UNIVERSIDAD CENFOTEC

Tesis de Maestría en Bases de Datos

Diseño de un Data Mart para JASEC, modelado mediante UML

Elaborado por
Lic. Osvaldo Navarro Navarro

Ingeniero en Sistemas de Información

Enero, 2014

Declaración Jurada

Yo, Osvaldo Navarro Navarro, mayor, casado, estudiante de la carrera de Maestría Profesional en Tecnologías de Bases de Datos de la universidad CENFOTEC, domiciliado en Tejar de El Guarco de Cartago, portador de la cédula de identidad número 1-893-548, en este acto, debidamente apercibido y entendido de las penas y consecuencias con las que se castiga en el Código Penal, el delito de perjurio ante quienes se constituyen en el Tribunal Examinador de mi trabajo de tesis para optar por el título de Maestría Profesional en Tecnologías de Bases de Datos, juro solemnemente que mi trabajo de investigación titulado **Diseño de un Data Mart para JASEC, modelado mediante UML** es una obra original que ha respetado todo lo preceptuado por las leyes Penales así como la Ley de Derechos de Autor y Derechos Conexos, número 6683 de 25 noviembre de 1982 y sus reformas, publicada en la Gaceta número 226 de 25 noviembre de 1982, incluyendo el numeral 70 de dicha ley que advierte: Artículo 70°: Es permitido citar a un autor transcribiendo los pasajes pertinentes siempre que estos no sean tantos y seguidos, que pueda considerarse como una producción simulada y sustancial, que redunde en prejuicio del autor y de la obra original. En fe de lo anterior firmo en la ciudad de San José, el 23 del mes de Enero del año dos mil catorce.

_____

Lic. Osvaldo Navarro Navarro.
Cédula: 1-893-548

Tribunal Examinador

Este proyecto fue aprobado por el Tribunal Examinador de la Maestría en Tecnologías de Bases de Datos de la Universidad CENFOTEC, como requisito para optar por el grado de Maestría Profesional en Tecnologías de Bases de Datos.

_____	_____
Msc. Luis Carlos Naranjo.	Msc. Ignacio Trejos Zelaya.
Tutor.	Lector.

_____

Msc. Guillermo Gómez Tenorio.

Lector.

_____

Lic. Osvaldo Navarro Navarro.

Sustentante.

Dedicatoria

Este trabajo se lo dedico a mi **DIOS** quien me ha guiado por el buen camino, me dio las fuerzas y esperanzas necesarias para seguir adelante.

A mis padres Saray Navarro Brenes y Edgar Navarro Brenes quienes me apoyaron e inculcaron el deseo de superación.

A mi esposa Guiselle Solís Torres, quien con su sacrificio y dedicación ha permitido que utilice el tiempo que ella merece, por alcanzar mis metas durante estos años.

A mis hijas Diana y Jimena, gracias por ser las estrellas que han guiado mi camino, hasta alcanzar esta meta.

Al profesor **Luis Naranjo Zeledón**, Ph.D., porque bajo su dirección obtuve una mejora y obtención de nuevos conocimientos para el desarrollo de mi tesis, además de ser una persona que demuestra un alto nivel de profesionalidad en sus actividades.

"*La dicha de la vida consiste en tener siempre algo que hacer, alguien a quien amar y alguna cosa que esperar*" **Thomas Chalmers.**

Tabla de Contenido

Declaración Jurada 2
Tribunal Examinador 3
Dedicatoria 4
Tabla de Contenido 5
Índice de figuras 8
Índice de Tablas 11
Introducción 12
Marco Institucional 12
Historia 12
Marco Jurídico 13
Misión 17
Visión. 17
Política de Calidad 17
Estructura Organizativa 18
Antecedentes 23
Justificación 24
Planteamiento del Problema 31
Problema General 31
Subproblemas 31
Objetivos 32
General 32
Objetivos Específicos 32
Alcances 33
Limitaciones 33
Marco Teórico 34
Data Warehouse 34
Data Mart 35
Data Mart vs. Data Warehouse 36
Sistemas de Extracción, Transformación y Carga (ETL) 37
Modelo Multidimensional 38
Tablas 38
Esquemas 39
Cubo OLAP 40
Modelamiento de Datos 42

Modelo de datos 42
Modelamiento de datos 42
Lenguaje de Modelado Unificado (U.M.L.) 43
Métodos de Diseño 44
Proceso de ingeniería para el desarrollo de un DW (DWEP) 46
Marco Metodológico 55
Tipo de investigación o proyecto 55
Sujetos y fuentes de información 56
Fuentes de Información 56
Sujetos de Información 56
Descripción de Instrumentos 57
Metodología de Desarrollo 59
Definición de la Metodología 59
Identificación de Requerimientos 60
Situación Actual 63
Análisis 65
Diseño Conceptual 65
Diseño Lógico 68
Diseño 71
Arquitectura 71
Conclusiones 73
Cronograma 75
Work Breakdown Structure (WBS) 77
Referencias 78
APÉNDICES 80
Apéndice #1: Cuestionario de Entrevista 80
Apéndice #2: Informe trimestral 83
Apéndice #3: Diagramas UML 84
Apéndice #4: Diagrama de Actividades 87
Apéndice #5: Documento de Análisis de Requerimientos 89
Apéndice #6. Diagrama de Casos de Uso 95
Apéndice #7. Esquema Conceptual/Análisis Dimensional. 97
Apéndice #8. Datos Fuente. Esquema Conceptual y Lógico 120
Apéndice #9. Mapeo de Datos 122

Apéndice #10. Esquema Conceptual del Datamart (DWCS)/ Esquema Lógico del Datamart(DWLS). 134
Apéndice #11: Procesos ETL. 136

Índice de figuras

Figura 1. División Distrital del Área Servida por JASEC (JASEC, 2013)..... 13
Figura 2. Generación propia y compras de potencia en MW (PDR, 2012). 14
Figura 3. Organigrama JASEC – Nivel 1 (RH, 2013). 19
Figura 4. Organigrama JASEC – Negocio Energía (RH, 2013). 20
Figura 5. Organigrama JASEC – Negocio Servicios Corporativos (RH, 2013)..... 21
Figura 6. Organigrama JASEC – Negocio Acueductos y Saneamiento Ambiental (RH, 2013). 22
Figura 7. Organigrama JASEC – Negocio Infocomunicaciones (RH, 2013). 22
Figura 8. Comportamiento Histórico de Consumo de Energía hasta 2035 (PDR, 2012). 25
Figura 9. Distribución de Consumo de Energía por Cantón, 2011 (PDR, 2012). 26
Figura 10. Cantidad de solicitudes de estudio, año 2013, primer y segundo trimestre (PDR, 2013). 27
Figura 11. Crecimiento de la Demanda de Energía vs. Generación de plantas de JASEC (PDR, 2012). 28
Figura 12. Tarifa Residencial (ARESEP, 2013) 29
Figura 13. Data Mart Dependiente (Elaboración Propia). 35
Figura 14. Data Mart independiente (Elaboración Propia). 36
Figura 15. ETL - Extraer – Transformar – Cargar (Elaboración Propia). 37
Figura 16. Esquema Copo de nieve (Kimball, 1996)..... 40
Figura 17. Esquema Estrella (Elaboración Propia). 40
Figura 18. Representación gráfica de un cubo de ventas (ORACLE 2010). 41
Figura 19. Tareas de la metodología de Kimball (Kimball, 1996) 45
Figura 20. Marco de diseño de un Data Warehouse (Luján, 2005). 48
Figura 21. Source Logical Schema (Luján, 2005). 51
Figura 22. DWCS nivel 1 (Luján, 2005)..... 51
Figura 23. DWCS nivel 2 (Luján, 2005)..... 52
Figura 24. DWCS nivel 3 (Luján, 2005)..... 52
Figura 25. DWPS (Luján, 2005). 53
Figura 26. Relación entre los Esquemas del DataMart (Zambrano, 2011). 54
Figura 27. Jerarquía de ubicación geográfica y de fecha (Elaboración Propia). 66
Figura 28. Esquema conceptual del modelo estrella de Solicitud de Estudio (Elaboración Propia). 66
Figura 29. Esquema Conceptual del Modelo Estrella de Inventario (Elaboración Propia). ... 67
Figura 30. Esquema Lógico, Modelo Solicitud Estudio (Elaboración Propia)..... 69
Figura 31. Esquema Lógico, Modelo Inventario (Elaboración Propia). 70
Figura 32. Diagrama de arquitectura (Elaboración Propia). 71
Figura 33. Diagramas UML 2.4. (UML, 2013). 84
Figura 34. Diagrama de Actividades. Solicitud de Estudios (Elaboración Propia). 87
Figura 35. Diagrama de Actividades. Movimiento del Almacén (Elaboración Propia). 88
Figura 36, Diagrama de contexto (Elaboración Propia). 91
Figura 37. Paquete Datamart (Elaboración Propia). 91
Figura 38. Paquete de Análisis de Solicitud de Estudios (Elaboración Propia). 92
Figura 39. Paquete de Análisis de Materiales (Elaboración Propia). 94
Figura 40. Diagrama Caso de Uso Análisis de Solicitudes (Elaboración Propia). 95

Figura 41. Diagrama Caso de Uso Análisis de Materiales (Elaboración Propia). 96
Figura 42. Relación entre los tres niveles (Luján, 2005). 97
Figura 43. Esquema conceptual del Datamart. Nivel 1 (Elaboración Propia). 99
Figura 44. Esquema estrella de Solicitud de Estudio (Elaboración Propia). 100
Figura 45. Esquema estrella de Materiales (Elaboración Propia). 101
Figura 46. Dimensión Cantón (Elaboración Propia)..... 102
Figura 47. Dimensión Estado (Elaboración Propia). 102
Figura 48. Dimensión Tipo de Cliente (Elaboración Propia). 103
Figura 49. Dimensión Fechas (Elaboración Propia). 103
Figura 50. Dimensión Tipo de Servicio (Elaboración Propia). 104
Figura 51. Dimensión Categoría de Obra (Elaboración Propia). 104
Figura 52. Dimensión Artículos (Elaboración Propia). 105
Figura 53. Dimensión Personas (Elaboración Propia). 105
Figura 54. Dimensión Almacén (Elaboración Propia). 106
Figura 55. Dimensión Solic. Estudio (Elaboración Propia). 106
Figura 56. Esquema Conceptual de la Tabla de Hechos Fact_SolicEstudio (Elaboración Propia). 107
Figura 57. Esquema Conceptual de la Tabla de Hechos Fact_Material (Elaboración Propia). 108
Figura 58. Diagrama SCS (Elaboración Propia). 120
Figura 59. Diagrama SLS (Elaboración Propia). 121
Figura 60. Diagrama de Mapeo. Niveles (Luján, 2005). 123
Figura 61. Mapeo de Datos. Nivel 0 (Elaboración Propia)..... 124
Figura 62. Mapeo de Datos. Nivel 1 (Elaboración Propia)..... 125
Figura 63. Mapeo de Datos. Nivel 2. Cantones (Elaboración Propia). 125
Figura 64. Mapeo de Datos. Nivel 2. Estados (Elaboración Propia). 126
Figura 65. Mapeo de Datos. Nivel 2. Tipo de Cliente (Elaboración Propia). 126
Figura 66. Mapeo de Datos. Nivel 2. Personal (Elaboración Propia). 126
Figura 67. Mapeo de Datos. Nivel 2. Almacén (Elaboración Propia). 127
Figura 68. Mapeo de Datos. Nivel 2. Estudio Solicitud (Elaboración Propia). 127
Figura 69. Mapeo de Datos. Nivel 2. Artículos (Elaboración Propia). 127
Figura 70. Mapeo de Datos. Nivel 2. Tipo de Servicio (Elaboración Propia). 128
Figura 71. Mapeo de Datos. Nivel 2. Categoría de Obra (Elaboración Propia). 128
Figura 72. Mapeo de Datos. Cantones. Nivel 3 (Elaboración Propia). 129
Figura 73. Mapeo de Datos. Estados. Nivel 3 (Elaboración Propia). 129
Figura 74. Mapeo de Datos. Tipo de Cliente. Nivel 3 (Elaboración Propia). 130
Figura 75. Mapeo de Datos. Personal. Nivel 3 (Elaboración Propia). 130
Figura 76. Mapeo de Datos. Almacén. Nivel 3 (Elaboración Propia). 131
Figura 77. Mapeo de Datos. Estudio Solicitud. Nivel 3 (Elaboración Propia). 131
Figura 78. Mapeo de Datos. Artículos. Nivel 3 (Elaboración Propia). 132
Figura 79. Mapeo de Datos. Tipo de Servicio. Nivel 3 (Elaboración Propia). 132
Figura 80. Mapeo de Datos. Categoría de Obra. Nivel 3 (Elaboración Propia). 133
Figura 81. DWCS – Esquema Solicitud Estudio/Inventario (Elaboración Propia). 134
Figura 82. DWLS – Solicitud Estudio (Elaboración Propia). 135
Figura 83. Diagrama de proceso ETL. Cantones (Elaboración Propia). 137
Figura 84. Diagrama de proceso ETL. Estados (Elaboración Propia). 138
Figura 85. Diagrama de proceso ETL. Tipo de Cliente (Elaboración Propia). 139

Figura 86. Diagrama de proceso ETL. Personal (Elaboración Propia)..... 140
Figura 87. Diagrama de proceso ETL. Almacén (Elaboración Propia). 141
Figura 88. Diagrama de proceso ETL. Solicitud Estudio (Elaboración Propia)..... 142
Figura 89. Diagrama de proceso ETL. Artículo (Elaboración Propia). 143
Figura 90. Diagrama de proceso ETL. Tipo de Servicio (Elaboración Propia)..... 144
Figura 91. Diagrama de proceso ETL. Categoría de Obra (Elaboración Propia). 145
Figura 92. Diagrama de proceso ETL. Fechas (Elaboración Propia). 145

Índice de Tablas

Tabla 1. Diferencia DW y Data Mart (ORACLE 2007). 36
Tabla 2. Desarrollando modelos de datos (Kriegel, 2011). 43
Tabla 3. Artefactos empelados en DWEP (Luján, 2005)..... 49
Tabla 4. Flujo de trabajo y diagramas (Elaboración Propia). 50
Tabla 5. Tareas-Instrumento (Elaboración Propia)..... 58
Tabla 6. Resumen de Diagramas por Flujo de Trabajo del DWEP (Elaboración Propia)..... 60
Tabla 7. Necesidades de información (Elaboración Propia)..... 62
Tabla 8. Relación de tabla de hechos con dimensiones (Elaboración Propia)..... 68
Tabla 9. Estereotipos (Luján, 2005)..... 98
Tabla 10. Estereotipos ETL (Luján, 2005). 136

Introducción

Marco Institucional

Historia

Citando la historia de la Junta Administrativa de Servicios Eléctricos de Cartago (JASEC, 2012), un 12 de octubre de 1964 el pueblo cartaginés se lanzó a las calles para hacer valer sus derechos, con pérdidas de vidas humanas. Los cartagineses le dieron fuerza a un ideal: tener su propia empresa de servicios eléctricos.

En el año 1961 el Instituto Costarricense Energía (ICE) era el proveedor de energía eléctrica, y un inapropiado aumento de tarifas, más una desmejorada instalación eléctrica provocó la llamada "huelga de pagos eléctricos". Al ver la falta de pago del servicio eléctrico de los diferentes abonados, se procedió a realizar una corta masiva en la calle del comercio, lo que generó violencia entre el pueblo cartaginés y el ICE. Detrás de las cuadrillas que desconectaban, iban los huelguistas reconectando.

Dos años y medio duró la huelga, se improvisaban manifestaciones cada 8 o 15 días y en horas no programadas, volantes, comunicados de radio y un alto civismo demostrado por los cartagineses provocó que la Asamblea Legislativa agilizara el Proyecto de Ley para una Junta Eléctrica en la provincia de Cartago.

Marco Jurídico

La Junta Administrativa del Servicio Eléctrico de Cartago (JASEC) es una empresa cartaginesa generadora y distribuidora de energía. Distribuye el servicio eléctrico en la provincia de Cartago, brinda sus servicios a cinco cantones, tal como se observa en la figura 1, en donde se muestra los cantones en los cuales presta sus servicios.

Su área de cobertura es de 309 Km^2 de un total de 3,120 km^2, lo cual representa una cobertura de un 10% de la provincia.

Figura 1. División Distrital del Área Servida por JASEC (JASEC, 2013)

En cuanto a la generación, la institución cuenta con sus propias plantas generadoras. En la figura 2 se muestra cuánta energía genera y cuánta energía compra a otra entidad generadora y distribuidora de energía a nivel nacional, en términos porcentuales.

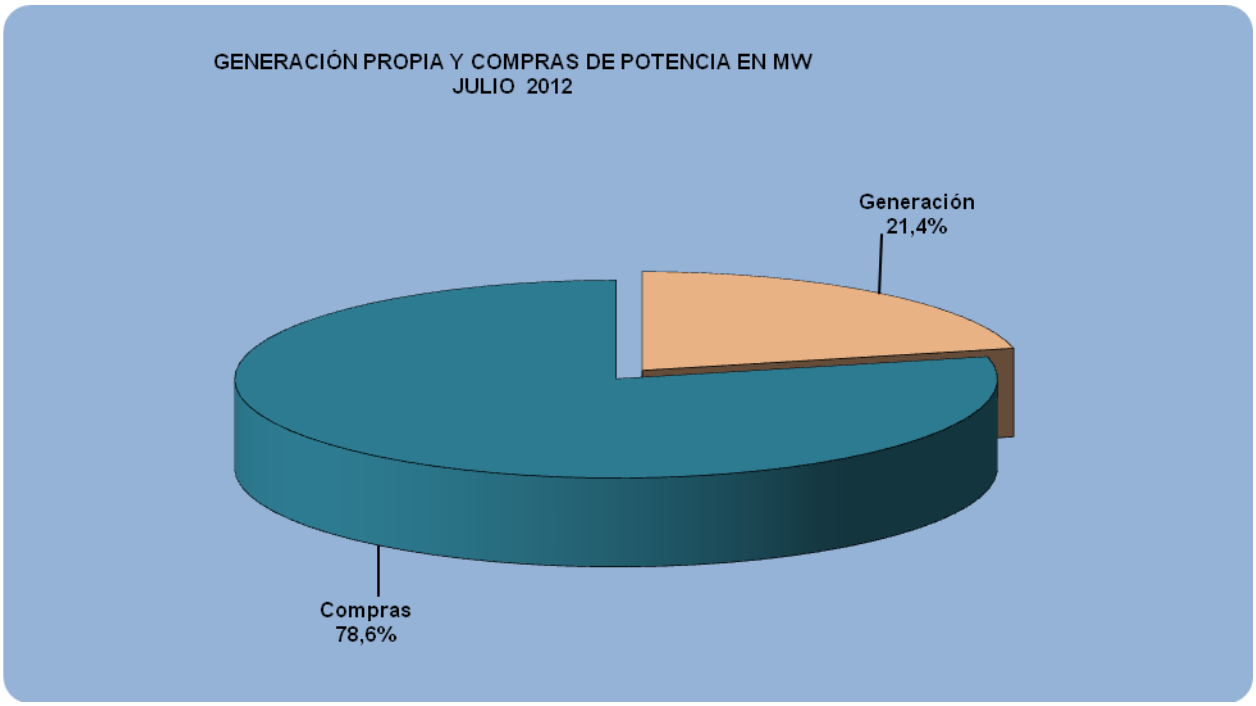


Figura 2. Generación propia y compras de potencia en MW (PDR, 2012).

JASEC fue constituida por medio de la Ley No. 3300 del 16 de julio de 1964 publicada en La Gaceta No. 166 del 23 de julio de 1964, con la finalidad de administrar la empresa eléctrica de la Municipalidad del Cantón Central de Cartago. Con base en la Ley No. 4881 del 9 de noviembre de 1971 publicada en la Gaceta No. 229 del 17 de noviembre de 1971, se autoriza a JASEC para prestar el servicio de alumbrado público brindado hasta ese momento por las municipalidades ubicadas dentro de su área servida.

