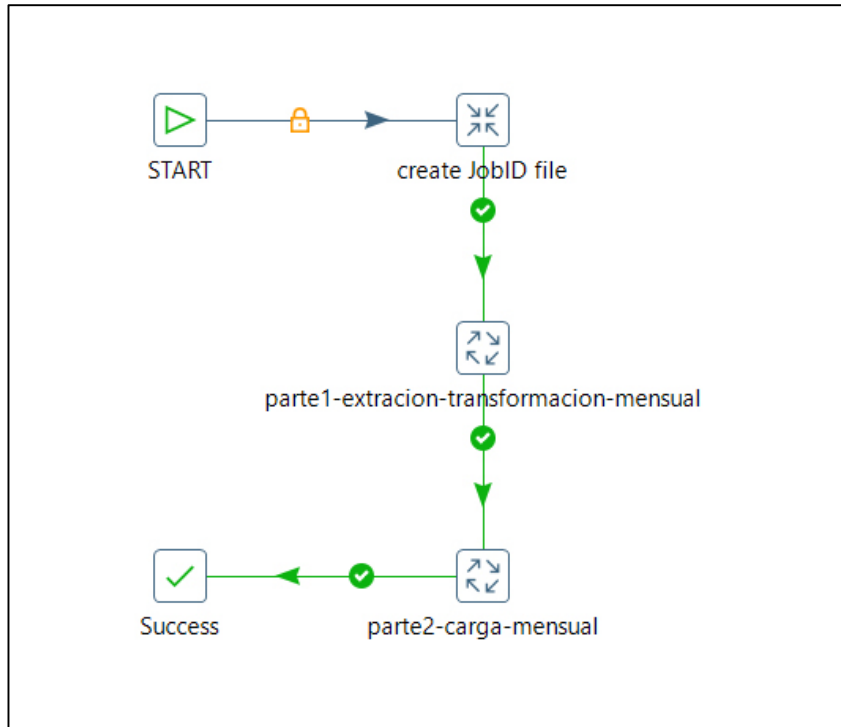


Figura 3.36: Actualizar Almacén Mensual

Proceso: PARTE1-EXTRACION-TRANSFORMACION-MENSUAL.KJB	
Objetivo:	Origen de los datos:
Punto de entrada para los procesos de extracción transformación y carga referente al manejo de las transacciones mensuales. Crea el archivo de control de la ejecución actual. Ejecuta los procesos: parte1-extracion-transformacion-mensual.kjb y parte2-carga-mensual.kjb.	Ninguno
	Destino de los datos:
	Ninguno

Tabla 3.47: Parte1 ETL Mensual

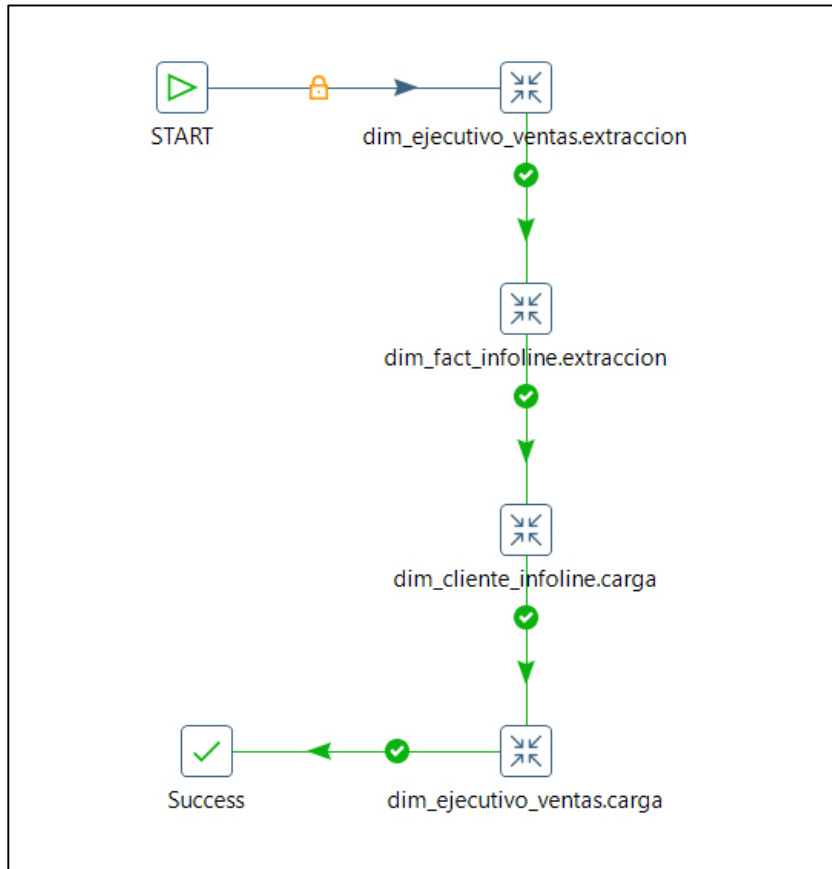


Figura 3.37: Parte1 ETL Mensual

Proceso: DIM_EJECUTIVO_VENTAS.EXTRACCION.KTR	
Objetivo:	Origen de los datos:
<p>Extrae por rango de fecha y almacena en archivo los registros asociados a los ejecutivos de cuenta, extraídos desde el origen que hayan sido creados o modificados desde la última ejecución del proceso.</p>	db_facturacion.ad_ejectuvo_cuenta
	Destino de los datos
	dim_ejecutivo_ventas.extraccion.out