La Ley No. 7420 del 12 de julio de 1994, según publicación de la Gaceta No. 155 del 17 de agosto de 1994, prorroga la concesión de aguas No. 762-H para el aprovechamiento de fuerza hidráulica, otorgada por el Servicio Nacional de Electricidad a JASEC, la cual se utiliza en la ampliación del Sistema Hidroeléctrico Birrís operado por JASEC desde 1971.

El 30 de abril de 1998 se publica en la Gaceta No. 103, la Ley No. 7799 que autoriza a JASEC la suscripción de contratos de asociación empresarial con empresas nacionales o extranjeras, públicas o privadas, cuya finalidad sea el emprendimiento conjunto en el desarrollo y explotación tanto de las obras como de los servicios que presta. Asimismo, se amplía los servicios públicos que pueden ser prestados por la institución, tales como:

- Suministro de energía eléctrica en las etapas de generación, transmisión, distribución y comercialización.
- Suministro del servicio de acueducto y alcantarillado, que incluye agua potable, recolección, tratamiento y evacuación de aguas negras, aguas residuales y pluviales.
- Suministro de combustibles derivados de hidrocarburos, dentro de los que se incluyen: los derivados del petróleo, asfaltos, gas y naftas destinados al consumidor final.

Adicionalmente en la Ley No. 7799 de previa cita, se autorizó a JASEC para prestar los servicios de televisión por cable, así como los servicios que prestan por sí mismas las municipalidades del país, previo convenio refrendado por la Contraloría General de la República.

Mediante la Ley No. 8345 del 26 de febrero del 2003 publicada en la Gaceta No. 59 del 25 de marzo del 2003, Ley de Participación de las Cooperativas de Electrificación Rural y de las Empresas de Servicios Públicos Municipales en el Desarrollo Nacional, se establecen las siguientes facultades:

- Se otorga a JASEC competencia nacional para desarrollar y operar plantas de generación de energía eléctrica.

- Se faculta al Ministerio de Ambiente y Energía para autorizar a JASEC concesiones de agua para el aprovechamiento de la fuerza hidráulica en centrales hidroeléctricas que no excedan de 60 megavatios.
- Se autoriza a JASEC para suscribir convenios de alianza empresarial con asociaciones cooperativas, empresas públicas nacionales y municipales del Sistema Eléctrico nacional conducentes al desarrollo y la explotación conjunta de obras y servicios de generación eléctrica.
- Se otorga a JASEC la posibilidad de suscribir convenios de cooperación, inversión y operación conjunta con entidades y empresas públicas nacionales para la prestación de servicios complementarios de sus actividades.

Mediante oficio No. 11121 del 8 de agosto del 2006, la Contraloría General de la República autoriza a JASEC la compra directa del Sistema Hidroeléctrico Barro Morado con una potencia nominal total de 1.570 kW, ubicado en la provincia de Cartago.

Mediante oficio No. 7238 del 6 de junio del 2006, la Contraloría General de la República autoriza el convenio de alianza empresarial entre JASEC y el Instituto Costarricense de Electricidad para el desarrollo del Proyecto Hidroeléctrico Toro 3.

Mediante la aplicación de la Ley de Expropiaciones, JASEC adquiere el Sistema Hidroeléctrico Tuis en el Cantón de Turrialba de la provincia de Cartago.

El 8 de mayo del 2008, ley No. 8660 “Fortalecimiento y Modernización de las Entidades Públicas del Sector de Telecomunicaciones”, se faculta a JASEC para la prestación de los servicios de Telecomunicaciones, Infocomunicaciones, y otros servicios en convergencia, a la

luz de los dispuesto en la ley No. 8642 del 4 de junio del 2008 “Ley General de Telecomunicaciones”.

Misión

“Somos una empresa cartaginesa que busca mejorar la calidad de vida de nuestros clientes, brindando servicios de interés público de excelencia, a precios competitivos con solidaridad y al mayor número de personas, administrándonos con los mejores principios éticos, de desarrollo sostenible y responsabilidad social.”

Visión.

“En el año 2014 seremos:

1. En el ámbito externo:

Una empresa corporativa de servicios, líder, visionaria, con calidad internacional, que supera las expectativas de los clientes.

Nuestros esfuerzos y resultados facilitan la generación de nuevos y mejores instrumentos de progreso, contribuyendo con el desarrollo sostenible y la calidad de vida de la Sociedad Costarricense.

2. En el ámbito interno:

Una Organización ágil y efectiva, con servicios de calidad certificada, a precios competitivos, sostenibilidad ambiental y rendición de cuentas.”

Política de Calidad

“Cumplimos nuestra Misión mediante los siguientes valores:

- Visión de negocios: observamos nuestro trabajo a través de la óptica del Cliente.

- Honestidad: hacemos uso productivo de los recursos de la Empresa, velamos por su custodia y garantizamos la calidad de los servicios.
- Competencia técnica: conocemos con detalle los productos/servicios que comercializamos y la reglamentación que rige la actividad.
- Responsabilidad social: contribuimos con la comunidad costarricense de manera activa, sostenible y voluntaria, en los ámbitos sociales, económicos y ambientales."

Estructura Organizativa

JASEC está dividida en cinco áreas de negocio:

- Energía: Brinda servicios de generación, transmisión y distribución de Energía Eléctrica.
- Servicios Corporativos: Brinda servicios financieros y de tecnologías a los diferentes negocios.
- Acueductos y Saneamiento Ambiental: Brindará servicios distribución de agua potable, alcantarillado y recolección de basura.
- Infocomunicaciones: Brindará servicios de Internet, Telefonía y Televisión, mediante tecnología Fibra Hasta la Casa (FTTH).
- Unidad Ejecutora: Brinda servicios financieros de proyectos realizados en convenio con otras instituciones.

La figura 3 permite observar la estructura organizativa a un primer nivel.

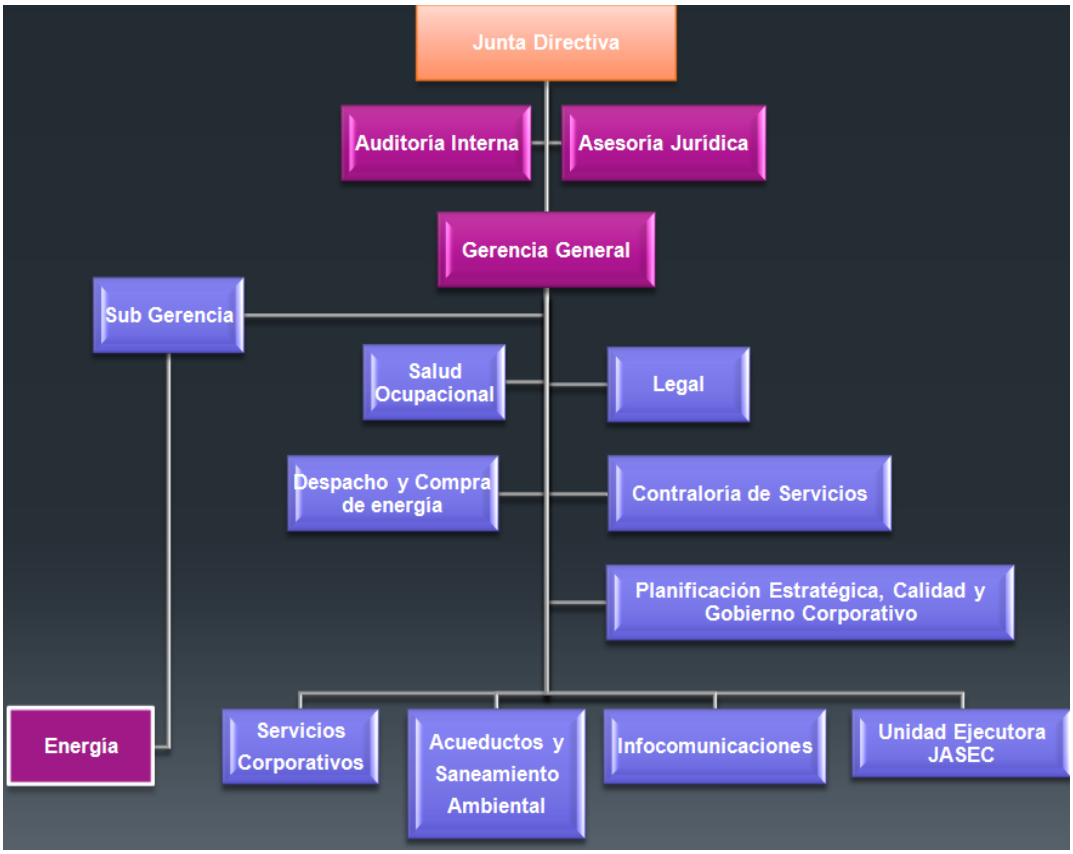


Figura 3. Organigrama JASEC – Nivel 1 (RH, 2013).

Cada negocio está dividido en seis Unidades Estratégicas de Negocios (UENs), y cada UEN tiene uno o más procesos o departamentos, las figuras 4, 5, 6 y 7 muestran las diferentes UENs con sus respectivos procesos de cada negocio.

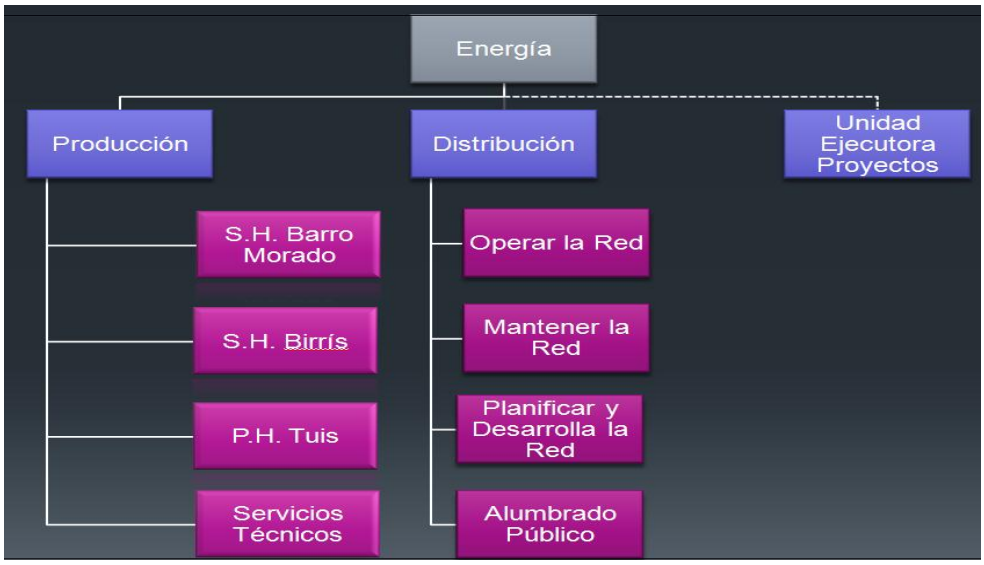


Figura 4. Organigrama JASEC – Negocio Energía (RH, 2013).

La figura 4 muestra que el negocio de Energía se compone de las UEN:

- Producción: Es la encargada del servicio de generación de energía eléctrica. Comprende desde la captación de recursos hídricos en las tomas, hasta la producción de energía en las plantas generadoras.
- Distribución: Tiene bajo su responsabilidad la subtransmisión, entrega y compra de energía, mediante el proceso Mantener y Controlar la red. El proceso (departamento) Planificar y Desarrollar la Red tiene como misión suministrar a los clientes de la red de transmisión y distribución de JASEC disponibilidad de acceso, tal que satisfaga los requerimientos de demanda y consumo de energía eléctrica a los distintos tipos de clientes.

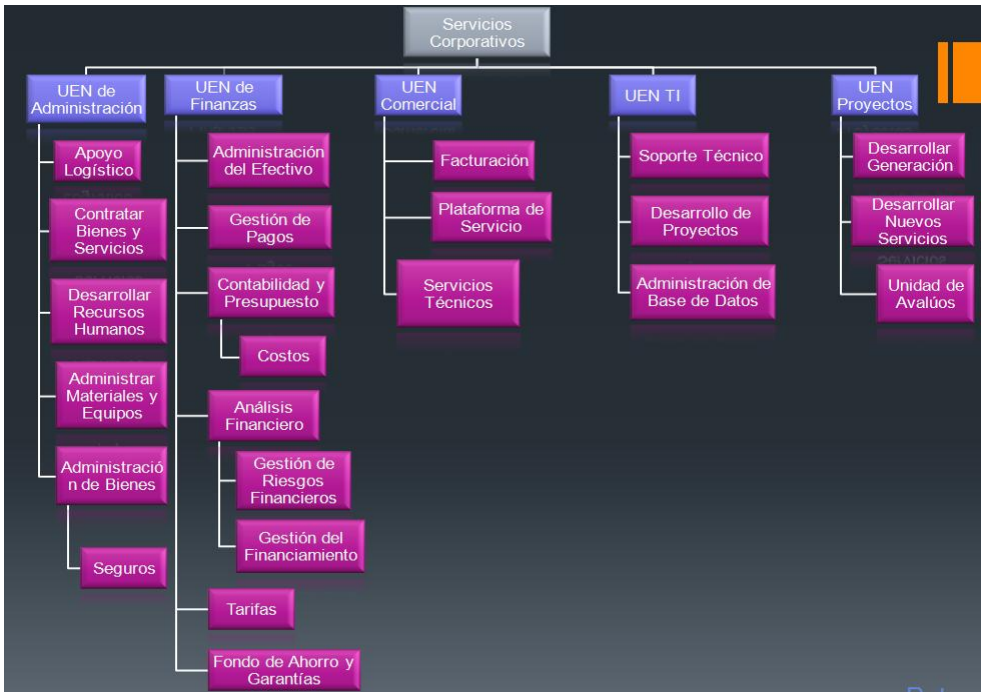


Figura 5. Organigrama JASEC – Negocio Servicios Corporativos (RH, 2013).

La figura 5 muestra que el Negocio de Servicios Corporativos se conforma de las UEN:

- Administración y Finanzas: Vela por una eficiente y efectiva gestión administrativa/financiera.
- Comercial: Se encarga de planificar, dirigir, coordinar, controlar y dar seguimiento a todas aquellas actividades orientadas al cliente.
- Tecnologías de Información: Planifica, dirige, coordina, da seguimiento y controla el desarrollo de los sistemas informáticos de la institución.
- Proyectos: Planifica, dirige, coordina, supervisa y controla la ejecución y formulación de proyectos tanto en el servicio de electricidad como en otros negocios.

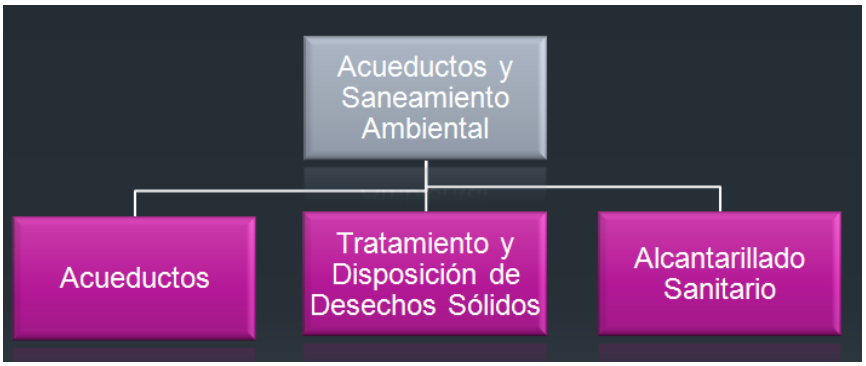


Figura 6. Organigrama JASEC – Negocio Acueductos y Saneamiento Ambiental (RH, 2013).

La figura 6 muestra que el negocio de Acueductos y Saneamiento Ambiental se compone de las UEN:

- Acueductos.
- Tratamiento y Disposición de Desechos Sólidos.
- Alcantarillado Sanitario.

Figura 7. Organigrama JASEC – Negocio Infocomunicaciones (RH, 2013).

La Figura 7 muestra que el negocio de Infocomunicaciones se compone de las UEN:

- Mercadeo.
- Control y Supervisión de Operaciones.
- Soporte de Operaciones TI.

Antecedentes

JASEC debe velar por una correcta planificación y ejecución de proyectos que garanticen el crecimiento, cobertura y optimización de la Red de Distribución Eléctrica, de tal forma que permita satisfacer la demanda actual y futura de los distintos tipos de clientes, mediante un proceso de planificación de corto, mediano y largo plazo.

En el Negocio de Energía, en la UEN de Distribución existe un proceso llamado Planificar y Desarrollar la Red, el cual es encargado de velar y cumplir la planificación y ejecución de proyectos.

Para garantizar una cobertura adecuada, JASEC ha realizado importantes inversiones en el área de Generación y Distribución eléctrica, tales como:

- Adquisición del Sistemas Hidroeléctricos
    - Tuis, ubicada en el cantón de Turrialba.
    - Barro Morado, ubicado en el cantón de Cartago.
- Alianza empresarial con el ICE para el desarrollo del Proyecto Hidroeléctrico Toro 3, ubicado en San Carlos y de la Subestación Tejar.
- Adquisición de servidores, estaciones de trabajo y equipos de comunicación.
- Adquisición de software que permita Supervisar, Controlar y Adquirir Datos en tiempo real (SCADA).
- Desarrollo e implementación de sistemas de información a la medida.

Todo ello en aras de contar con más generación propia y así poder garantizar la demanda actual y futura.

Se han realizado importantes esfuerzos para poder tener un buen control sobre el desarrollo y ejecución de nuevas obras de electrificación, tal como el desarrollo e implementación de un Sistema de Control de Obras. Este no cumple con las expectativas y necesidades actuales, por lo que han proliferado un alto "desarrollo" de sistemas llamados "Shadow Systems", tales como:

- Desarrollo de reportes a la medida.
- Hojas de cálculo, en donde el personal transcribe la información de diferentes reportes para analizar información sobre solicitudes de nuevas obras.

Justificación

La electricidad es un producto que, en términos generales, no puede ser almacenado, por lo que las empresas proveedoras, como el caso de JASEC, deben asegurar que las cantidades de energía disponible (producida y adquirida) satisfaga la demanda de manera continua.

Hoy en día hay nuevas necesidades de información que los líderes y coordinadores exigen para tomar decisiones, para lo cual se requiere una mayor cantidad de información a analizar y que esta sea confiable y oportuna.

La figura 8 muestra las proyecciones de Consumo de Energía para el año 2035, se espera que sea casi el triple de la demanda actual.

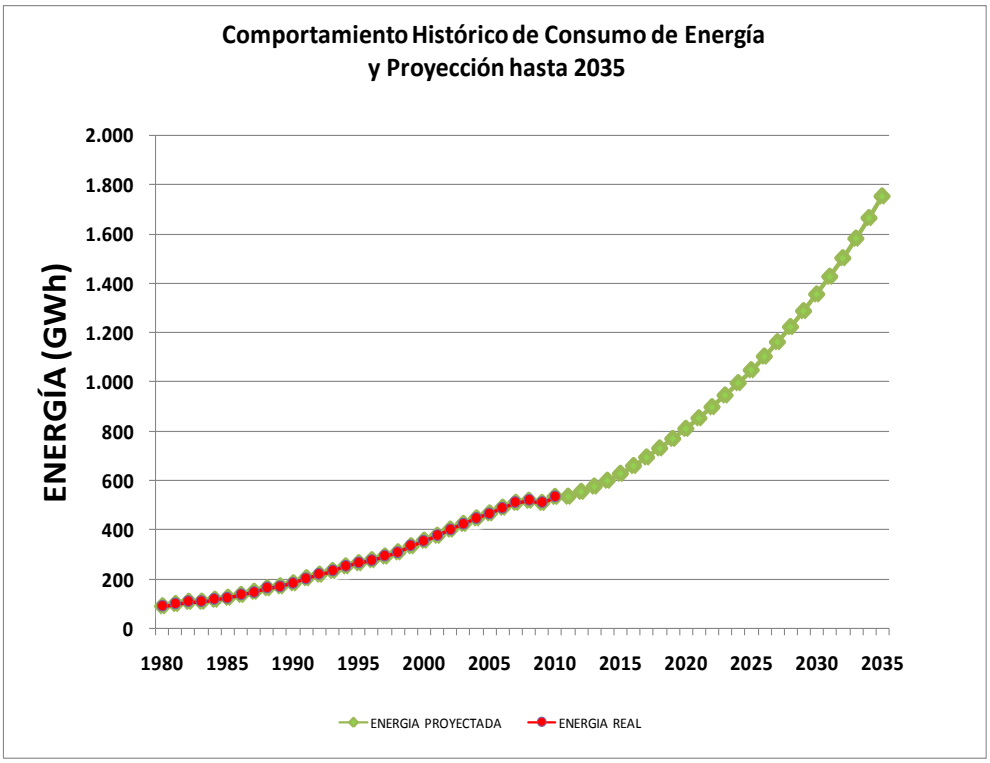


Figura 8. Comportamiento Histórico de Consumo de Energía hasta 2035 (PDR, 2012).

La figura 9 muestra que de los cinco cantones en los que la empresa presta el servicio, el cantón de Cartago es el de mayor consumo.

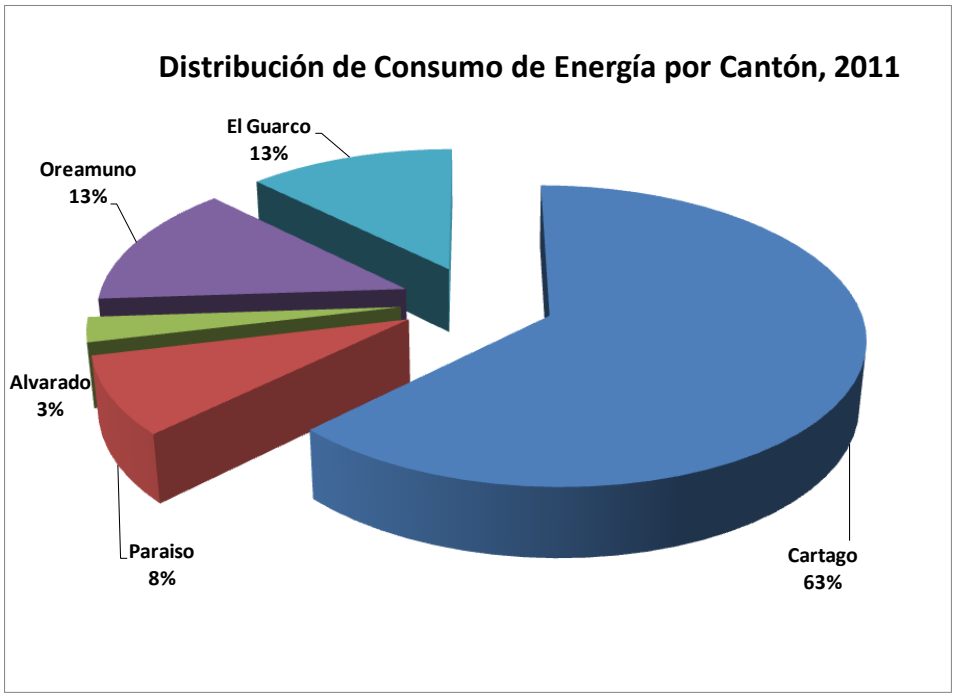


Figura 9. Distribución de Consumo de Energía por Cantón, 2011 (PDR, 2012).

En el año 2013 hay una tendencia de crecimiento en necesidades de consumo de energía, la figura 10 muestra un crecimiento en el sector urbanístico y comercial, esto debido a una marcada cantidad de solicitudes de Visado de Planos, Instalación de Bancos de Transformadores y extensión de Líneas Primarias y Secundarias.

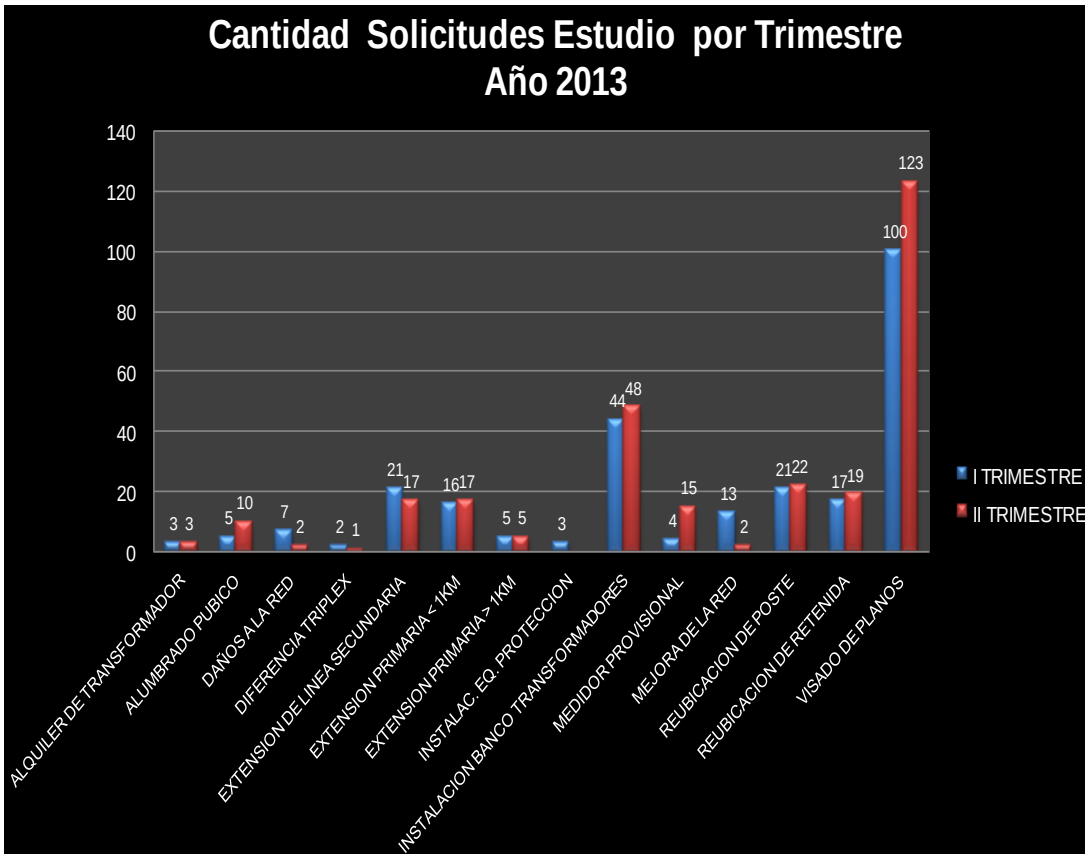


Figura 10. Cantidad de solicitudes de estudio, año 2013, primer y segundo trimestre (PDR, 2013).
Fuente: Departamento Planificar y Desarrollar la Red.

Se espera que para el año 2016 JASEC sea capaz de satisfacer en un alto porcentaje el consumo de energía con generación propia, como lo muestra la figura 11.

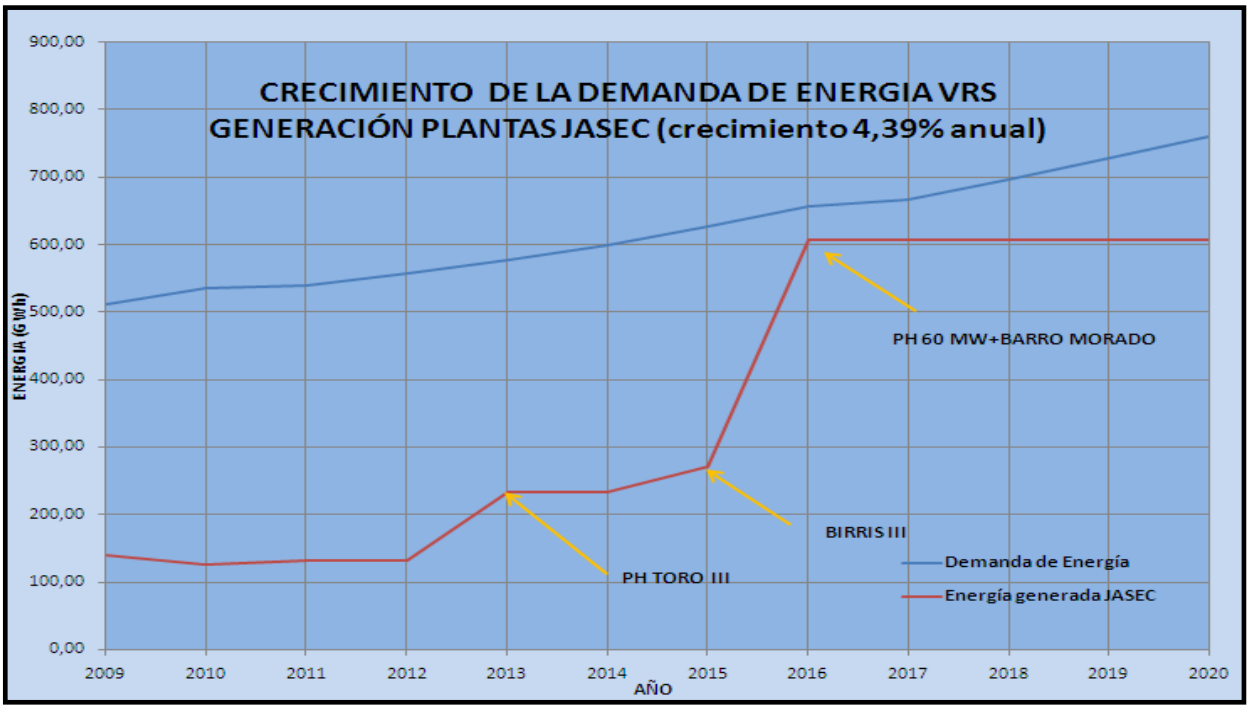


Figura 11. Crecimiento de la Demanda de Energía vs. Generación de plantas de JASEC (PDR, 2012).

Esto obliga a JASEC a realizar una buena planificación para satisfacer la demanda y suministro de energía eléctrica.

Día a día JASEC realiza esfuerzos para ser una empresa competitiva en el sector nacional, al buscar traer más inversión a la provincia de Cartago y brindar un buen servicio de energía, además de ser una de las empresas a nivel nacional que brinda las tarifas más bajas por consumo de KWh. La figura 12 muestra una comparación de tarifas residenciales entre las diferentes empresas de distribución eléctrica a nivel nacional, lo cual afirma que JASEC es la empresa con tarifa residencial más baja. Los datos fueron tomados de la Autoridad Reguladora de Servicios Públicos (ARESEP, 2013).

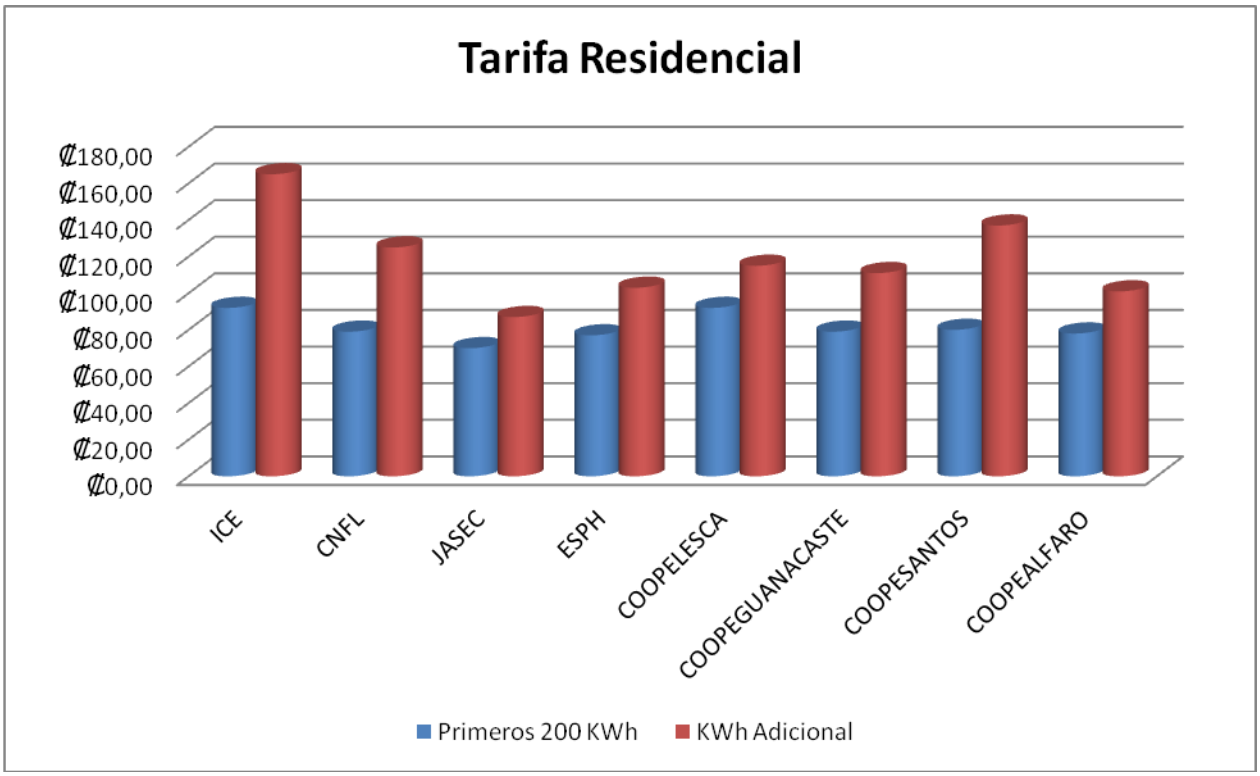


Figura 12. Tarifa Residencial (ARESEP, 2013)

Ante el panorama descrito, se puede afirmar que es necesario contar con información oportuna y confiable que permita tomar decisiones oportunas. La información deberá ir acompañada de un adecuado almacenamiento y custodia, así como un plan para lograr que la información fluya en forma oportuna a todas las unidades de la institución que la requieran. Es por ello que se propone el Diseño de un Data Mart para JASEC, modelado mediante UML.

Se utilizará UML como herramienta para modelar el diseño del Data Mart. Milicev (2009) indica que UML tiene varios inconvenientes, entre ellos:

- Complejo y Heterogéneo: Debido a que varios sectores incorporarán nuevos conceptos para poder modelar diferentes tipos de problemas.
- UML no cuenta aún con precisión semántica. Al estar dirigido a todas las fases de desarrollo de software permite modelos incompletos, inconsistentes e informales, ya

que toma en cuenta las fases iniciales del desarrollo de software, tal como la etapa requisitos, aun cuando los analistas solo tienen ideas vagas e incompletas del sistema que se está construyendo.

A pesar de sus inconvenientes tiene sus grandes ventajas, Milicev (2009) y Luján (2005) indican:

- Es un lenguaje común para desarrollar e intercambiar modelos de software, lo cual lo hace un lenguaje estandarizado y aceptado hoy en día.
- Puede ser utilizado en diferentes dominios de aplicativos (profiles), tales como sistemas de información, sistemas de telecomunicaciones, aplicaciones Web y sistemas empresariales, entre otras. De esta manera, analistas/diseñadores no necesitan entender en un 100% el UML, sino solamente que conozcan los elementos necesarios para la aplicación que desean modelar.
- Sigue un paradigma de orientación a objetos (OO), el cual se ha demostrado que es semánticamente más rico que otros paradigmas porque los modelos OO se acercan a la concepción que tiene el usuario final del problema.
- UML es ampliamente aceptado en el área de Tecnologías de Información. Hoy en día muchas herramientas de Ingeniería de Software Asistidas por Computadora (CASE) soportan UML, tales como IBM Rational Rose, Microsoft Visio y JDeveloper, entre otros.
- UML genera un lenguaje intermedio que facilita el entendimiento de requerimiento entre los interesados y los desarrolladores.

Para el modelamiento lógico y conceptual del Data Mart se utilizará la metodología propuesta por Luján (2005) llamada Proceso de Ingeniería para el Desarrollo de Bodega de Datos (Data Warehouse Engineering Process – DWEP), y basada en el Proceso Unificado (Unified Process UP). Esto debido a las ventajas que nos ofrece UP, entre ellas:

- Proceso iterativo e incremental.
- Basado en componentes.
- Se basa en UML para el modelar.

Planteamiento del Problema

Problema General

¿Cómo mejorar la obtención de información en forma oportuna y confiable que permita al personal del área de Planificar y Desarrollar la Red realizar una buena planificación y ejecución de proyectos que permita garantizar un adecuado crecimiento y cobertura de la red de distribución eléctrica?

Subproblemas

¿Cómo se realiza el proceso de planificación y ejecución de proyectos?

¿Cómo unificar datos de diferentes fuentes en un único repositorio?

¿De dónde se obtienen los diferentes datos para realizar el proceso de planificación?

¿Cómo mejorar la oportunidad de información para una adecuada toma de decisiones?

Objetivos

Existen una serie de taxonomías para la definición de objetivos, tales como:

- Bloom, 1956,
- Bloom Revisada, 2001,
- Marzano y Quendal, 2006,
- Era Digital de Churches, 2008.

Los objetivos se definieron con base en la taxonomía Bloom 1956, debido a que han pasado más de cincuenta años y al día de hoy continúa siendo la herramienta fundamental para establecer objetivos de aprendizaje en forma exitosa. Además, los verbos que utiliza son los adecuados para la definición de objetivos del proyecto.

General

Diseñar un Data Mart para el área de Planificar y Desarrollar la Red de la JASEC modelado mediante UML.

Objetivos Específicos

- Conocer el proceso actual de planificación y ejecución de proyectos.
- Definir los requisitos y funciones deseadas para el Data Mart.
- Revisar las fuentes de datos.
- Construir el modelo conceptual y lógico del Data Mart.

Alcances

- El trabajo se desarrollará en el área de Planificar y Desarrollar la Red, de la UEN de Distribución Eléctrica del Negocio de Energía de la JASEC.
- Abarcará el modelo conceptual y lógico de un Data Mart que permita al encargado respectivo tomar una decisión más acertada referente a la planificación y ejecución de nuevas obras eléctricas.
- La metodología a utilizar será el Proceso de Ingeniería para el Data Warehouse planteado en la tesis doctoral Luján-Mora (Sergio Luján-Mora, 2005), utilizando como herramientas de modelado al Lenguaje Unificado de Modelado (UML) citada en su tesis doctoral.

Limitaciones

El trabajo propuesto consiste en el modelo conceptual y lógico del Data Mart y no en su implementación. Además se verá limitado por la disponibilidad del personal del área de Planificar y Desarrollar la Red, el cual no tendrá un efecto negativo en el proyecto debido a que se cuenta con todo el respaldo del coordinador del área interesada.

Marco Teórico

Data Warehouse

Inmon (1992) indica "Un warehouse es una colección de datos orientado a temas, integrado, variable en el tiempo y no volátil para ayudar al proceso de toma de decisiones gerenciales."

Orientado temas: Datos que brindan información sobre un "sujeto" del negocio en particular, en lugar de concentrarse en la dinámica de las transacciones de la organización.

Integrados: Los datos con los que se nutre el DW vienen de diferentes fuentes y son integrados para dar una visión de un "todo" coherente.

Variante en el tiempo: Todos los datos en el DW son asociados con un periodo de tiempo específico.

No volátiles: Los datos son estables en el DW. Se pueden agregar más datos pero los existentes no se eliminan.

Kimball (1992) indica que el DW es "una copia de los datos transaccionales específicamente estructurados para consultas y análisis".

El DW y el Data Mart nacen debido a la necesidad de contar con información útil que permite apoyar a la toma de decisiones, dado que los datos operacionales no cumplen con este objetivo.

Data Mart

Ralph Kimball et al. (1998) indican "…los departamentos y divisiones van a crear sus propios Data Warehouses para responder a preguntas urgentes de negocios…"

Podríamos deducir que un Data Mart es como un DW específico que apunta a un área de negocio en particular (Compra, Ventas y Productos, entre otros.).

Inmon (1992) indica que existen dos tipos de Data Mart:

- Dependiente, el cual se alimenta de un DW corporativo. Es decir, desde un DW pueden emerger varios Data Marts, tal como lo muestra la figura 13.