Tabla 3.48: Proceso Dimensión Ejecutivo de Ventas

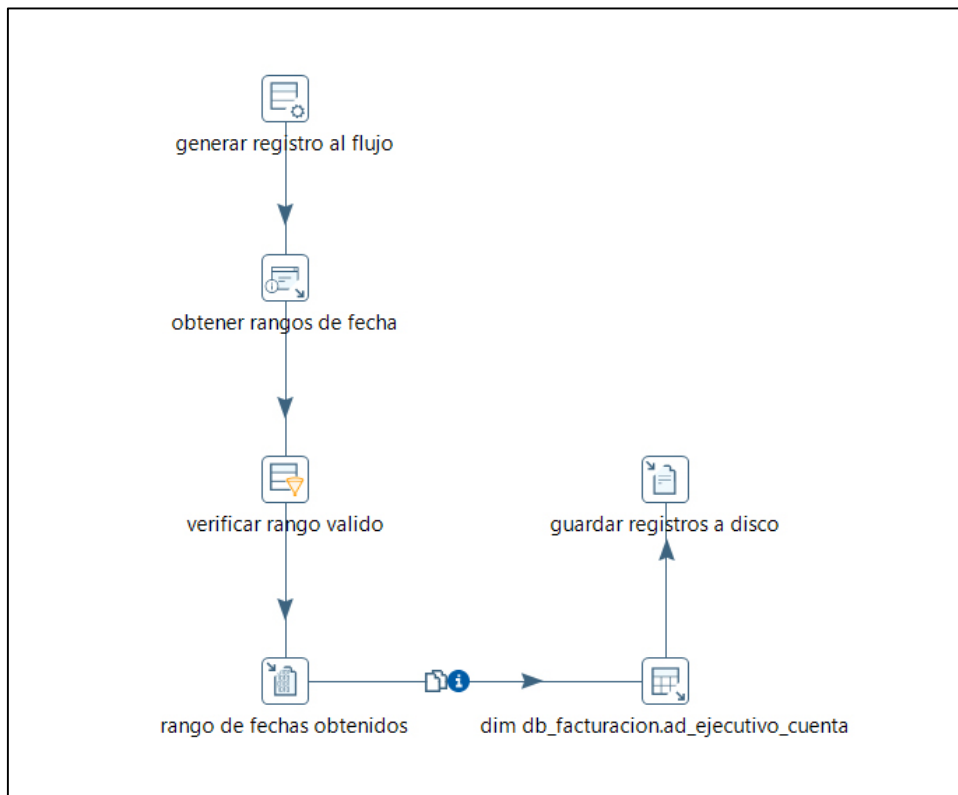


Figura 3.38: Proceso Dimensión Ejecutivo de Ventas

Proceso: FACT_INFOLINE.EXTRACCION.KTR	
Objetivo:	Origen de los datos:
Extrae, transforma y almacena por rango de fecha en archivo los registros extraídos asociados al servicio infoline desde el origen que hayan sido creados o modificados desde la última ejecución del proceso.	db_infoline_mercadeo.movimientos_infoline db_infoline_mercadeo.telefonos
	Destino de los datos
	fact_infoline.extraccion.out