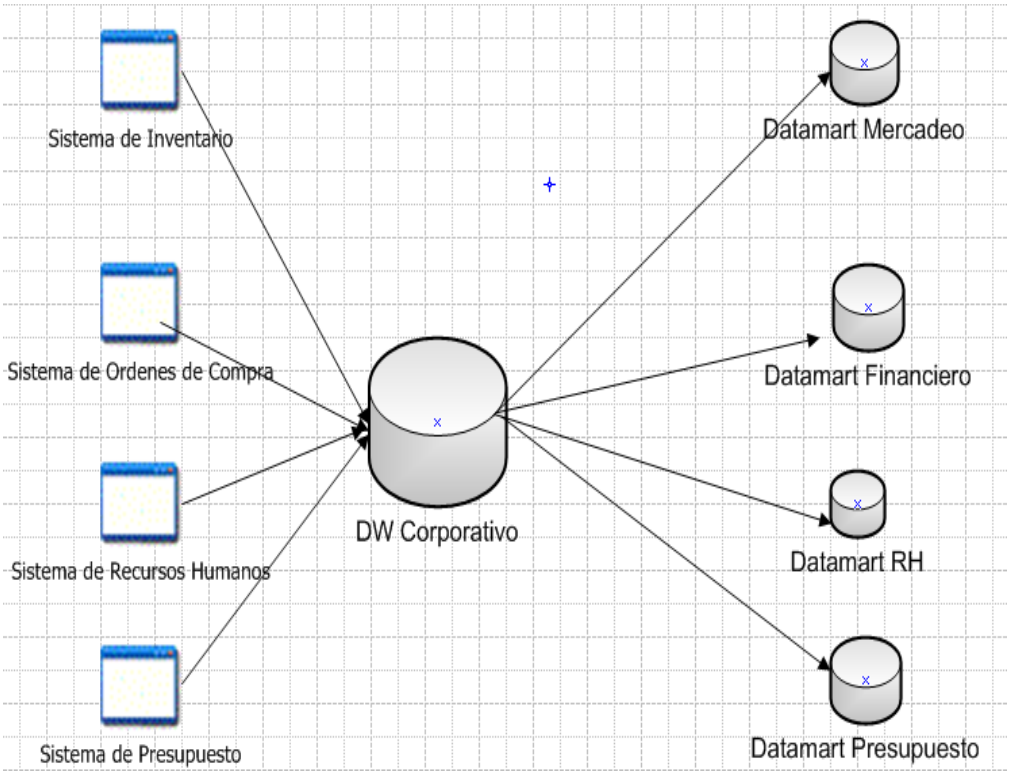


Figura 13. Data Mart Dependiente (Elaboración Propia).

- Independiente, el cual se alimenta de diferentes sistemas transaccionales, tal como lo muestra la figura 14.

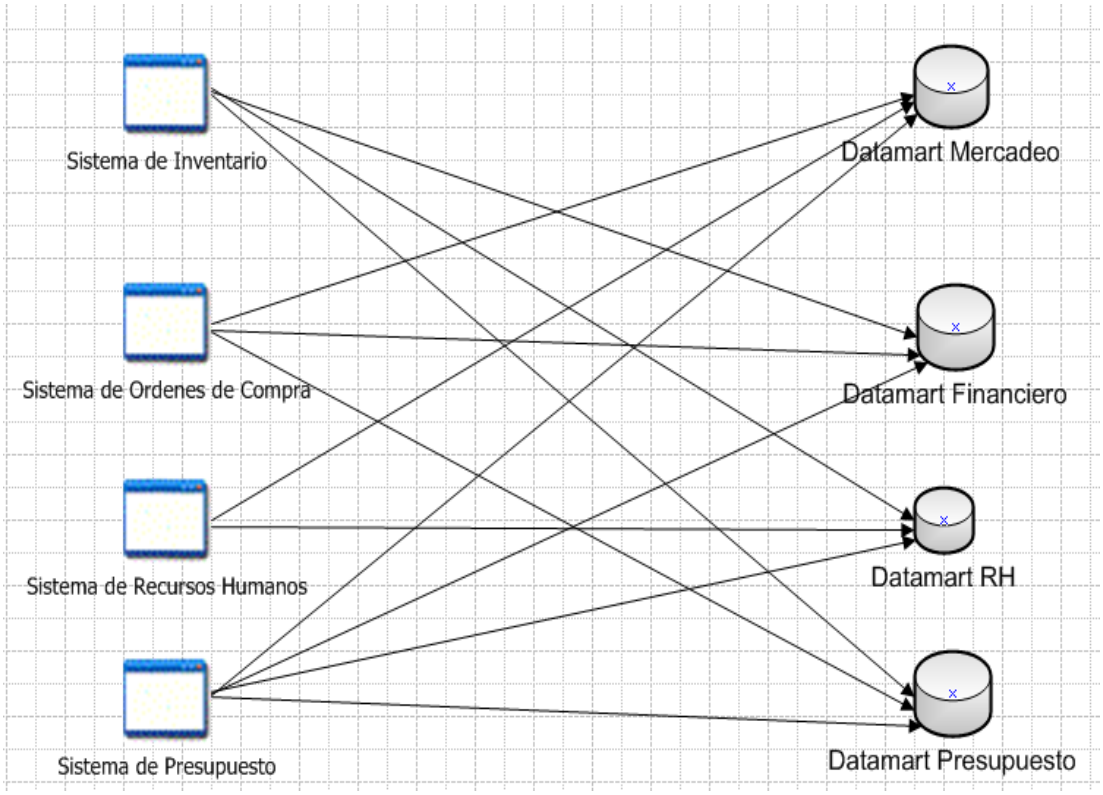


Figura 14. Data Mart independiente (Elaboración Propia).

Data Mart vs. Data Warehouse

Un DW que maneja información de distintas áreas típicamente es implementado como el repositorio central de información de toda una organización. El acceso a toda la información por parte de los usuarios de la empresa no es conveniente, ya que muchas veces necesita un subconjunto de esta información. En estos casos se utiliza un Data Mart. La tabla 1 nos muestra las diferencias que se dan entre un Data Warehouse y un Datamart.

Tabla 1. Diferencia DW y Data Mart (ORACLE 2007).

Categoría	Data Warehouse	Data Mart
Alcance	Corporativo	Línea de Negocios
Temas	Múltiples	Un solo tema
Fuentes de Datos	Muchas	Pocas
Tamaños	100 GB-TB+	< 100 GB
Tiempo de implementación	De meses a años	Meses

Se puede concluir que un DW como un Data Mart, es el proceso de extraer datos de las operaciones diarias de la empresa, procedente de diversos sistemas transaccionales, con el objetivo de integrarlos, sumarizarlos y almacenarlos en un único depósito de datos.

Sistemas de Extracción, Transformación y Carga (ETL)

Los sistemas ETL se encargan de una serie de funciones:

- Extracción de datos de las distintas fuentes de datos, ya sean provenientes de sistemas transaccionales o de fuentes externas.
- Transformación, son tareas de limpieza y consolidación de datos.
- Carga de los datos transformados al Data Warehouse o Datamart.

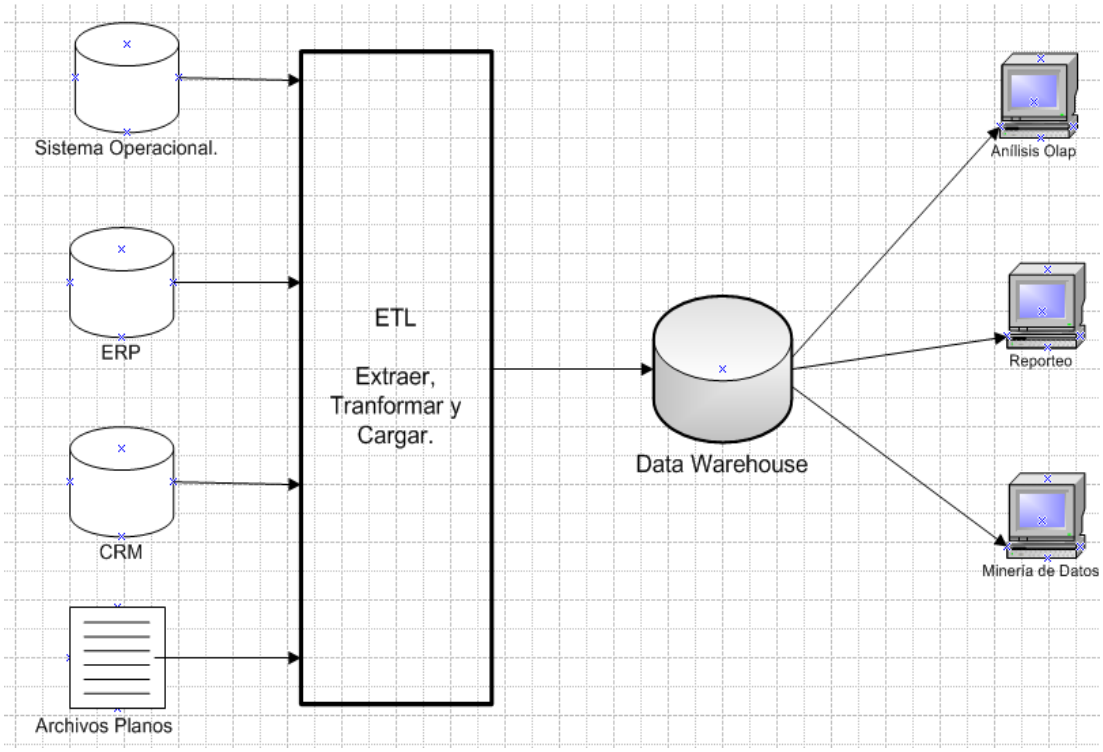


Figura 15. ETL - Extraer – Transformar – Cargar (Elaboración Propia).

La figura 15 muestra que el DW es alimentado por información obtenida de diferentes fuentes, tales como Sistemas Operacionales, Sistemas de Planificación de Recursos

Empresariales (ERP), Sistemas de administración de clientes (Customer Relationship Management – CRM), y/o archivos. Los datos son extraídos, transformados y cargados (ETL) al DW, con el objetivo de que herramientas de análisis, de reporteo y de minería de datos puedan explotar dicha información para una adecuada toma de decisiones.

Luján (2005) nos indica: "Es ampliamente reconocido que el diseño y mantenimiento de estos procesos ETL es un factor clave de éxito en los proyectos de Data Warehouse. Debido a la dificultad de diseñar y mantener este tipo de procesos, existe muy poca proliferación de herramientas de este tipo e igualmente respecto al modelado de estos procesos."

Modelo Multidimensional

Un modelo de datos es un sistema formal y abstracto que permite documentar y organizar la información para una adecuada comunicación entre el personal técnico y usuarios.

Un Data Warehouse se basa en un modelo de datos multidimensional, el cual permite modelar bases de datos en una forma simple y entendible para los usuarios finales, ya que presenta la información de una forma estándar e intuitiva, al permitir accesos a la información con un alto rendimiento.

El modelo multidimensional comprende dos grandes componentes:

- Tablas.
- Esquemas.

Tablas

En el modelo multidimensional se utilizan dos tipos de tablas:

- Tablas de hechos.

Adamson (2010) indica que son tablas que representan hechos, donde un hecho es una medida u atributo de tipo cuantitativo que es de interés para una empresa, como por ejemplo ventas totales del mes.

- Tablas de dimensiones.

Contienen atributos que le dan un contexto a un determinado hecho, son de tipo cualitativo.

Esquemas

En el modelo multidimensional existen dos tipos de esquemas:

- Estrella.

Se le llama estrella debido a que forma un diagrama en forma de estrella, que tiene en el centro de la estrella una tabla de hechos rodeado de tablas de dimensiones. En la figura 17 el modelo estrella consta una tabla de hechos llamada FactVentas y tres tablas de dimensión (DimProducto, DimCliente y DimFecha).

- Copo de nieve.

Utilizan un modelo de datos donde las dimensiones están normalizadas. En la figura 16 se nota una semejanza al modelo relacional, donde la dimensión de producto se relaciona con las marcas, y esta se relaciona con la subcategoría, categoría y departamento.

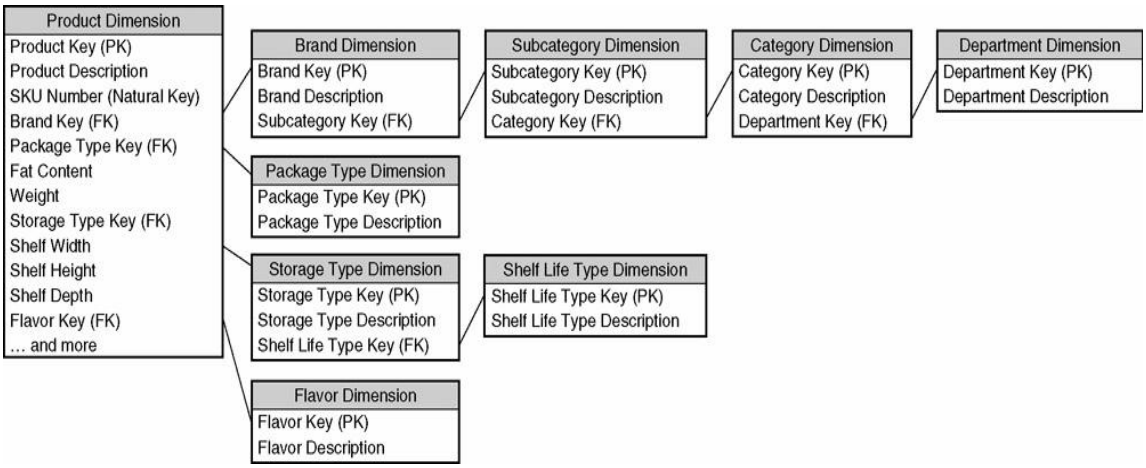


Figura 16. Esquema Copo de nieve (Kimball, 1996).

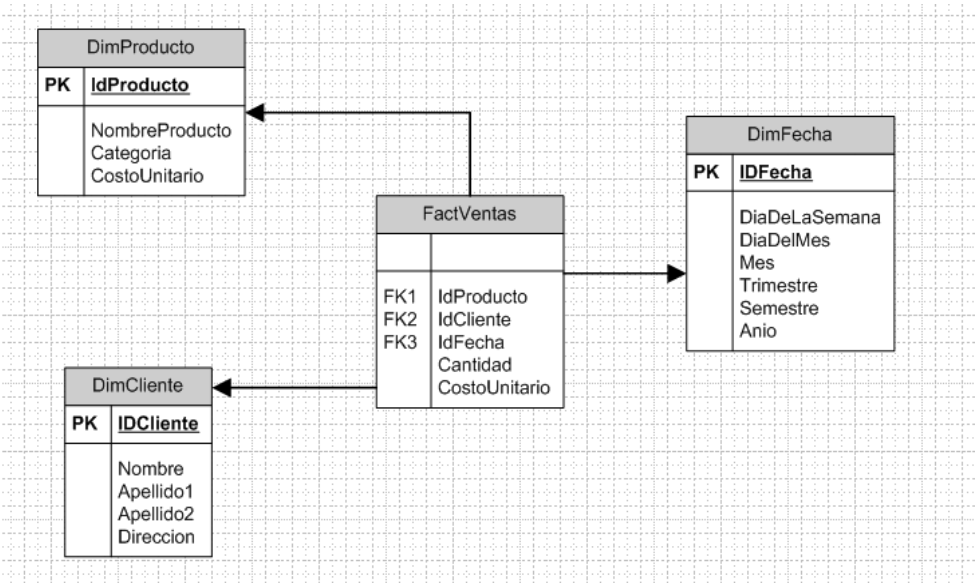


Figura 17. Esquema Estrella (Elaboración Propia).

Cubo OLAP

Adamson (2010) señala que los "Modelos dimensionales no siempre son implementados en una base de datos relacional. Una base de datos multidimensional (MDB) almacena información dimensional en un formato llamado cubo. El concepto básico detrás de un cubo es precalcular valores provenientes de varias tablas de dimensión y de hechos".

Un Cubo o Hipercubo es la estructura básica del modelo multidimensional, representa la información como una matriz multidimensional. En los ejes se encuentran los criterios de análisis y en los cruces están los valores a analizar.

La Figura 18 muestra una representación gráfica de un cubo de datos de ventas de diferentes productos en diferentes años y en diferentes ciudades.

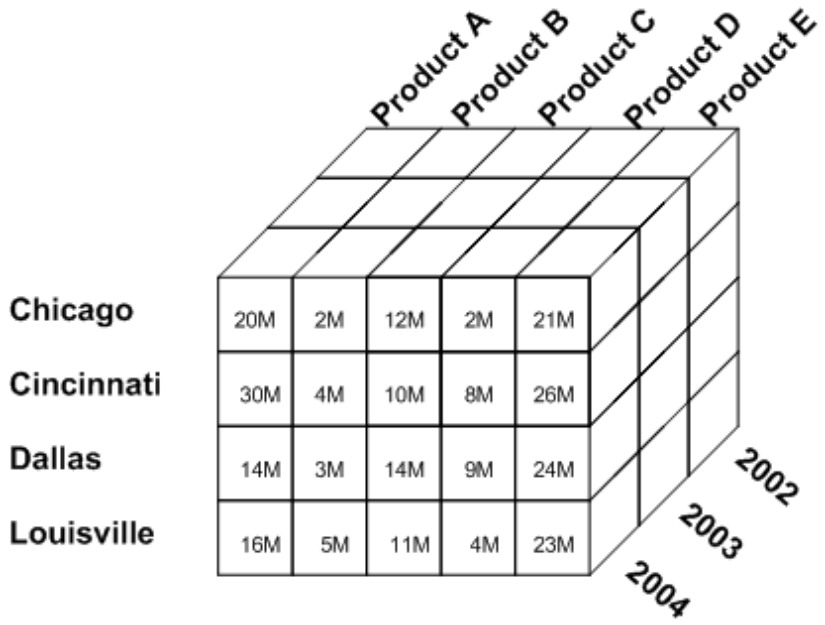


Figura 18. Representación gráfica de un cubo de ventas (ORACLE 2010).

Modelamiento de Datos

Modelo de datos

Oppel (2009) lo describe como el dato en un sistema de información que es representado y accesado. Son abstracciones de una base de datos propuesta o existente.

Modelamiento de datos

Kriegel (2011) lo describe como el proceso de crear un modelo de datos al aplicar descripciones de modelos de datos formales y usar técnicas de modelamiento de datos. Es decir, es el primer paso donde la abstracción del software se enfrenta al mundo real, es la hora en plasmar la idea en un sistema.

Existen tres niveles de modelamiento:

- Conceptual: Tiene un alto nivel de abstracción, representa las interacciones entre entidades (objetos físicos o conceptos abstractos, por ejemplo, libro, autor, cliente) y relaciones (por ejemplo, un autor puede tener varios libros). Se acerca más al problema que a la solución y es en donde el usuario representa sus ideas, y puede ser implementado en cualquier base de datos, ya que es independiente de la tecnología. El que realiza este tipo de modelamiento se le llama Modelador de Datos.
- Lógico: Se piensa en atributos de una entidad. Se utiliza un modelo relacional para el diseño de tablas, hechos, dimensiones y su interacción, además de la definición del proceso ETL, y el mapeo entre la fuente de datos y el DW. Consiste en trasladar o mapear el diseño conceptual a un diseño lógico, el cual encaja en un modelo de base de datos. El que realiza este proceso es el Diseñador de Base de Datos o un Administrador de Base de Datos (DBA).

- Físico: Es en donde sus ideas abstractas se concretan. Se escribe código para generar objetos de la base de datos tales como tablas de hechos, dimensiones y reglas de integridad, entre otros. Consiste en mapear el diseño lógico en uno o más diseños físicos, cada uno orientado a un determinado motor de base de datos. La persona encargada de su diseño frecuentemente es el DBA.

El modelo de datos consta de una serie de pasos, los cuales dependiendo del nivel de modelamiento se utilizan o no se utilizan, tal como lo muestra la tabla 2.

Tabla 2. Desarrollando modelos de datos (Kriegel, 2011).

Pasos	Conceptual	Lógico	Físico
Entidades	X		
Relaciones	X	X	
Atributos		X	
Llaves primarias		X	X
Llaves Foráneas		X	X
Tablas/Vistas			X
Columnas			X
Tipos de datos			X

Lenguaje de Modelado Unificado (U.M.L.)

El Object Management Group (OMG, 2011) lo define como un lenguaje que ayuda a especificar, visualizar y documentar modelos de sistemas informáticos, incluyendo su estructura y diseño, de manera tal que cumpla con todos los requisitos. La versión más reciente del UML es la 2.4.1, publicada en el mes de agosto del 2011.

UML es un lenguaje de modelado visual de propósito general orientado a objetos que sirve para visualizar, especificar, construir y documentar, además de ser independiente de cualquier fabricante comercial. Ofrece un vocabulario y reglas que permiten:

- Crear y leer modelos.
- Construir planos de un sistema.

UML maneja varios tipos de diagrama, los cuales se agrupan en:

- Diagramas de estructura: Muestran la estructura estática del sistema, sus partes de abstracción y los diferentes niveles de implementación, además de sus interrelaciones. Sus elementos representan los conceptos más importantes de un sistema y pueden incluir niveles de abstracción del mundo real y los conceptos de la aplicación.
- Diagramas de comportamiento: Permiten observar el comportamiento dinámico de los objetos en un sistema, el cual puede describirse como una serie de cambios en el sistema a través del tiempo.

En el apéndice 3 se mencionan los diferentes tipos de diagramas utilizados por UML.

Métodos de Diseño

Kimball (1996) presenta varios casos de diseño en los que se aplica un esquema estrella (star schema), con sus variantes tal como copo de nieve y constelación de hechos (snowflake, fact constellation), para realizar un modelado multidimensional. Además nos propone una matriz de arquitectura de BUS para integrar el diseño de varias estrellas y conseguir así un almacén de datos corporativo y global.

Todos los modelos en mención son de gran importancia y son grandes referentes para el modelado multidimensional, pero podrían carecer de un método formal para el diseño de DW ya que el esquema estrella es muy orientado al diseño de estructuras de base de datos, el BUS indica la integración entre varias estrellas, ambos dejan de lado los procesos de requerimientos y ETL, entre otros.

Una metodología muy utilizada para el diseño de un DW es el propuesto por Kimball (1998), el cual presenta el ciclo de vida de un proyecto DW para el cual se proponen distintas herramientas y técnicas a seguir, sin que se proponga un método o modelo global y formal a lo largo de todo el proceso. La figura 19 muestra las diferentes tareas de la metodología Kimball.

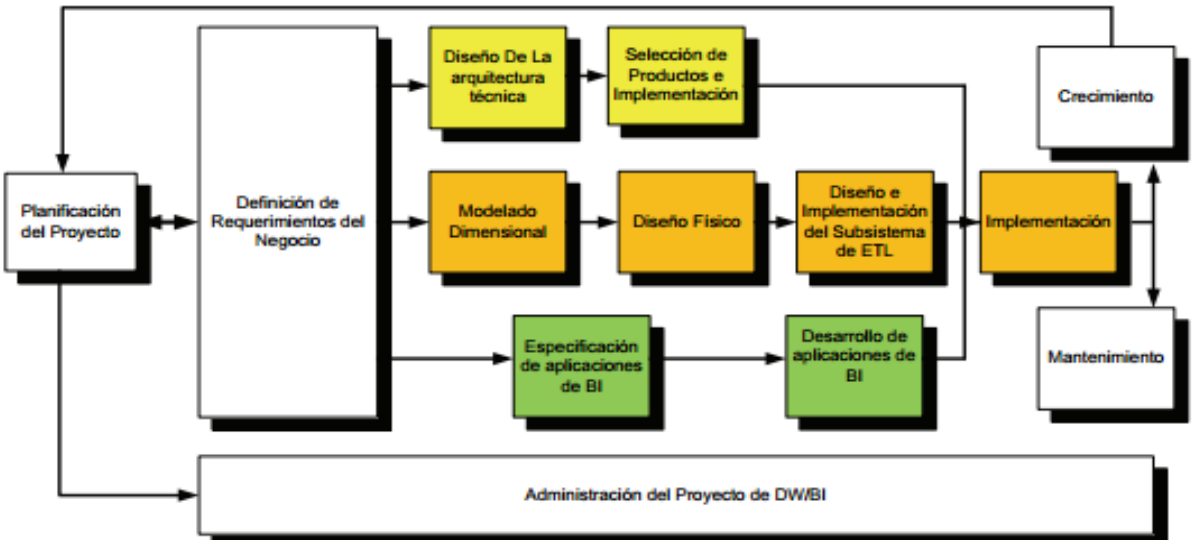


Figura 19. Tareas de la metodología de Kimball (Kimball, 1996)

El Centro de Investigación en Inteligencia de Negocios (2013) realizó un foro de Business Intelligence, en donde indican: "En términos de una metodología para BI, no existe una estándar 100% efectivo para generar proyectos, sino más bien, se importan una serie de diferentes metodologías de otros ámbitos, que pueden adaptarse según el problema en particular. Queda

abierta la discusión en torno a cuál o cuáles son las mejores prácticas que deberían ser incluidas en una metodología básica."

Proceso de ingeniería para el desarrollo de un DW (DWEP)

Se hace referencia a la metodología de "El proceso de ingeniería para la construcción de un Data Warehouse" y las extensiones a UML de "Diagramas del Data Warehouse" propuesto en la tesis doctoral "Diseño de Data Warehouse con UML" (Sergio Luján-Mora, 2005). En ella se presenta un Proceso de Ingeniería para el Desarrollo de DW (DWEP) basado en UP, el cual nos permite modelar (lógico, conceptual y físico) un DW en todos sus aspectos fundamentales.

DWEP plantea que el desarrollo de un Datamart se divide en pequeños pasos (iteraciones), lo cual nos permite un mejor control y un mejor desarrollo del proyecto. DWEP divide el desarrollo de un DW en cuatro fases: Concepción, Elaboración, Construcción y Transición. Cada fase tiene una serie de objetivos y entregables que son propios de un proyecto de DW.

Fase de Concepción: Consiste en analizar el proyecto para justificar su desarrollo e implementación. Se describe el proyecto, se detectan riesgos y se establece la funcionalidad básica.

Fase de Elaboración: Se forma una arquitectura para construir el software, estableciendo la base lógica de la aplicación con los casos de uso definitivos y los artefactos del sistema que lo componen. En él se efectúa una primera iteración con un prototipo funcional.

Fase de Construcción: Su fin es el desarrollo del sistema final en un entorno del usuario final.

Fase de Transición: Su fin es implantar el sistema en un entorno operativo.

Además de tener fases, DWEP propone una serie de flujos de trabajo. Un flujo de trabajo es un conjunto de actividades realizadas en un área determinada cuyo resultado es la construcción de artefactos, como por ejemplo un diagrama.

Dentro de los flujos de trabajo utilizados están: Requerimientos, Análisis, Diseño, Implementación, Pruebas, Mantenimiento y Revisión postdesarrollo. En cada flujo de trabajo se crearán una serie de entregables basados en diagramas UML.

Requerimientos: Se enfoca a las necesidades del usuario final. Durante este flujo de trabajo el usuario debe especificar qué tipo de análisis y agregaciones le serán de utilidad para generar reportes que le ayuden a la toma de decisiones.

Análisis: Define y analiza la especificación generada en los Requerimientos. Se analizan los datos que servirán como fuente de información del Datamart.

Diseño: Se define la estructura del Datamart tanto a nivel lógico como físico. Se mostrará la relación de cómo los datos fuentes se relacionan con los datos del modelo multidimensional.

Implementación: En este flujo de trabajo se construye el Datamart, se construyen sus estructuras físicas, se desarrollan los procesos ETL.

Pruebas: El objetivo de este flujo de trabajo es verificar que la implementación funcione de acuerdo con lo deseado. Más específicamente, el propósito de las pruebas es:

- Planear las pruebas requeridas.
- Diseñar y realizar pruebas a través de los casos de prueba requeridos.
- Ejecutar las pruebas y analizar los resultados de las mismas.

Mantenimiento: A diferencia de muchos sistemas, el DW o Datamart nunca está terminado. El objetivo principal de este flujo de trabajo es el de mantener los proceso de carga y actualización necesarios para que la información del DW esté actualizada. Este flujo de trabajo comienza cuando el DW o Datamart ha sido construido y entregado a los usuarios finales, pero no tiene una fecha de fin. Durante este flujo de trabajo puede ocurrir que los usuarios generen nuevos requerimientos como nuevas consultas, lo cual genera una nueva iteración.

Revisión Postdesarrollo: Esta tarea no forma parte del flujo de trabajo en el desarrollo del DW, pero es una revisión de los procesos para realizar mejoras en futuros proyectos. Se debe mirar y revisar el DW, la documentación creada, y tratar de identificar oportunidades de mejora y éxito para tomar en cuenta. Si durante el proceso se realizó un registro de los tiempos y esfuerzo utilizado, esta puede ser una fuente de información muy útil para futuros proyectos.

La figura 20 muestra cómo los flujos se relacionan con las fases DWEP.

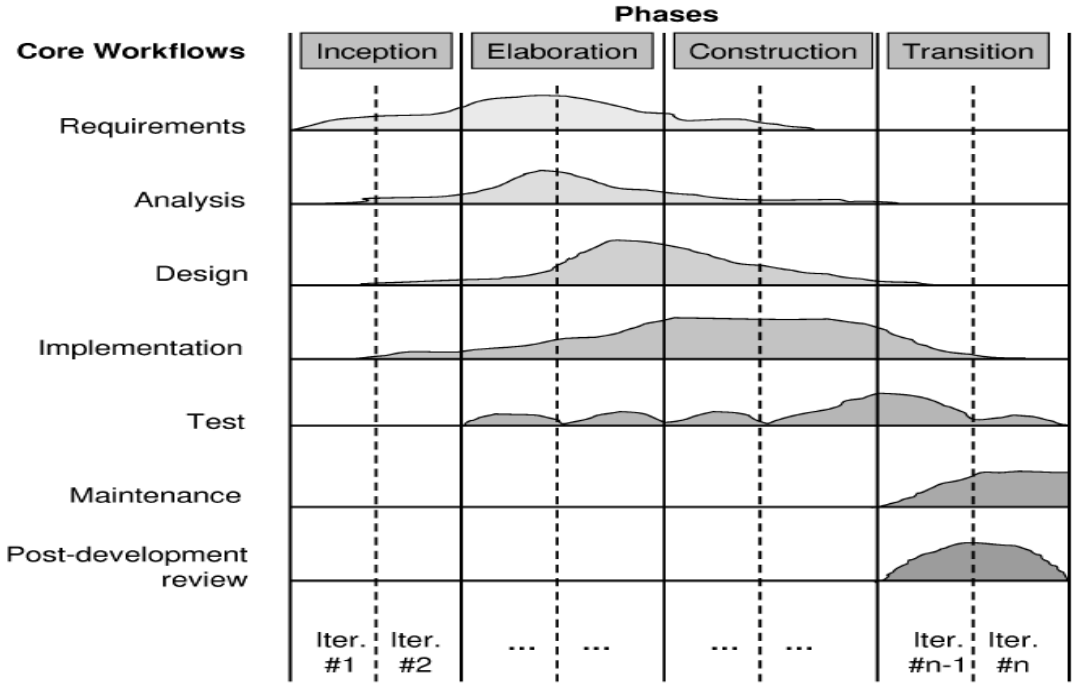


Figura 20. Marco de diseño de un Data Warehouse (Luján, 2005).

Como se muestra en la tabla 3, DWEP emplea 15 artefactos, basados en diagrama de casos de uso, diagrama de clases, diagrama de componentes y diagrama de despliegue.

Tabla 3. Artefactos empelados en DWEP (Luján, 2005).

	Source (S)	Integration	Data Warehouse (DW)	Customization	Client (C)
Conceptual	**SCS** Class diagram *Standard UML*	**DM** Class diagram *Data Mapping Profile*	**DWCS** Class diagram *Standard UML* *Multidimensional Profile*	**DM** Class diagram *Data Mapping Profile*	**CCS** Class diagram *Standard UML* *Multidimensional Profile*
Logical	**SLS** Class diagram *Different data modeling profiles*	**ETL Process** Class diagram *ETL Profile*	**DWLS** Class diagram *Different data modeling profiles*	**Exporting Process** Class diagram *ETL Profile*	**CLS** Class diagram *Different data modeling profiles*
Physical	**SPS** Comp. & deploy. diagrams *Database Deployment Profile*	**Transportation Diagram** Deployment diagram *Database Deployment Profile*	**DWPS** Comp. & deploy. diagrams *Database Deployment Profile*	**Transportation Diagram** Deployment diagram *Database Deployment Profile*	**CPS** Comp. & deploy. diagrams *Database Deployment Profile*

LEGEND: CS: Conceptual Schema, LS: Logical Schema, PS: Physical Schema, Comp. & deploy: Component and deployment

Para la mayoría de flujo de trabajo se puede utilizar diferentes diagramas UML, la tabla 4 se muestra el flujo de trabajo de requerimientos, análisis, diseño e implementación con los posibles diagramas UML o técnicas a utilizar.

Tabla 4. Flujo de trabajo y diagramas (Elaboración Propia).

Flujo de trabajo	Diagrama UML o Técnica.
Requerimientos	Diagramas de caso de uso.
Análisis	Source Conceptual Schema (SCS), Source Logical Schema (SLC) y Source Physical Schema (SPS)
Diseño	Data Warehouse Conceptual Schema (DWCS), Client Conceptual Schema (CCS) y Data Mapping (DM).
Implementación	Data Warehouse Logical Schema (DWLS), Data Warehouse Physical Schema (DWPS), Client Logical Schema (CCLS), Client Physical Schema (CPS), ETL Process, Exportation Process y el Transportation Diagram.

El diagrama de caso de uso nos permite un modelamiento visual. Cada caso de uso podría constar de:

- Nombre.
- Identificador.
- Actores involucrados.
- Precondiciones.
- Flujo de eventos que deben realizarse.
- Postcondiciones.

En la figura 21 se muestra un diagrama SLS (Source Logical Schema) de un sistema transaccional que administra las ventas, que muestra las tablas y sus relaciones.

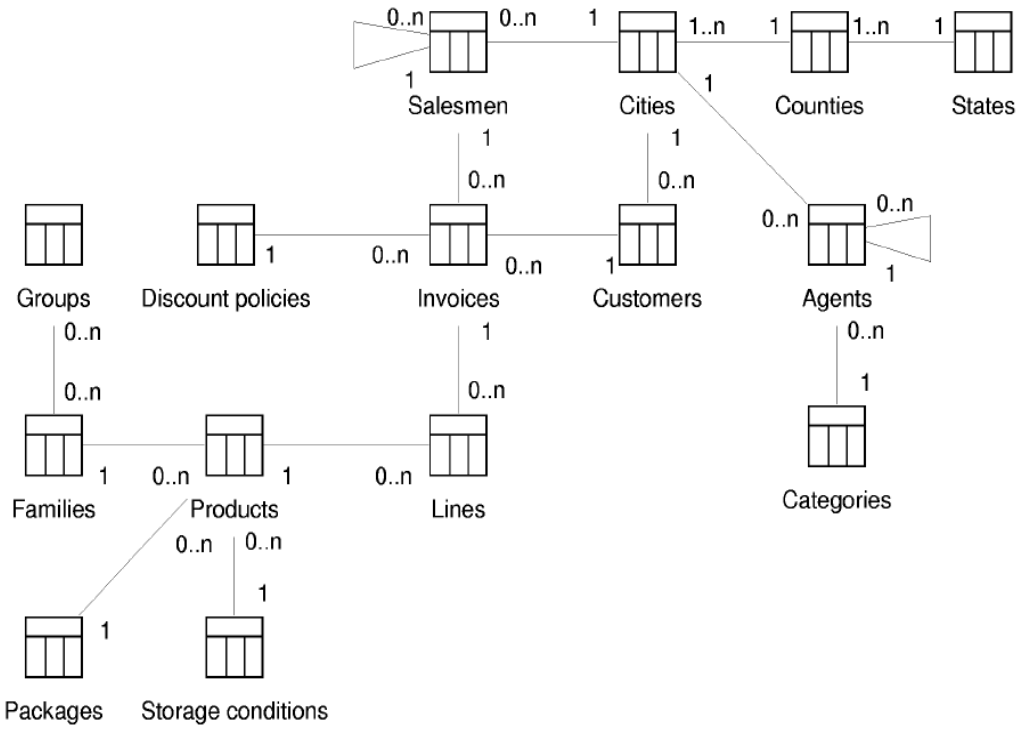


Figura 21. Source Logical Schema (Luján, 2005).

En la figura 22 se muestra un diagrama DWCS (Data Warehouse Conceptual Schema) de nivel 1, el cual se compone de tres esquemas, donde cada uno es un modelo estrella.

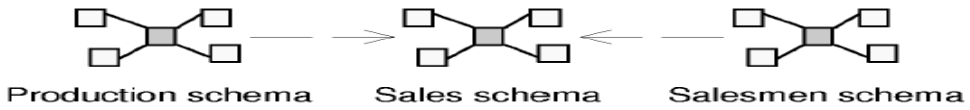


Figura 22. DWCS nivel 1 (Luján, 2005).

La figura 23 muestra el diagrama DWCS nivel 2 del esquema de ventas, compuesto de una tabla de hechos y cuatro tablas de dimensión.

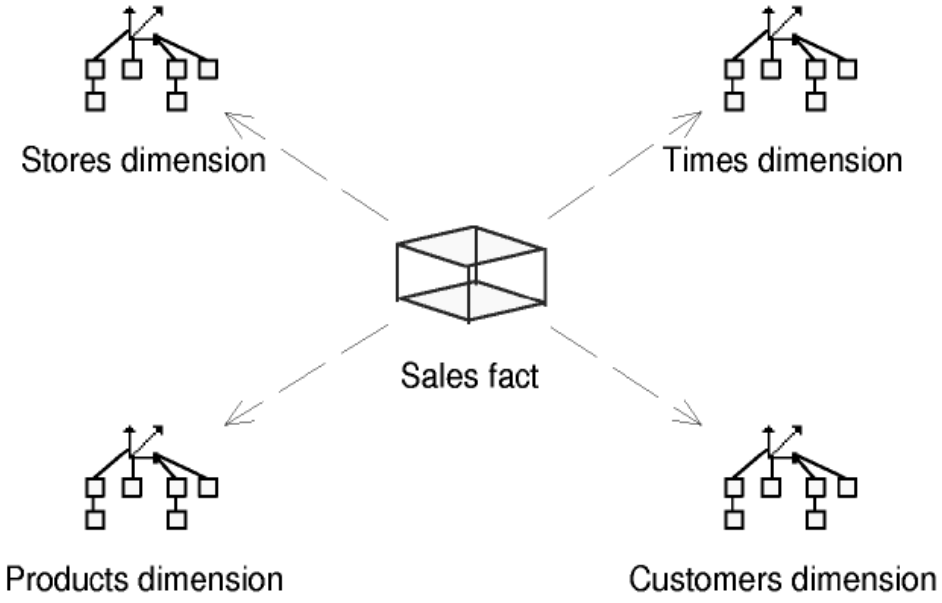


Figura 23. DWCS nivel 2 (Luján, 2005).

La figura 24 muestra diagrama DWCS nivel 3, donde se muestra la definición de la dimensión de Clientes.

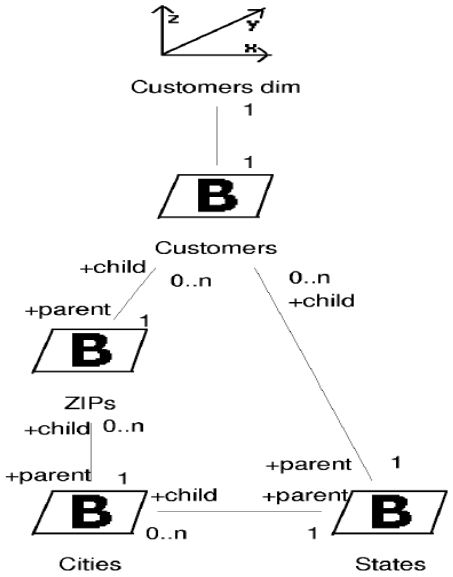


Figura 24. DWCS nivel 3 (Luján, 2005).

La figura 25 muestra el diagrama DWPS (Data Warehouse Physical Schema).