Tabla 3.49: Proceso tabla de hechos Infoline

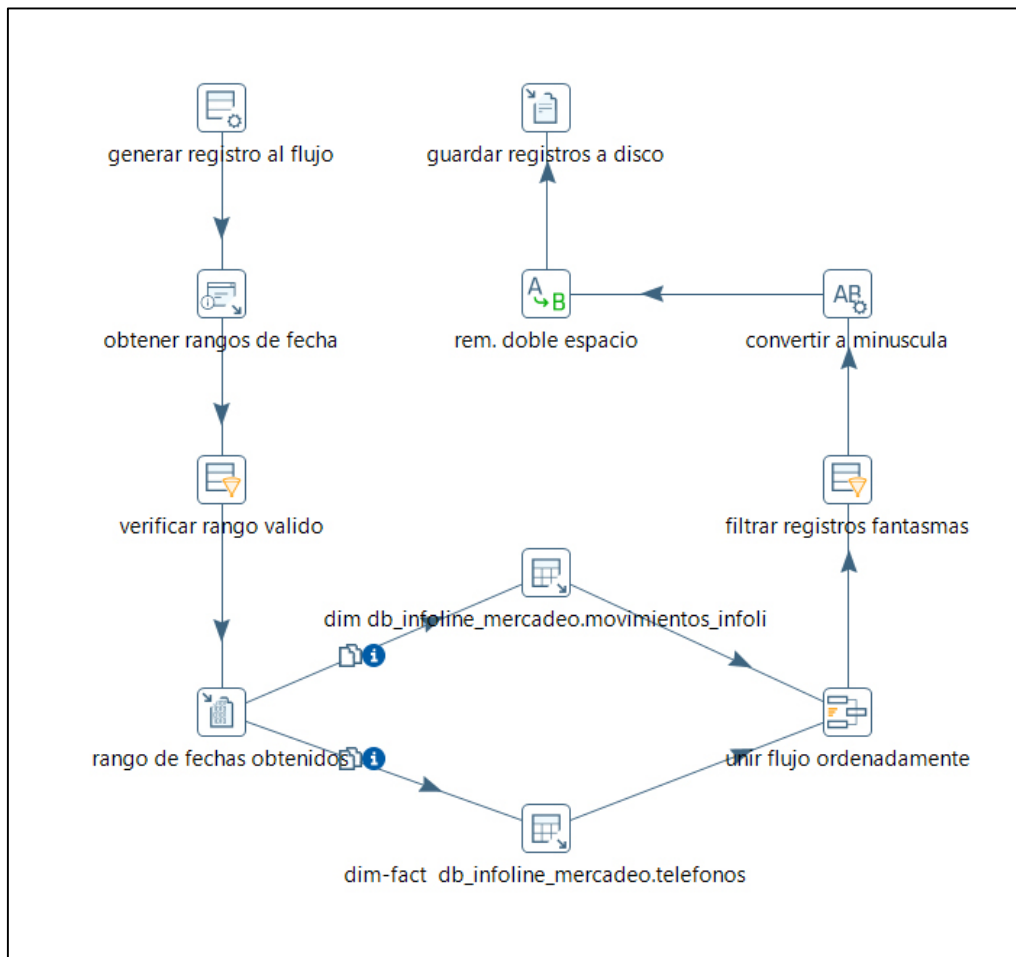


Figura 3.39: Proceso tabla de hechos Infoline

Proceso: DIM_CLIENTE_INFOLINE.CARGA.KTR	
Objetivo:	Origen de los datos:
Extrae de archivo y carga en la base de datos destino cliente_infoline con la información asociada al servicio infoline, desde su creación de origen o los modificados desde la última ejecución del proceso.	fact_infoline.extraccion.out
	Destino de los datos:
	destino_bi.dim_cliente_infoline

Tabla 3.50: Dim Cliente Infoline

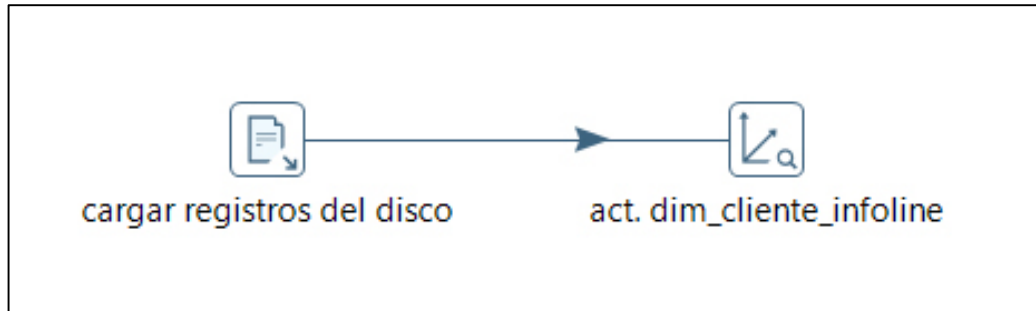


Figura 3.40: Dim Cliente Infoline

Proceso: DIM_EJECUTIVO_VENTAS.CARGA.KTR

Objetivo:

Extrae de archivo y carga en la base de datos destino ejecutivo_ventas con la información asociada al servicio ejecutivo ventas, desde su creación de origen o los modificados desde la última ejecución del proceso.

Origen de los datos:

db_facturacion.ad_asociado

Destino de los datos:

destino_bi.dim_ejecutivo_ventas

Tabla 3.51: Dim Ejecutivo Ventas

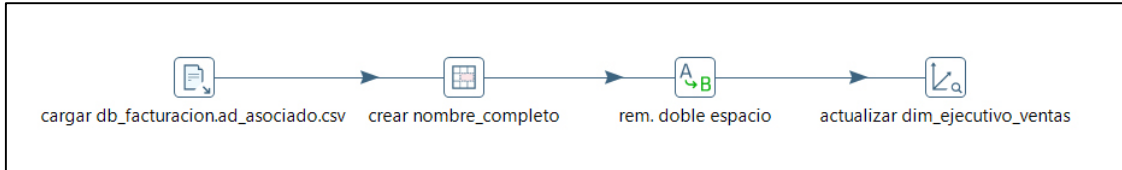


Figura 3.41: Dim Ejecutivo Ventas

Proceso: PARTE2-CARGA-MENSUAL	
Objetivo:	Origen de los datos:
Punto de entrada para los procesos de extracción transformación y carga referente al manejo de las transacciones mensuales.	Ninguno
	Destino de los datos:
	Ninguno