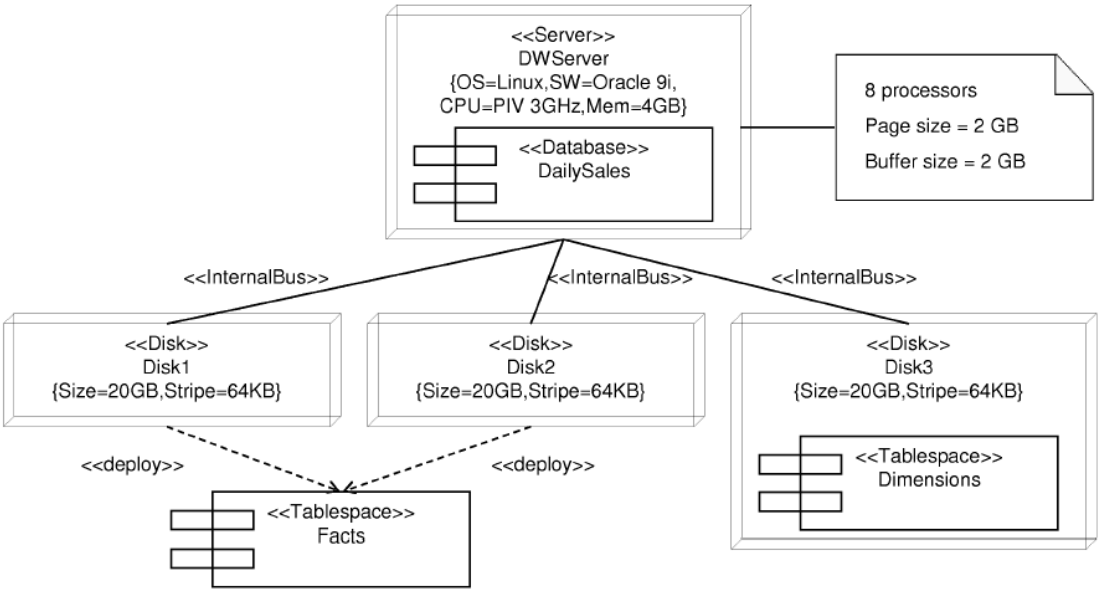


Figura 25. DWPS (Luján, 2005).

DWEP esquematiza el desarrollo de un Datamart en Esquema Conceptual, Lógico y Físico. En la figura 26 se muestra una relación entre ellos:

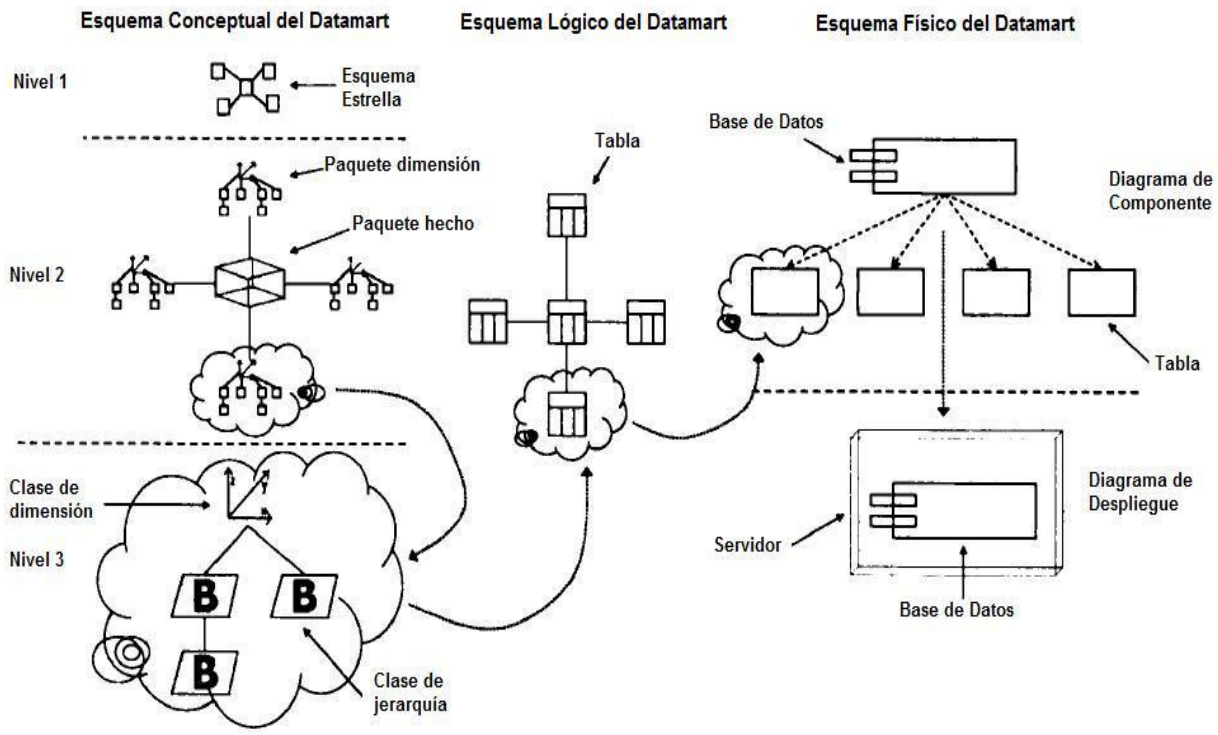


Figura 26. Relación entre los Esquemas del DataMart (Zambrano, 2011).

Como se observa en la figura 26, el Esquema Conceptual consta de 3 niveles. El nivel 1 define el modelo multidimensional. El nivel 2 especifica cada esquema del nivel 1 y muestra las dimensiones que se relacionan con cada hecho. Por último, el nivel 3 especifica cada dimensión.

Marco Metodológico

Tipo de investigación o proyecto

El proyecto se enmarca dentro del enfoque cualitativo ya que pretendemos comprender y conocer un proceso y una realidad dada en el área de Planificar y Desarrollar la Red en un entorno real. Para conceptualizar este enfoque, Hernández, Fernández & Baptista (2010) señalan:

- Para la generación de la idea existe una realidad en la mente de los involucrados, que el investigador desea descubrir, comprender e interpretar.
- Para definir el objetivo se intenta descubrir, comprender e interpretar sus fenómenos, a partir de percepciones y significados de los participantes.
- Para el diseño de la investigación, se inicia con un bosquejo de un método, pero este debe ser flexible y adaptable en el proceso. Se utilizan diversas técnicas para recolectar datos, tales como observación, entrevistas e historias de vida, entre otros.
- Para el análisis de resultados, se trata de comprender a las personas.
- Para la presentación de resultados, se pueden utilizar narraciones, fragmentos de textos, videos, audios, fotografías, mapas.

En cuanto al alcance, se define como descriptivo, pues pretende describir el proceso de planear y ejecutar nuevas obras eléctricas.

El diseño de este proyecto es de Investigación Acción, ya que tratará de mejorar el proceso de Planificar y Desarrollar la Red, mediante el diseño de un Data Mart. Elliott (1993) indica:

"El objetivo prioritario de la investigación-acción consiste en mejorar la práctica en vez de generar conocimiento, así, la producción y utilización del conocimiento se subordina a este objetivo fundamental y está condicionado por él".

Sujetos y fuentes de información

Fuentes de Información

Las principales fuentes de información utilizadas para este proyecto son:

- Documentación brindada por la JASEC concerniente al marco jurídico.
- Documentación brindada por el área de Planificar y Desarrollar la Red de JASEC.
- Entrevistas con personal de planta del área de Planificar y Desarrollar la Red de JASEC.
- Documentos, artículos, libros y material que brinde una guía y apoyo en lo referente a temas de DW y UML.
- Libros de UML como referencias para el uso de UML.
- Tesis relacionadas con Data Warehouse.
- Guías para el desarrollo de proyectos de graduación.

Sujetos de Información

- Con respecto a los sujetos de información para el proyecto se han realizado investigaciones en la página web de JASEC, mediante las cuales se ha podido obtener información histórica y organizacional de la institución.

- La información referente a la planificación y desarrollo de nuevas obras eléctricas ha sido proporcionada por personas relacionadas al área de Planificar y Desarrollar la Red.
- Preliminarmente las entrevistas se dieron con el coordinador del área de Planificar y Desarrollar la Red, Lic. Roger Carrillo, con quien se trató las necesidades del proyecto.
- La mayor parte de entrevistas se dieron con el personal del área de Planificar y Desarrollar la Red, con las cuales se aclararon dudas, se delimitó el proyecto, se obtuvieron opiniones con respecto a la importancia de desarrollar el proyecto, además de los beneficios que obtendrán una vez finalizado.

Descripción de Instrumentos

Una investigación de estudio cualitativo utiliza un conjunto de técnicas e instrumentos propios de su paradigma, tales como entrevistas, relatos y grupos de discusión, entre otros.

Para poder estudiar el despacho de energía, se utilizaron los siguientes instrumentos:

- Entrevistas.
    - Nos permita recolectar información en una forma más clara y específica acerca de las vivencias personales.
- Observación.
    - Nos permitirá conocer el proceso
- Grupos de discusión.

- Nos permitirá identificar necesidades, comprender y a su vez mejorar y valorar la eficacia del proceso.

La tabla 5 muestra en qué tareas del proyecto se utilizará un instrumento en específico.

Tabla 5. Tareas-Instrumento (Elaboración Propia).

Tarea	Instrumento
DISEÑO DATA MART	
MEJORA AL DOCUMENTO DEL PROYECTO	
Entregable 1	
Revisión y Mejora al Marco Organizacional.	
Revisión y Mejora al Marco Teórico.	
Revisión y Mejora al Marco Metodológico.	
Entregable 2.	
ANÁLISIS PROCESO	
Proceso Actual	Observación, Entrevista, Grupo de discusión.
Entregable 3.	
Revisión y Mejora del análisis de proceso	
ELABORACIÓN	
Esquema conceptual de la fuente de datos.	Entrevista.
Esquema lógico de la fuente de datos.	Entrevista
Entregable 4.	
Revisión y mejora	
Esquema conceptual del DM.	Entrevista
Diagrama de mapeo de datos.	Entrevista
Entregable 4.	
Revisión y mejora	
CONCLUSIONES Y RECOMENDACIONES	
Conclusiones	
Recomendaciones.	
Revisión final de tutor.	
Mejora	

Metodología de Desarrollo

Definición de la Metodología

Se eligió la metodología llamada Data Warehouse Engineering Process (DWEP), propuesta por Sergio Luján-Mora (2005), basada en el lenguaje UML para el desarrollo de un Datamart.

Para proyecto se desarrollarán los primeros tres flujos de trabajo: Requerimientos, Análisis y Diseño, mencionados en del Marco de Diseño de un Data Ware House, figura 20, donde:

- Requerimientos: Se enfoca en las necesidades del usuario final, la captura de los mismos se hará mediante diagramas de Actividades y de Caso de Uso.
- Análisis: Se define y analiza la especificación de los Casos de Uso, datos que servirán como fuente para alimentar el Datamart. Se utilizarán Diagramas de Esquemas Conceptual y Lógico, los cuales permitirán modelar los datos y realizar esquematizaciones de procesos de carga (ETL).
- Diseño: Se define la estructura del Datamart a nivel lógico y físico, se relacionarán los datos fuente con los datos definidos en el modelo multidimensional, se utilizarán los Diagramas de Esquema Conceptual y el de Mapeo de Datos.

La tabla 6 resume los diagramas a utilizar en cada flujo de trabajo que se utilizará en el proyecto.

Tabla 6. Resumen de Diagramas por Flujo de Trabajo del DWEP (Elaboración Propia).

Flujo de trabajo	Diagramas
Requerimientos	Diagrama de Actividades. Diagrama de Casos de Uso.
Análisis	Esquema Conceptual de los Datos Fuente Esquema Lógico de los Datos Fuente
Diseño	Esquema Conceptual del Datamart. Diagrama de Mapeo de Datos.

Identificación de Requerimientos

El proyecto es basado en las necesidades de información del personal del proceso de Planificar y Desarrollar la Red, el cual está conformado por personas que toman algún grado de decisión a diferente nivel, entre ellos está: coordinador del área y Técnicos en Redes Eléctricas. Luego de las entrevistas (apéndice 1), con el personal del proceso y de un análisis del reporte trimestral presentado a la gerencia (apéndice 2), a nivel general se obtienen las siguientes necesidades:

- Contar con una herramienta que les facilite la toma de decisiones, la cual permita acceder a la información requerida lo más rápido posible y en cualquier momento.
- Contar con información actualizada y organizada.

En el apéndice 6 se representan los requerimientos mediante diagramas de Casos de Uso. Podemos agrupar los requerimientos en dos grandes temas:

- **Solicitudes de Estudio.** Una buena planificación de la red eléctrica permitirá cubrir todas las necesidades de crecimiento en una forma adecuada, con tiempos

razonables y convenientes para los ciudadanos del área servida por JASEC, atender potenciales beneficiarios que no cuentan con servicio eléctrico y abastecer a clientes que se encuentran en comunidades organizadas en la periferia del sistema. Todo ello permitirá que JASEC siga siendo una empresa competitiva, por lo que es de gran importancia poder analizar información referente a las diversas solicitudes de estudio realizado por los clientes actuales y potenciales.

- **Materiales**. El proceso de Planificar y Desarrollar la Red trabaja en conjunto con el proceso de Control de Materiales, el cual le provee todo tipo de material necesario para desarrollo de nuevas obras de electrificación, por lo que es de gran importancia contar con información de consumo de material, con el objetivo de mantener todo tipo de material necesario para el desarrollo oportuno de la red eléctrica.

La tabla 7 muestra las diferentes necesidades de información que se requieren para realizar una adecuada toma de decisiones. En el apéndice 5 se muestra el Documento de Análisis de Requerimientos.

Tabla 7. Necesidades de información (Elaboración Propia).

Tema	Necesidades	Filtros	Busca responder
Solicitudes de estudios.	Costos de materiales, mano de obra y transporte, presupuestados y reales de solicitudes formalizadas. Costos administrativos, de imprevistos e ingeniería.	Por fecha, por tipo de estudio, por clase de trabajo, por categoría, por cantón.	Con esta información se busca responder: ¿Cuál es el cantón con mayor crecimiento? ¿Cuál es el ingreso económico por concepto de nuevas obras? ¿Cuál es el costo real en la construcción de nuevas obras? ¿Cuál es la duración promedio entre la solicitud y la formalización de una solicitud? ¿Cuál es el tipo de estudio más solicitado?
Materiales	Consumo de materiales eléctricos, en cantidad y en costo.	Por artículo, por cantón, por año.	Con esta información se busca responder: ¿Cuál es el artículo de más consumo? ¿Cuál es el artículo de consumo de mayor costo?

Situación Actual

Actualmente la gerencia general de JASEC solicita informes de desarrollo de obras de electrificación y mejoramiento de la red eléctrica, con el objetivo de verificar avances de nuevas obras y crecimiento de la red. Esto para poder realizar una adecuada planificación, lo cual permitirá a JASEC ser una empresa más consolidada en el territorio nacional.

El departamento de Planificar y Desarrollar la Red emplea hojas de Microsoft Excel para elaborar reportes que le son de insumo para realizar una adecuada toma de decisiones, además de utilizar este tipo de hojas de cálculo para realizar análisis de datos. La información necesaria para los informes proviene de dos sistemas de transaccionales:

- Sistema de Inventario.
- Sistema de Control de Obras.

La única forma de obtener la información es por medio de los reportes programados en dichas aplicaciones, los cuales muchas veces no son flexibles ni de utilidad a las necesidades del usuario. Esto involucra que estén interactuando frecuentemente con personal de Tecnologías de Información para que le generen un nuevo reporte o extracción de datos por cada nueva necesidad, lo cual al usuario le consume trabajo y tiempo en la elaboración de los reportes, sin mencionar el riesgo inherente de manipulación de datos por parte del usuario final.

Con el sistema de inventario se obtienen datos de aquellos artículos que se utilizaron en una determinada solicitud de estudios, así como su número de requisición (documento de salida de inventario), costo del artículo y cantidad, además de determinar el stock necesario para la realización de nuevas obras de electrificación.

Con el sistema de Control de Obras, se obtiene información que les permita realizar proyecciones de crecimiento de la red y zonas con mayores crecimientos, entre otros.

En el apéndice 4 se muestra gráficamente los pasos que se siguen en el momento en que un cliente se presenta a realizar una nueva solicitud de estudio, esto mediante un Diagrama de Actividades del Negocio.

Análisis

Diseño Conceptual

Se han identificado dos esquemas de modelo estrella que permitan satisfacer las necesidades de información del proceso Planificar y Desarrollar la Red:

- Solicitud estudio.
- Inventario.

Es importante conocer algunos conceptos relacionados con el diseño conceptual, tales como:

- Granularidad: Aplicado a la tabla de hechos, lo cual indica el nivel de detalle de los registros contenidos. La tabla de hechos puede tener un nivel granular detallado o agregado. Por ejemplo, en una tabla de hechos de ventas de artículos, se podría tener un registro por cada artículo vendido o un registro que indique el monto total del artículo por mes.
- Agregación: No indica la forma en que está agrupada la información. Se relaciona con la tabla de hechos, ya que contiene un resumen de datos, como por ejemplo, monto total de ventas en un determinado mes, semestre o año.
- Jerarquía de dimensiones: Una jerarquía se da entre tablas de dimensiones. Una tabla de dimensiones está compuesta de una o más tablas hijas, es decir, se da una relación padre-hijo. La figura 27 muestra dos ejemplos de jerarquía, una es de ubicación geográfica y la otra de fecha, en donde se podría realizar análisis de solicitudes de estudio por provincia, cantón o distritos.

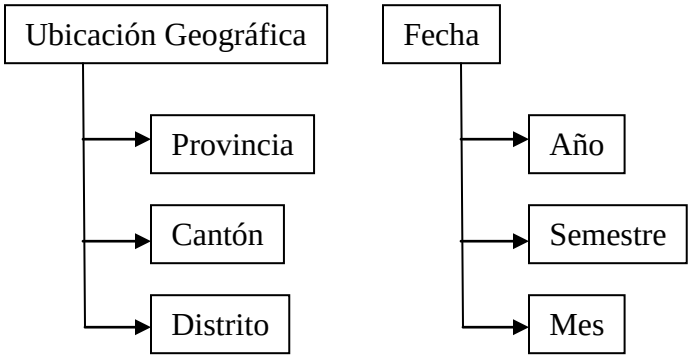


Figura 27. Jerarquía de ubicación geográfica y de fecha (Elaboración Propia).

Esquema de Solicitud Estudio

Satisface los requerimientos de información relacionados con las solicitudes de estudio. Mantiene un nivel de granularidad detallado. Consta de las dimensiones de Estado de la solicitud, Tipo de Cliente, Categoría de la Obra, Tipo de Servicio, Cantón y Fecha. La figura 28 muestra el esquema conceptual del modelo estrella.

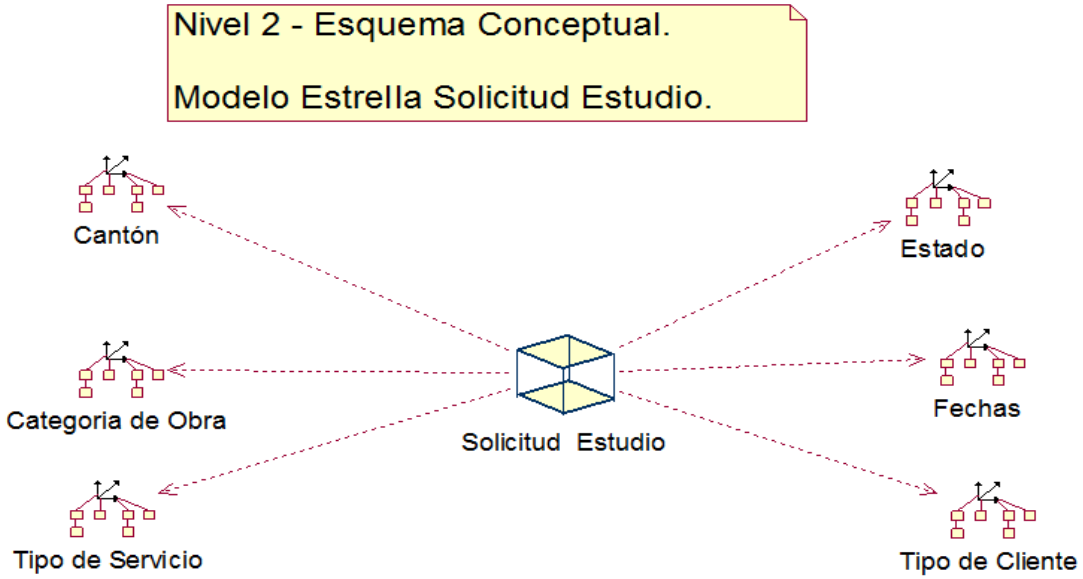


Figura 28. Esquema conceptual del modelo estrella de Solicitud de Estudio (Elaboración Propia).

Esquema de Inventario

Satisface los requerimientos de materiales. Mantiene un nivel de granularidad detallado. Consta de las dimensiones de Fecha, Solicitud de estudio, Número de requisición, Persona a la que se le entregó el material, almacén del que salió el material y artículo requisado.

La figura 29 muestra el esquema conceptual del modelo estrella.

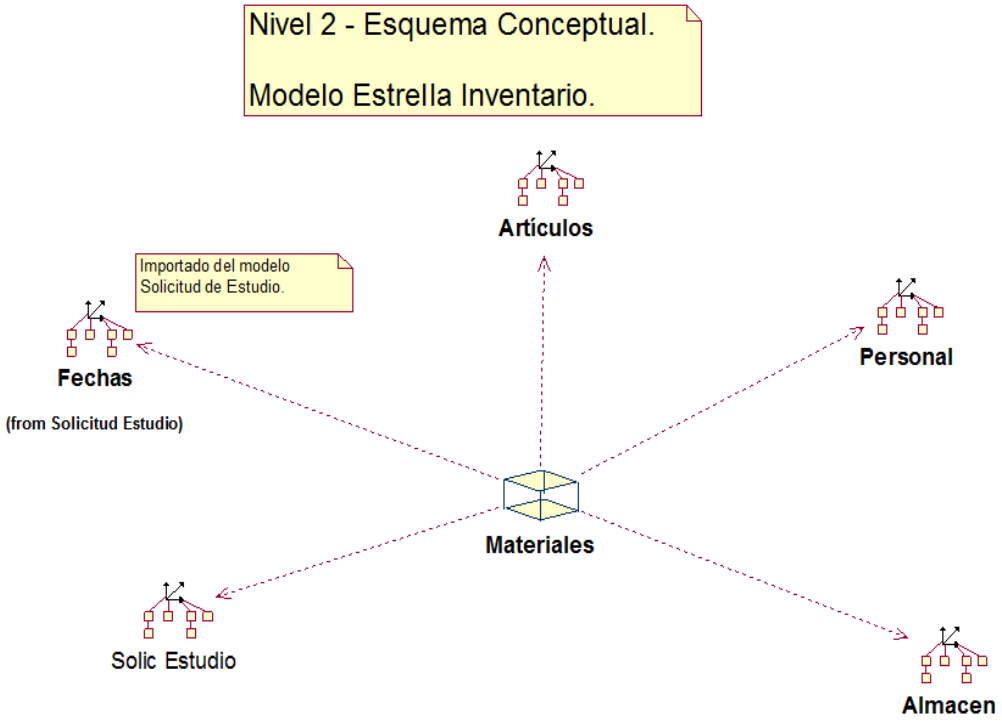


Figura 29. Esquema Conceptual del Modelo Estrella de Inventario (Elaboración Propia).

Como resumen, para poder satisfacer las necesidades de información, se ha propuesto el diseño de dos modelos estrella (Solicitud Estudio e Inventario). En la tabla 8 se muestra la relación entre las diferentes tablas de hechos con las tablas de dimensiones, en donde notamos que la dimensión de Fechas es compartida por los hechos.

Tabla 8. Relación de tabla de hechos con dimensiones (Elaboración Propia).

	Hechos	
	Solicitud Estudio	Inventario
Cantones	X	
Estados	X	
Tipo de Cliente	X	
Fechas	X	X
Tipo de Servicio	X	
Categoría de la Obra	X	
Solicitud de Estudio		X
Artículos		X
Personal		X
Almacén		X

El apéndice 7 muestra el Esquema Conceptual de Datamart.

Diseño Lógico

En el diseño lógico definiremos los campos de las tablas con sus tipos de datos. Dentro de los tipos de datos, tenemos datos:

- Numéricos.
- Alfanuméricos.
- Fecha, entre otros.

Esquema Solicitud Estudio

El esquema de Solicitud Estudios brinda información de las Solicitudes de Estudios. Las medidas relacionadas dan a conocer los diferentes ingresos, tal como monto de solicitud, monto presupuestado de material, monto presupuestado de mano de obra, monto presupuestado de transporte, monto de ingeniería, imprevistos y gastos administrativos, que son montos que debe de cancelar el cliente, además de los montos reales de la obra, tal como el monto real de material, mano de obra y transportes. La figura 30 muestra el esquema lógico del modelo de Solicitud Estudio (DWLS).

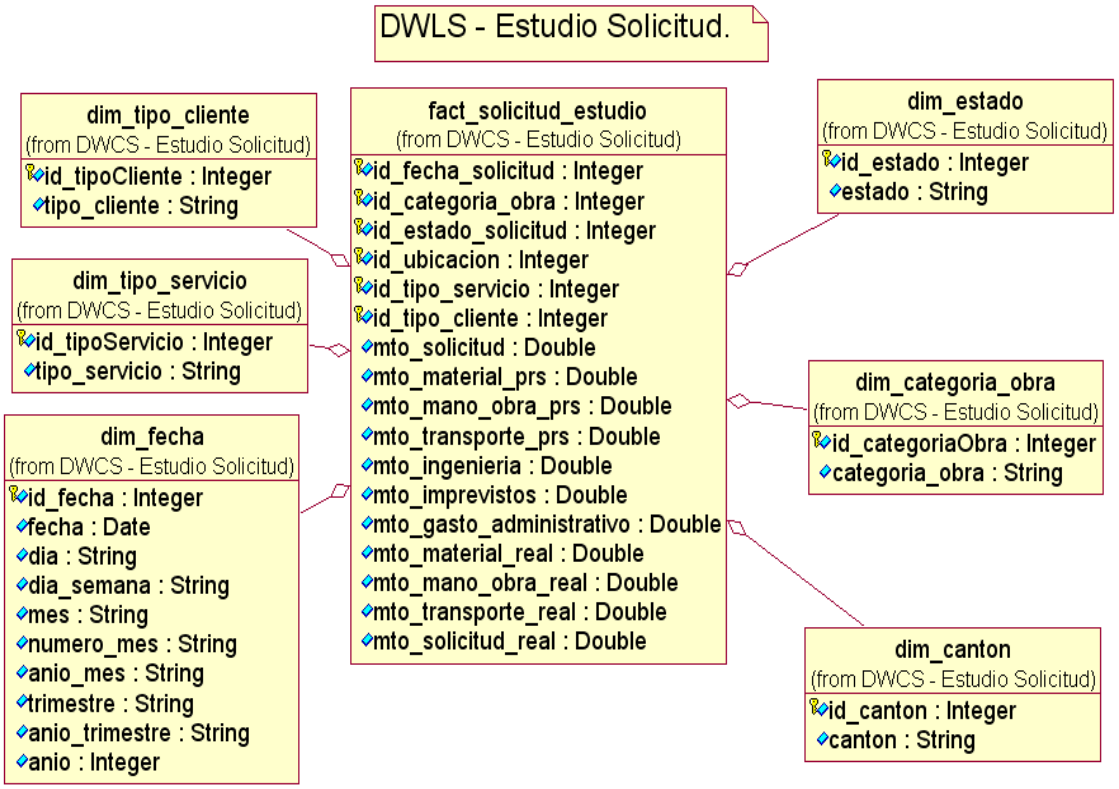


Figura 30. Esquema Lógico, Modelo Solicitud Estudio (Elaboración Propia).

Esquema Inventario

El esquema de Inventario brinda información de las salidas de artículos asociado a una Solicitud de Estudio. Las medidas relacionadas dan a conocer la cantidad requisada y costo de materiales. En la figura 31 se muestra el esquema lógico del modelo de Inventario (DWLS).

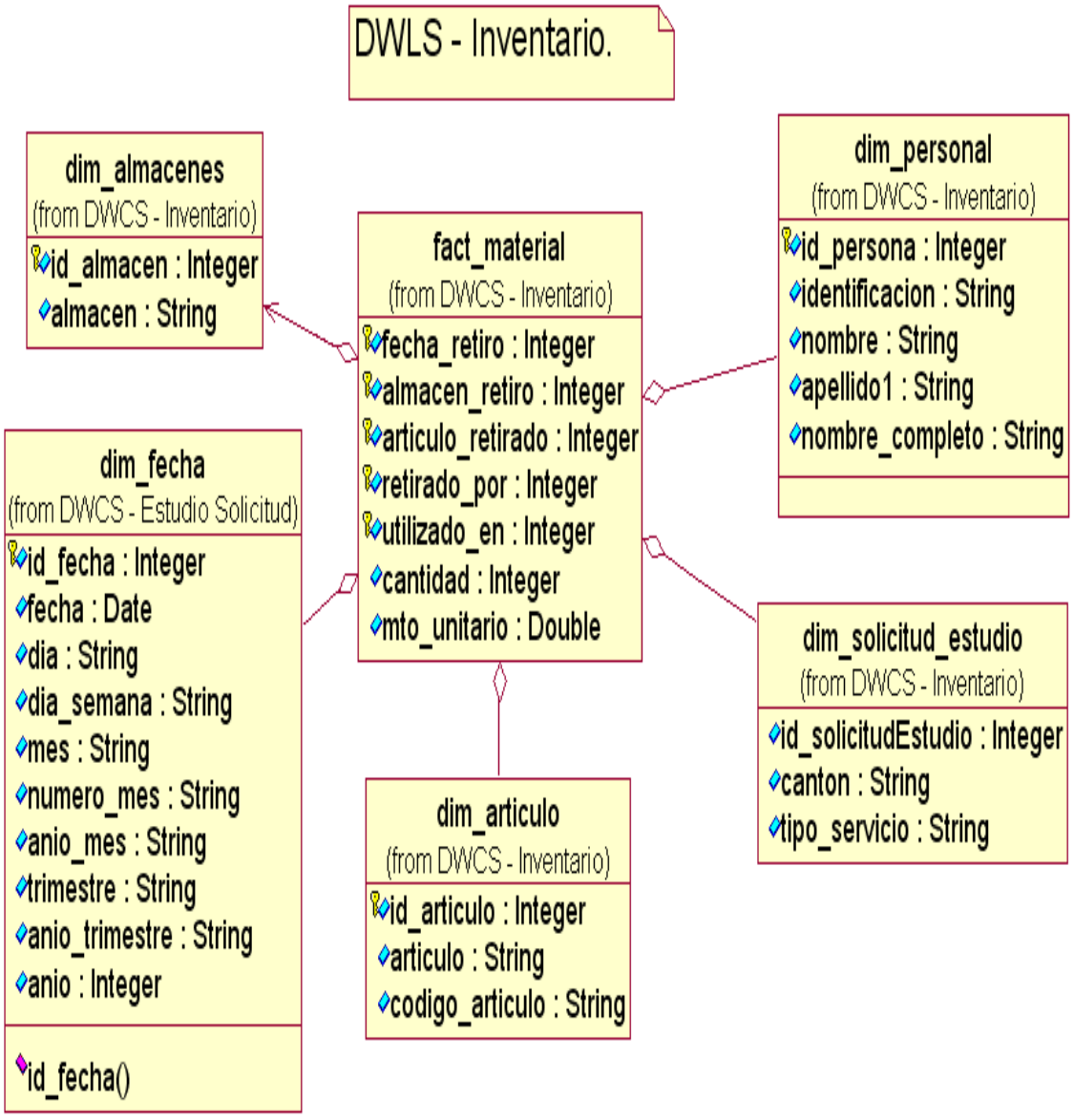


Figura 31. Esquema Lógico, Modelo Inventario (Elaboración Propia).

Diseño

Arquitectura

La figura 32 permite mostrar los componentes de los cuales se conforma la arquitectura de la solución.

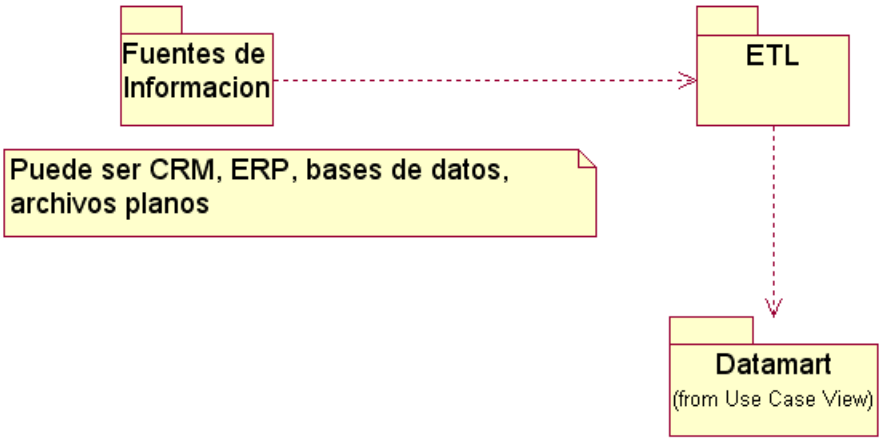


Figura 32. Diagrama de arquitectura (Elaboración Propia).

La arquitectura se conforma de varios componentes:

- Fuente de Información: Indica la procedencia de los datos que alimentan al Datamart. La fuente de datos puede ser documentos Excel, documentos de texto, bases de datos transaccionales y CRM, entre otros. Para nuestro proyecto, la fuente de datos identificada es una base de datos transaccional, la cual contiene la información referente a la solicitud de estudio y de material utilizado en cada solicitud.
- ETL: Este componente involucra una serie procesos que permiten extraer información de la fuente, transformarlos y cargarlos al Datamart.
- Datamart: Es la base de datos multidimensional. En ella se integran todos los datos provenientes de diversas fuentes de información.

Fuentes de Información

Para el esquema de Solicitud Estudio e Inventario se tomará como fuente de información la base de datos transaccional de la empresa, específicamente aquellos objetos de base de datos que almacenan información referente a las solicitudes de estudio y consumo de material por cada solicitud de estudio. En el apéndice 8 se muestra el Esquema Lógico y Conceptual de la base de datos que contiene los insumos necesarios para el Datamart.

ETL

Es el proceso encargado de obtener la información de las diferentes fuentes y pasarlas al Datamart. Durante el proceso de pase la información es transformada, de tal forma que esta pueda ser cargada a las tablas de dimensiones y de hechos.

Se podría ver el proceso de diseño del ETL en dos pasos:

- Mapear los datos de las fuentes con el Datamart. La metodología DWEP propone un Diagrama de Mapeo de Datos, el cual es mostrado en el apéndice 9.
- Diseñar un flujo de proceso de ETL, el cual muestra la carga de cada tabla de dimensión y de hecho. La metodología DWEP propone un diagrama de Diseño de Procesos ETL, el cual es mostrado en el apéndice 11.

Conclusiones

En el presente proyecto se plantearon los siguientes objetivos:

General

- Diseñar un Data Mart para el área de Planificar y Desarrollar la Red de la JASEC modelado mediante UML.

Objetivos Específicos

- Conocer el proceso actual de planificación y ejecución de proyectos.
- Definir los requisitos y funciones deseadas para el Data Mart.
- Revisar las fuentes de datos.
- Construir el modelo conceptual y lógico del Data Mart.

El primer objetivo específico, "Conocer el proceso actual de planificación y ejecución de proyectos", se logró alcanzar al aplicar la técnica de Observación y de Entrevista realizada en el proceso de Planificar y Desarrollar la Red. Al analizar los resultados del proceso actual, se pudo observar y entender que el proceso actual para poder planificar y desarrollar nuevas obras eléctricas es un proceso ineficiente, propenso a errores y con un alto consumo de tiempo, ya que deben estar interactuando constantemente con el departamento de TI para obtener datos y poder manipularlos en hojas de cálculo. La entrevista realizada se encuentra en el apéndice #1 y la especificación del proceso actual se encuentra definida dentro del apartado de Marco Metodológico, además del apéndice #4 que muestra un diagrama de actividades del proceso actual.

El segundo objetivo específico, "Definir los requisitos y funciones deseadas para el Data Mart", fue alcanzado aplicando técnicas de grupos de discusión y reuniones sostenidas con personal de Planificar y Desarrollar la Red, en donde se logró determinar qué tipo de información se requiere para poder satisfacer las necesidades referentes a la planificación de nuevas obras. Se determinó que se necesitan datos de dos áreas, tales como Solicitud de Estudio y de Materiales. Para este objetivo se generó un documento de análisis de requerimientos, el cual se muestra en el apéndice #5.

El tercer objetivo específico, "Revisar las fuentes de datos", se logró alcanzar mediante una serie de reuniones con el personal de TI, en donde se logró determinar de cuáles bases de datos, esquemas, tablas y campos se debe extraer la información necesaria para poder satisfacer las necesidades del proceso de Planificar y Desarrollar la Red. La información debe ser extraída de una única base de datos y de tres esquemas. Para este objetivo se generó el documento Mapeo de Datos, el cual se muestra en el apéndice #9.

El cuarto objetivo específico, "Construir el modelo conceptual y lógico del Data Mart", se logró alcanzar con la creación de los diagramas de Esquema Conceptual del Datamart y Esquema lógico del Datamart, el cual se muestra en el apéndice #10. Antes se tuvo que determinar cuáles son las tablas de hechos y dimensiones a utilizar, por lo que se generó el apéndice #7.

En cuanto al objetivo general, "Diseñar un Data Mart para el área de Planificar y Desarrollar la Red de la JASEC modelado mediante UML", se logró al utilizar la metodología DWEP, y se utilizó un lenguaje común UML para el desarrollo de un Data Warehouse/Datamart, el cual permitió facilitar el proceso de diseño del Datamart.