Tabla 3.52: Parte2-Carga-Mensual

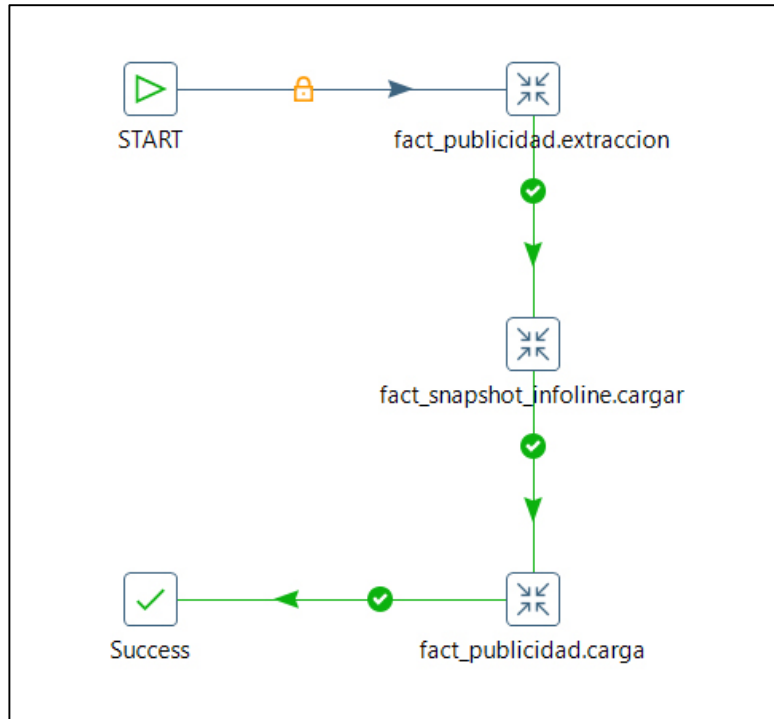


Figura 3.42: Parte2-Carga-Mensual

Proceso: FACT_PUBLICIDAD.EXTRACCION.KTR	
Objetivo:	Origen de los datos:
Extrae, transforma y almacena por rango de fecha en archivo los registros extraídos asociados al servicio publicidad desde el origen que hayan sido creados o modificados desde la última ejecución del proceso.	Ninguno
	Destino de los datos:
	Memoria

Tabla 3.53: Fact Publicidad Extracción

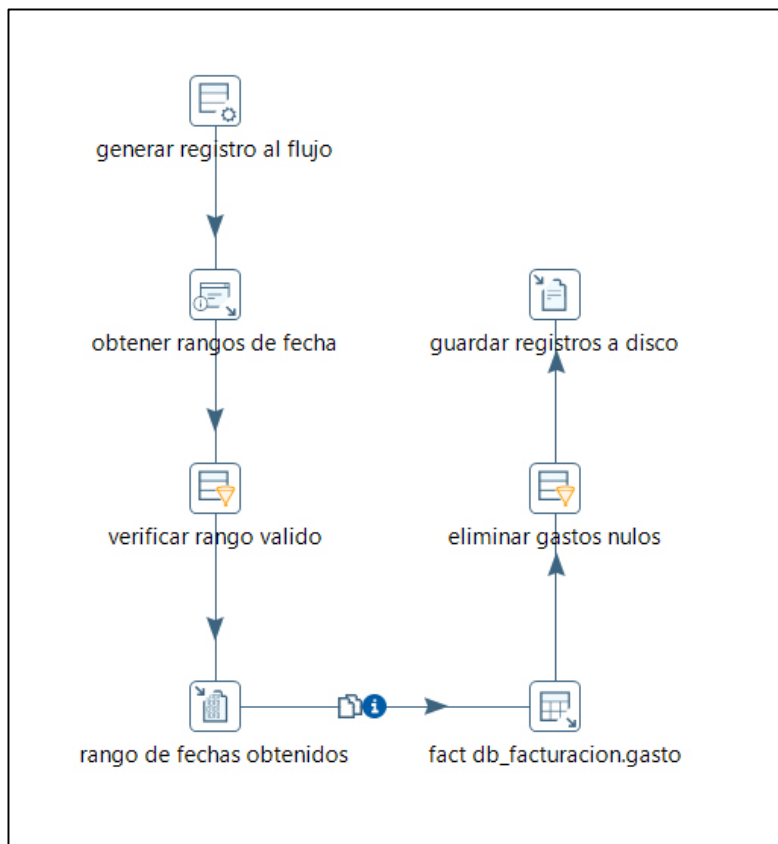


Figura 3.43: Fact Publicidad Extracción

Proceso: FACT_SNAPSHOT_INFOLINE.CARGA.KTR	
Objetivo:	Origen de los datos:
Extrae, transforma y almacena por rango de fecha en archivo los registros extraídos asociados al servicio infoline desde el origen que hayan sido creados o modificados desde la última ejecución del proceso.	fact_infoline.extraccion.out
	Destino de los datos:
	destino_bi.fact_snapshot_infoline_cliente

Tabla 3.54: Fact Snapshot Infoline

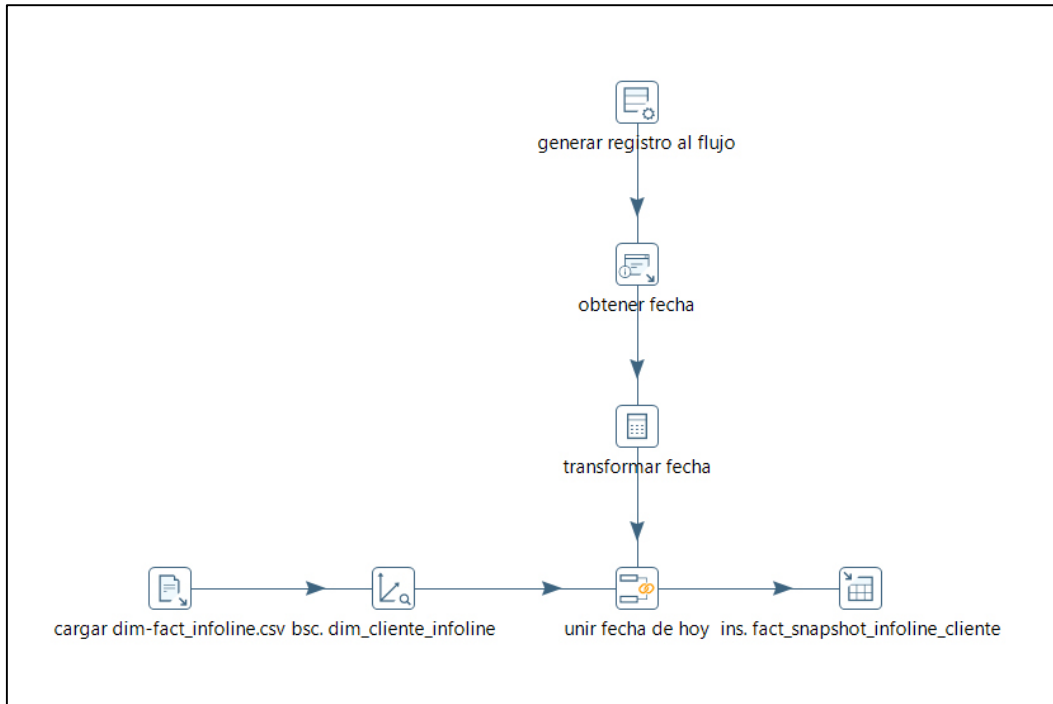


Figura 3.44: Fact Snapshot Infoline

Proceso: DIM_EJECUTIVO_VENTAS.CARGA.KTR	
Objetivo:	Origen de los datos:
Extrae, transforma y almacena por rango de fecha en archivo los registros extraídos asociados al servicio ejecutivo_ventas desde el origen que hayan sido creados o modificados desde la última ejecución del proceso.	db_facturacion.gastos
	Destino de los datos:
	destino_bi.fact_publicidad

Tabla 3.55: Dim Ejecutivo Ventas

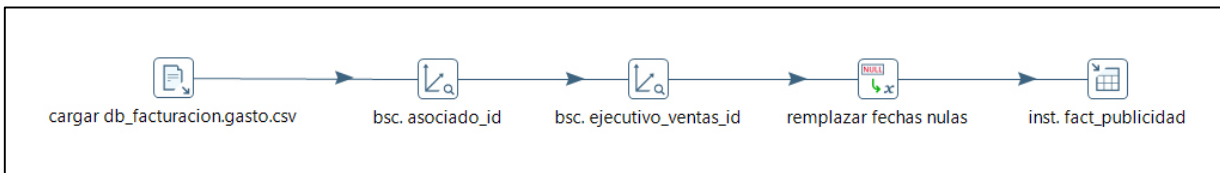


Figura 3.45: Dim Ejecutivo Ventas

3.7 Fase IV: Despliegue

Teniendo en cuenta que la no disposición de permisos de administrador, especialmente en los entornos de desarrollo y calidad, contextualizó el alcance de este TEG, es importante acotar que el mismo se mantuvo en desarrollo hasta el proceso de calidad, debido a que la estrategia de la organización para cumplir con la metodología sugerida se centró en trasladar al personal interno de la empresa los procesos de prueba en los ambientes de producción.

Este personal fue adiestrado y acompañado para tal fin, es decir, realizar las pruebas dentro de la organización con los máximos responsables de los procesos de toma de decisión, todo ello para verificar que efectivamente el sistema respondía satisfactoriamente las necesidades del negocio, teniendo en cuenta también la limitación de que el desarrollo era a distancia. En este sentido, y aun cuando los protocolos internos de la empresa a nivel de pruebas técnicas y funcionales se desconocían a nivel del desarrollador de este TEG, lo que se si obtuvo fue la verificación de que el sistema operaba correcta y satisfactoriamente, según pruebas de aceptación realizadas entre 21/10/2013 y 28/10/2013.

Estas pruebas de aceptación consistieron en la verificación de todos los componentes del sistema, así como de su funcionamiento, todo ello para verificar el grado general de satisfacción y logro de objetivos. Se muestran a continuación algunas capturas de pantalla relacionadas con consultas específicas:



Figura 3.46: Cuadro de mandos Demo Total de llamadas por operadora y total de SMS por operadora



Figura 3.47: Clientes nuevos y desafiliados

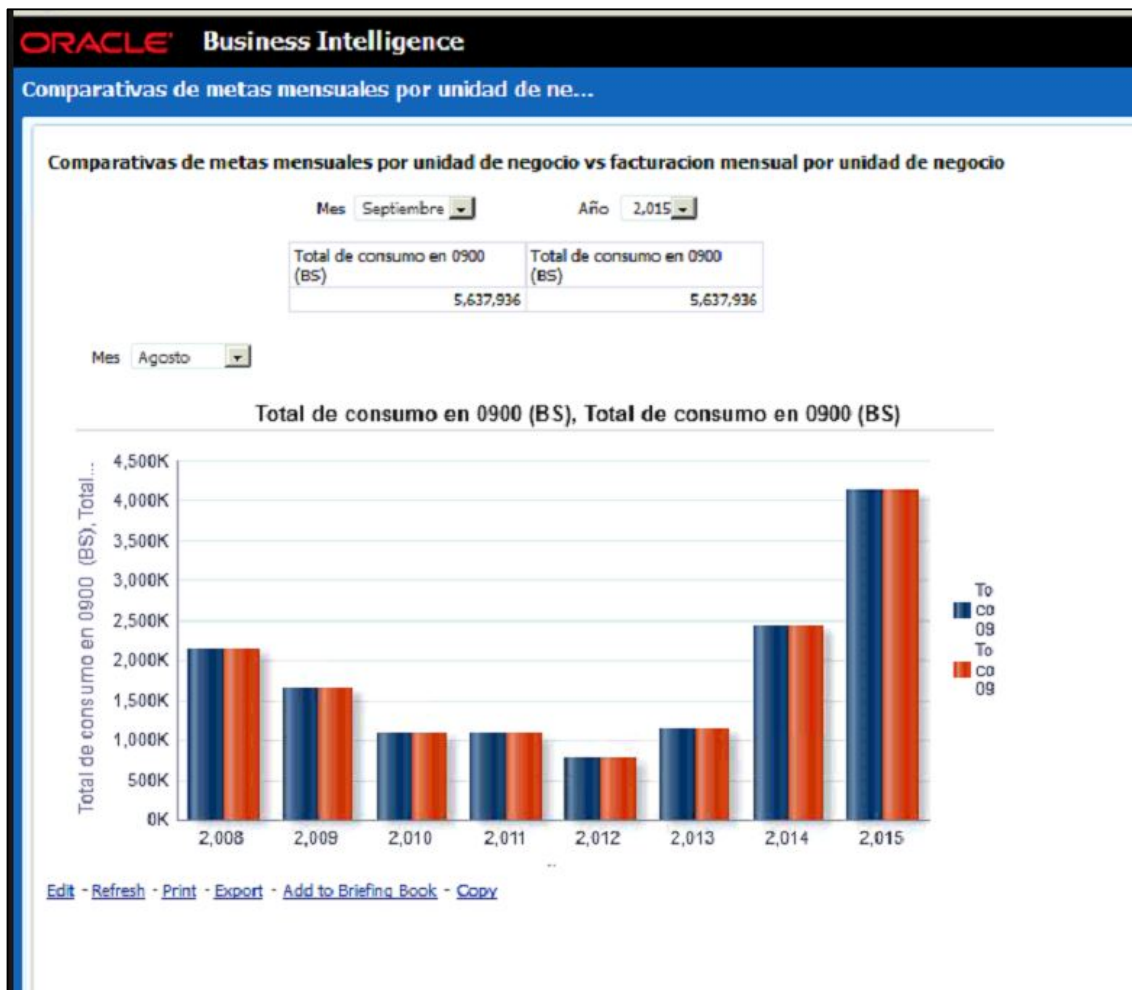


Figura 3.48: Meta

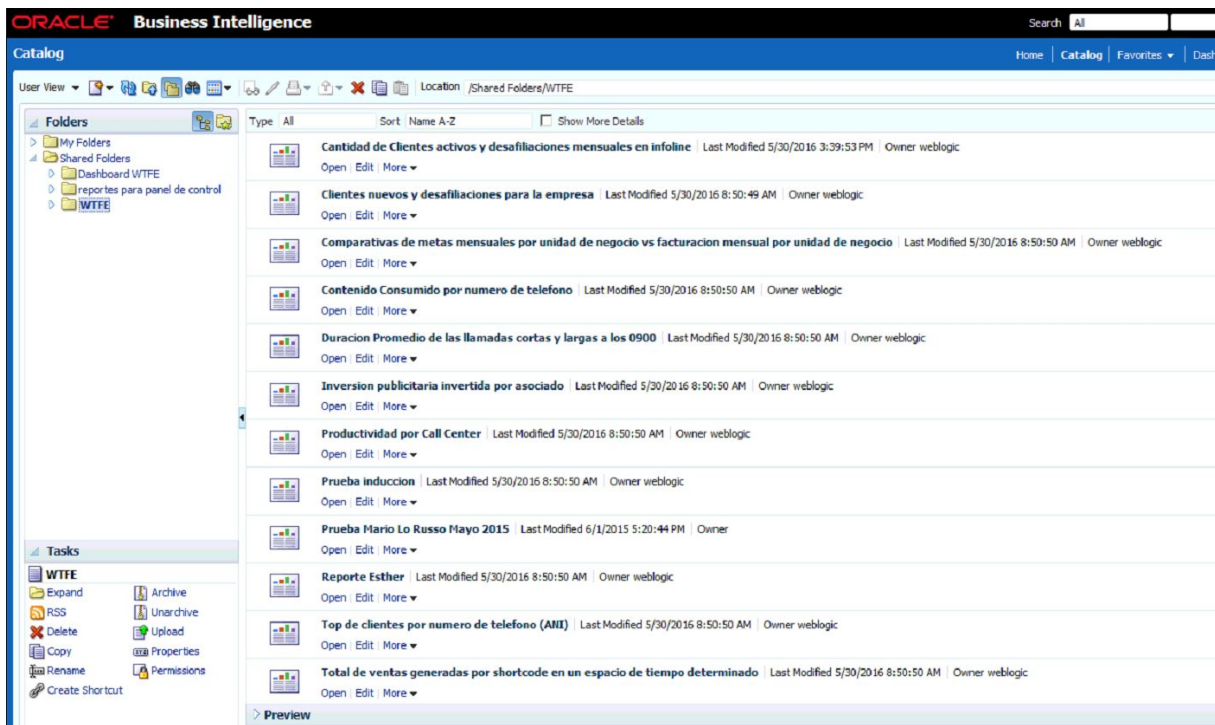


Figura 3.49: Catálogo

Siendo esta la última fase de la solución a implementar, se buscó ante todo garantizar el buen funcionamiento del sistema tanto en el ambiente de desarrollo como en el de calidad.

Se ha de aclarar que los ambientes de desarrollo no fueron descartados, ya que son útiles al momento de extender el sistema, una vez se encuentre desarrollado y puesto en operación.

Se muestran a continuación las características del ambiente de calidad a los efectos de este despliegue:

- Servidores Privado Virtual, administrado con VMware Vsphere 5.0
- Sistema operativo: Oracle Linux 5.5 / Windows Server 2008 R2
- 1 TB de espacio en disco y 6 GB de memoria RAM
- OpenJDK 1.7.0_40
- Oracle BI 11G
- Pentaho Data Integrator 4.4.0
- MySQL 5.5.34

Como comentario final al entorno aplicativo se acota que todos los elementos metodológicos son determinados por el entorno y su posibilidad de desarrollo, por lo que aspectos como extensibilidad y mantenimiento del sistema, así como aquellos asociados a estos, no se consideran aplicables para el logro de los objetivos de este TEG.

Conclusiones

De acuerdo con lo mostrado en los capítulos anteriores, se confirma el logro del desarrollo de un Sistema de Inteligencia de Negocio basándose en una mezcla de tecnologías de código abierto y sistemas propietarios.

Se logra reducir muy considerablemente los tiempos de los reportes y consultas analíticas del negocio, gracias a las múltiples bondades del sistema mencionadas, explicadas y desarrolladas en este trabajo.

El modelo dimensional planteado no solamente cubre las necesidades actuales, sino que también tiene la capacidad de afrontar las nuevas inquietudes asociadas a los procesos de las áreas del negocio atendidas, permitiendo la correcta extensibilidad del sistema a futuros cambios en los requerimientos. En este sentido, el sistema de gestión de bases de datos utilizado, cumple con las exigencias requeridas para el sistema de inteligencia de negocio propuesto y ejecutado, manteniendo el estándar definido por la organización para el manejo y control de sus datos, por lo que se afirma que efectivamente se desarrolló un almacén de datos con un modelo congruente, que satisface el conjunto de necesidades analíticas del negocio.

Por otro lado, y dado que todos los entregables propuestos en cada una de las fases del desarrollo fueron alcanzados, se comprueba la efectividad de la metodología seleccionada, Kimball Lifecycle, para el desarrollo de Sistemas de Inteligencia de Negocios.

Así mismo, y con base en las distintas pruebas realizadas y en la buena recepción por parte de los usuarios técnicos y funcionales, se puede afirmar que el Sistema de Inteligencia de Negocios disminuye notablemente no solo los tiempos de desarrollo de reportes con fines analíticos, sino los reportes y consultas en sí mismas, excediendo las expectativas en algunos casos.

Finalmente, y de acuerdo a todo lo anterior, se logra efectiva y rotundamente el desarrollo de un sistema eficiente de inteligencia de negocio que apoya el proceso de toma de decisiones en el departamento de mercadeo de la organización estudiada.

Observaciones y recomendaciones

Es importante resaltar que dada la naturaleza personalizada del proyecto y al tiempo de ejecución que transcurre desde las reuniones iniciales a la entrega final del producto los requerimientos iniciales del cliente pueden cambiar durante la implementación. Esto debido a rotación del personal en la organización, nuevos servicios que surgen dentro del negocio y que se quisieran incluir en los reportes, servicios que se dejan de ofrecer y que ya no tendría sentido analizar, etc. Lo recomendado para evitar inconvenientes es:

- Apegarse fielmente a metodologías de manejo de proyectos.
- Contar totalmente con el apoyo de los patrocinadores del proyecto.
- Asegurar el entendimiento de los requerimientos.
- Realizar documentación de cada una de las fases (Firmar los requerimientos levantados, los modelos, los procesos de extracción y los reportes o indicadores junto a los tableros finales); no avanzar hasta tener la aprobación de los involucrados.
- Entender los cambios de alcances y documentarlos como tal, para su posterior implementación.

De la misma manera es importante realizar un buen dimensionamiento de los equipos dispuestos por la organización para la implementación del proyecto, tomarse el tiempo para estimar la cantidad de recursos que se utilizaran para los procesos de ETLs y el impacto que pueden ocasionar en las operaciones diarias de la empresa.

Referencias Bibliográficas

Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Thomas, D. (2001). Manifesto for Agile Software Development. Retrieved March 23, 2014, from <http://agilemanifesto.org/>

Darmawikarta, D. (2007). Dimensional Data Warehousing with MySQL. Boston: BrainySoftware.

Eckerson, W. (n.d.). Four Ways to Build a Data Warehouse. Retrieved from <http://www.bi-bestpractices.com/view-articles/4770>

EcuRed.cu (2012). EcuRed. Recuperado el 21 de mayo de 2012, de: <http://www.ecured.cu/index.php>

Elliott, G. (2004). Global Business Information Technology: An Integrated Systems Approach. Pearson Addison Wesley.

Guarente, M. (2002). ¿HOW is your business doing? London: Sunday Business.

Hamel, G., & Prahalad, C. (1994). Competing for the future. Boston: Harvard Business School Press.

Inmon, B. (1998). "Data Mart Does Not Equal Data Warehouse". Recuperado el 13 de marzo de 2012, de: <http://www.information-management.com/infodirect/19991120/1675-1.html>

Inmon, W. H. (1999). Building the Operational Data Store. Wiley.

Inmon, W. (2000). Data Mart Does Not Equal Data Warehouse. California: DM Review's.

Inmon, W. H. (2002). Building the Data Warehouse (3rd ed.). Wiley.

Inteligencia de Negocios. (2001). Recuperado el 24 de Agosto de 2011, de Idensa: <http://www.idensa.com/>

Julie, S., & Steinbart. (2000). Data Warehousing and Data Mining: Opportunities for Internal Auditors. Florida: The Institute of Internal Auditors Research.

Kimball, R., & Margy, R. (2002). The Data Warehouse Toolkit the Complete Guide to Dimensional Modeling (2da ed.). New York: Wiley.

Kimball, R. (1997). "A Dimensional Model Manifesto". Recuperado el 2 de marzo de 2012, de: http://www.ralphkimball.com/html/articles_search/articles1997/9708d15.html

Kimball, R., & Caserta, J. (2008). The Data Warehouse ETL Toolkit (2da ed). New York: Wiley.

Kimball, R., & Ross, M. (2013). The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling (3rd ed.). John Wiley & Sons.

Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., & Becker, B. (2008). *The Data Warehouse Lifecycle Toolkit* (2nd ed.). Wiley Publishing.

Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., & Becker, B. (2011). *The Data Warehouse Lifecycle Toolkit* (2nd ed.). John Wiley & Sons.

Laudon, K. C. y Laudon J. P. (2004). *Sistemas de Información Gerencial*. México: Pearson Prentice Hall.

Loshin, D. (2003). "Business Intelligence: The Savvy Manager's Guide". Morgan Kaufmann Publishers.

Martínez, J. (2002). *¿Qué es la Inteligencia de Negocios y qué beneficios ofrece?* México: MicroStrategy.

Moraga, M (2001). "El modelo de datos jerárquico". Universidad de Castilla - La Mancha. Recuperado el 20 de marzo de 2012, de: http://alarcos.inf-cr.uclm.es/doc/bda/doc/trab/T0001_MAMoraga.pdf

Oppel, A. (2009). "Databases". Estados Unidos de Norteamérica: McGraw Hill.

Saroka, R. (2002). *Sistemas de Información en la era digital*. Buenos Aires: Fundación OSDE.

SIPEC. (2011). Reportes. Recuperado el 13 de octubre de 2011, de SIPEC: <http://sipec.sep.gob.mx/WebHelp/reportes/reporte.htm>

Tegarden, D. P., Dennis, A., & Wixom, B. H. (2009). *Systems Analysis and Design with UML* (3ra ed.). New York, NY, USA: John Wiley & Sons, Inc.

Vélez, C. (2000). *Indicadores*. Recuperado el 13 de octubre de 2011, de SIRAC: <http://www.sirac.info/Hospitales/html/indicadores.asp>

Verástegui, H. (2007). *Modelo Dimensional de Datos*. Recuperado el 24 de agosto de 2011, de Db-system: http://www.db-system.com/pls/portal/docs/PAGE/SITIOWWDB/ARTICULOS/MODELADO%20DIMENSIONAL%20DE%20DATOS_V2.PDF

Williams, S., & Williams, N. (2007). "The Profit Impact of Business Intelligence". Morgan Kaufmann Publishers.