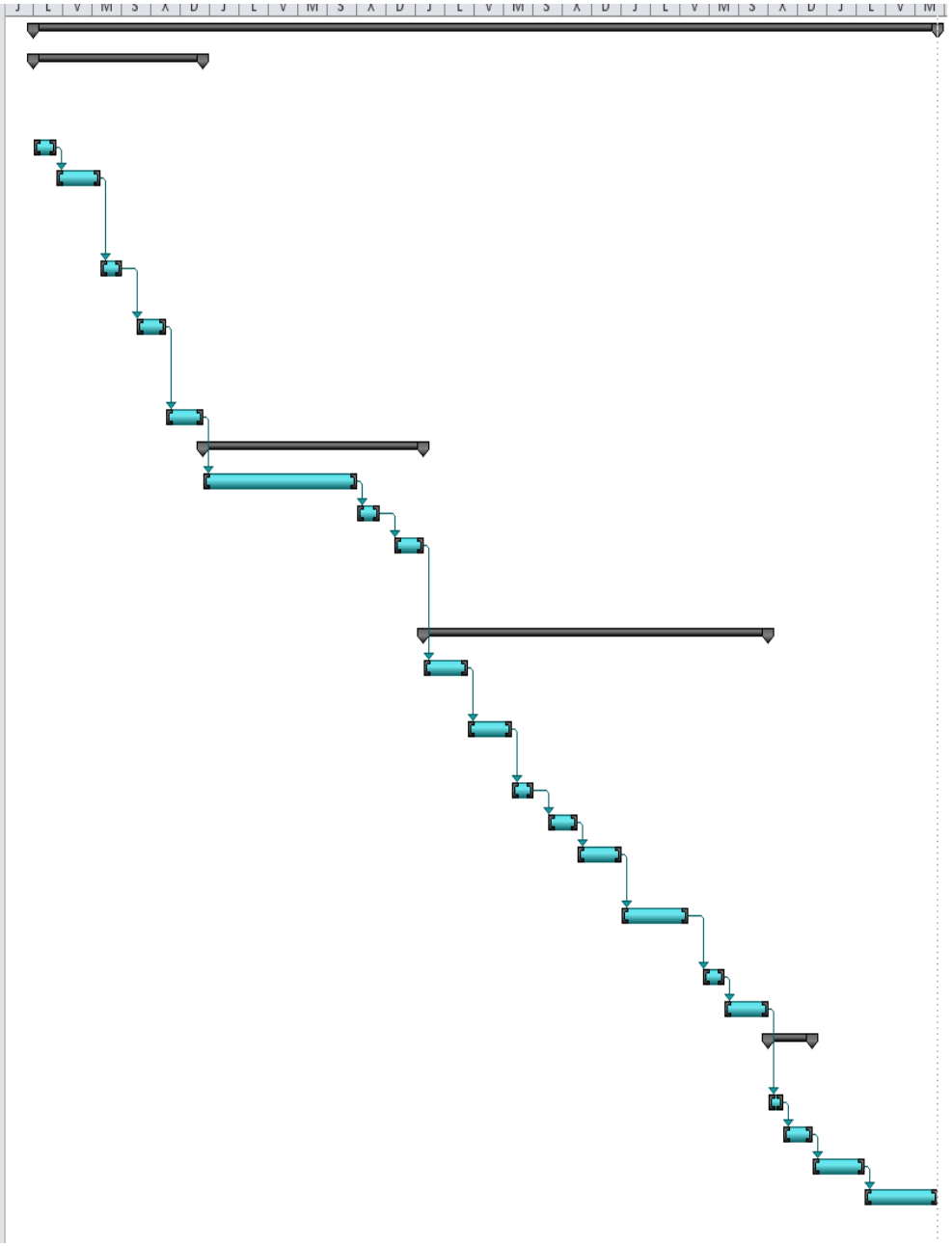
Cronograma

Nota:

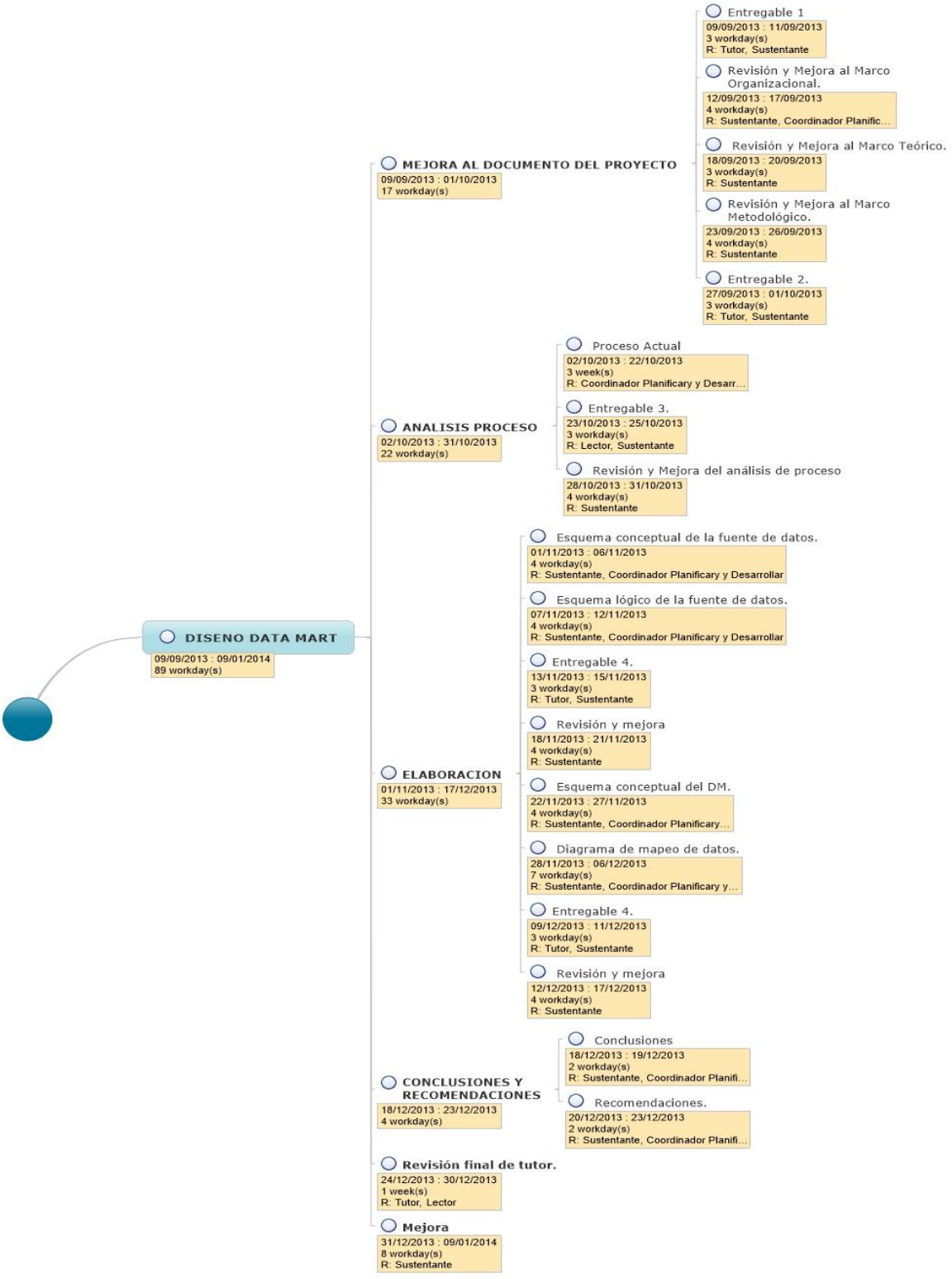
Este cronograma se encuentra sujeto a disposición de la JASEC.

Nombre de tarea	Duración	Comienzo	Fin
DISEÑO DATA MART	89 días	lun 09/09/13	jue 09/01/14
MEJORA AL DOCUMENTO DEL PROYECTO	17 días	lun 09/09/13	mar 01/10/13
Entregable 1	3 días	lun 09/09/13	mié 11/09/13
Revisión y Mejora al Marco Organizacional.	4 días	jue 12/09/13	mar 17/09/13
Revisión y Mejora al Marco Teórico.	3 días	mié 18/09/13	vie 20/09/13
Revisión y Mejora al Marco Metodológico.	4 días	lun 23/09/13	jue 26/09/13
Entregable 2.	3 días	vie 27/09/13	mar 01/10/13
ANÁLISIS PROCESO	22 días	mié 02/10/13	jue 31/10/13
Proceso Actual	15 días	mié 02/10/13	mar 22/10/13
Entregable 3.	3 días	mié 23/10/13	vie 25/10/13
Revisión y Mejora del análisis de proceso	4 días	lun 28/10/13	jue 31/10/13
ELABORACIÓN	33 días	vie 01/11/13	mar 17/12/13
Esquema conceptual de la fuente de datos.	4 días	vie 01/11/13	mié 06/11/13
Esquema lógico de la fuente de datos.	4 días	jue 07/11/13	mar 12/11/13
Entregable 4.	3 días	mié 13/11/13	vie 15/11/13
Revisión y mejora	4 días	lun 18/11/13	jue 21/11/13
Esquema conceptual del DM.	4 días	vie 22/11/13	mié 27/11/13
Diagrama de mapeo de datos.	7 días	jue 28/11/13	vie 06/12/13
Entregable 4.	3 días	lun 09/12/13	mié 11/12/13
Revisión y mejora	4 días	jue 12/12/13	mar 17/12/13
CONCLUSIONES Y RECOMENDACIONES	4 días	mié 18/12/13	lun 23/12/13
Conclusiones	2 días	mié 18/12/13	jue 19/12/13
Recomendaciones.	2 días	vie 20/12/13	lun 23/12/13
Revisión final de tutor.	5 días	mar 24/12/13	lun 30/12/13
Mejora	8 días	mar 31/12/13	jue 09/01/14

DISENO DATA MART	**89 días**	**lun 09/09/13**	**jue 09/01/14**
MEJORA AL DOCUMENTO DEL PROYECTO	**17 días**	**lun 09/09/13**	**mar 01/10/13**
Entregable 1	3 días	lun 09/09/13	mié 11/09/13
Revisión y Mejora al Marco Organizacional.	4 días	jue 12/09/13	mar 17/09/13
Revisión y Mejora al Marco Teórico.	3 días	mié 18/09/13	vie 20/09/13
Revisión y Mejora al Marco Metodológico.	4 días	lun 23/09/13	jue 26/09/13
Entregable 2.	3 días	vie 27/09/13	mar 01/10/13
ANALISIS PROCESO	**22 días**	**mié 02/10/13**	**jue 31/10/13**
Proceso Actual	15 días	mié 02/10/13	mar 22/10/13
Entregable 3.	3 días	mié 23/10/13	vie 25/10/13
Revisión y Mejora del análisis de proceso	4 días	lun 28/10/13	jue 31/10/13
ELABORACION	**33 días**	**vie 01/11/13**	**mar 17/12/13**
Esquema conceptual de la	4 días	vie 01/11/13	mié 06/11/13
Esquema lógico de la fuente de datos.	4 días	jue 07/11/13	mar 12/11/13
Entregable 4.	3 días	mié 13/11/13	vie 15/11/13
Revisión y mejora	4 días	lun 18/11/13	jue 21/11/13
Esquema conceptual del DM.	4 días	vie 22/11/13	mié 27/11/13
Diagrama de mapeo de datos.	7 días	jue 28/11/13	vie 06/12/13
Entregable 4.	3 días	lun 09/12/13	mié 11/12/13
Revisión y mejora	4 días	jue 12/12/13	mar 17/12/13
CONCLUSIONES Y RECOMENDACIONES	**4 días**	**mié 18/12/13**	**lun 23/12/13**
Conclusiones	2 días	mié 18/12/13	jue 19/12/13
Rec omendaciones	2 días	vie 20/12/13	lun 23/12/13
Revisión final de tutor.	**5 días**	**mar 24/12/13**	**lun 30/12/13**
Mejora	**8 días**	**mar 31/12/13**	**jue 09/01/14**



Work Breakdown Structure (WBS)



Referencias

Adamson, C. (2010). Complete Reference. Star Scheme. McGraw-Hill.

ARESEP (2013). Tarifas Eléctricas por Empresa Distribuidora. Recuperado el 11 de noviembre del 2013: http://aresep.go.cr/index.php/inicio/index.php?option=com_content&view=article&id=1024&catid=72&Itemid=392

Centro de Investigación en Inteligencia de Negocios (2013). Resultados del 2do foro en Business Intelligence. Recuperado el 02 de octubre del 2013: http://www.ceine.cl/resultadosforo2bi/

Elliott, J. (1993). El cambio educativo desde la investigacion-accion. Madrid: Morata.

Hernández, R. Fernández-Collado, C. & Baptista, P. (2010). Metodología de la investigación. 5ed. México: Ed. McGraw-Hill.

Inmon, W. (1992). Building the Data Warehouse. Wiley-QED John Wiley & Sons, Inc.

JASEC (2013). División Distrital del Área Servida por JASEC. Recuperado el 20 de noviembre del 2013: http://www.jasec.co.cr/index.php/jasec-quienes-somos/informacion-general/mapa-virtual.

Kimball, R. (1992). The Data Warehouse Toolkit. Wiley Computer Publishing.

Kimball, R. (1996). The Data Warehouse Toolkit. John Wiley & Sons. (2nd Edition).

Kimball, R. & Reeves, L. & Ross, M., Thornthwaite, W. (1998). The Data Warehouse Lifecycle Toolkit, Wiley Computer Publishing.

Kriegel, A. (2011). Discovering SQL. A Hands-On Guide for Beginners. Wiley Publishing, Inc.

Luján, S. (2005). Data Warehouse Design with UML. Universidad de Alicante.

Milicev, D. (2009). Model-Driven Development with Executable UML.Wiley Publishing, Inc.

Object Management Group (2011). OMG Unified Modeling Language (OMG UML) Infrastructure. Recuperado el 12 de noviembre del 2013: http://www.omg.org/spec/UML/2.4.1/Infrastructure/PDF/

Oppel, A. (2009). Data Modeling: A Beginner's Guide. McGraw-Hill.

ORACLE (2007). Oracle Business Intelligence Standard Edition One Tutorial Release 10g (10.1.3.2.1). Recuperado el 12 de noviembre del 2013: http://docs.oracle.com/html/E10312_01/dm_concepts.htm

ORACLE (2010). Oracle Business Intelligence Discoverer Plus User's Guide 10g Release 2 (10.1.2.1). Recuperado el 12 de Noviembre del 2013: http://docs.oracle.com/html/B13915_04/i_olap_chapter.htm

PDR (2012). Plan de Desarrollo de la Red, 2012. Proceso Gestionar Tarifas, JASEC.

PDR (2013). Informe Planificar y Desarrollar la Red II Trimestre, Año 2013. Proceso Planificar y Desarrollar la Red, JASEC.

RH (2013). Organigrama de la Junta Administrativa del Servicio Eléctrico de Cartago. Departamento de Recursos Humanos, JASEC.

UML (2013). UML 2.4 Diagrams Overview. Recuperado el 12 de diciembre del 2013: http://www.uml-diagrams.org/uml-24-diagrams.html

Zambrano, J. (2011). Análisis, diseño e implementación de un Datamart para el área de mantenimiento y logística de una empresa de transporte público de pasajeros. Pontifica Universidad Católica de Perú.

APÉNDICES

Apéndice #1: Cuestionario de Entrevista

El siguiente cuestionario se le realizó al coordinador de proceso de Planificar y Desarrollar la Red. Tiene por objetivo determinar la viabilidad e importancia que tendrá el diseño e implementación del Datamart.

1. ¿Qué tan frecuente la Gerencia General de JASEC le solicita informes?

Mensual	Trimestral	Semestral	Anual	Otro: _____

2. ¿En la actualidad el proceso de Planificar y Desarrollar la Red cuenta con un sistema automatizado que le permita realizar una adecuada toma de decisiones?

SI	NO

3. ¿De cuál sistema de información obtiene datos para realizar una adecuada toma de decisiones?

Sistema Contable	
Sistema de Control de Obras	
Sistema de Inventario	
Sistema Presupuestario	
Sistema de Recursos Humanos	
Sistema de Control de Averías	

4. ¿De los sistemas mencionados, los reportes que le generan satisfacen a las necesidades del departamento?

Nada	Poco	Mucho

5. ¿De los sistemas anteriormente seleccionados, estos le permiten obtener todos los insumos de información necesarios?

SI	NO

6. ¿Tiene que recurrir frecuentemente al departamento de TI para que le brinde información primaria o le realice reportes a la medida para una adecuada toma de decisiones?

SI	NO

7. ¿Qué tipo de herramienta utiliza para elaborar reportes y ejecutar análisis sobre los datos?

Sistema Automatizado	Hoja de cálculo	Otro: _____

8. ¿Tiene problemas al tener que transcribir la información enviada por TI para redactar los informes gerenciales?

Errores de transcripción	Información Inoportuna (retrasos)	Dependencia de TI	Otros: _____ _____

9. ¿Qué beneficio le traerá tener una futura implementación un Datamart para su departamento?

Ninguna	Poca	Mucha

10. ¿En qué área le beneficiaría la implementación?

Consulta rápida de información	Toma de decisiones	Nada

Apéndice #2: Informe trimestral

El siguiente muestra un informe gerencial trimestral, generado por el área de Planificar y Desarrollar la Red para una adecuada toma de decisiones.

Apéndice #3: Diagramas UML

La figura 33 muestra que UML tiene dos grandes agrupaciones:

- Diagramas de estructura.
- Diagramas de comportamiento.

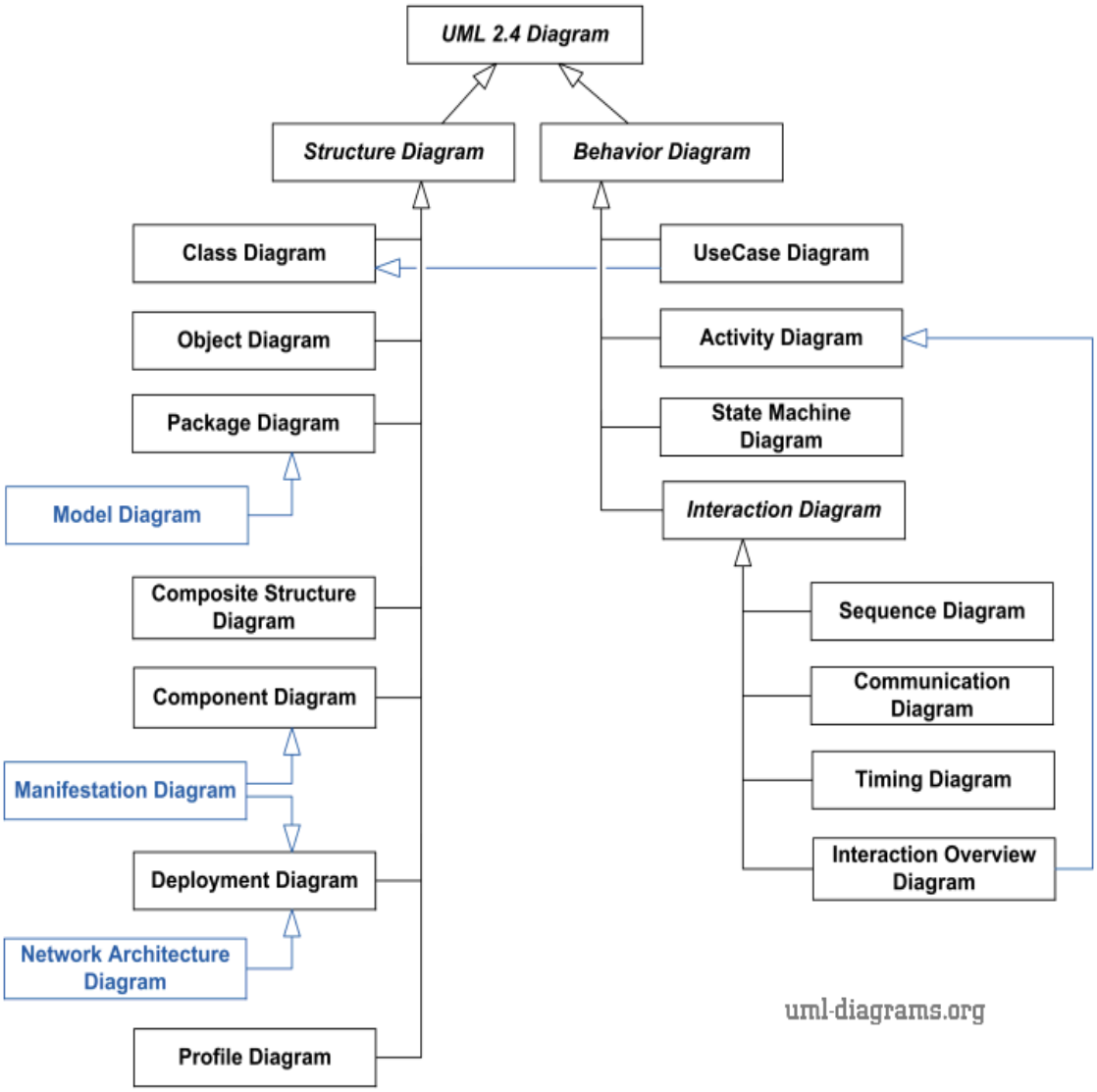


Figura 33. Diagramas UML 2.4. (UML, 2013).

Tipos de Diagramas

Diagrama	Descripción
Diagrama de clases	Permite visualizar las relaciones entre clases que involucra un sistema. Son muy utilizados en el análisis y diseño de un sistema. Principalmente contiene clases (atributos, métodos y visibilidad) y relaciones (herencia, composición, agregación, asociación y uso).
Diagrama de objetos	Se modelan las instancias de las clases mostradas en el diagrama de clases. Permite mostrar un conjunto de objetos y sus relaciones. Son útiles para modelar estructuras complejas.
Diagrama de paquetes	Muestra la descomposición del modelo en unidades (paquetes) y sus dependencias. Permiten simplificar los diagramas de clases.
Diagrama de Estructura Compuesta	Muestra la estructura interna (partes y conectores) de una clase.
Diagrama de componentes	Muestra las dependencias lógicas entre los diferentes componentes de software.
Diagramas de despliegue	Muestra la configuración física del Sistema en términos de nodos y sus respectivas conexiones, tales como conexiones de red. Esencialmente es un diagrama de clases centrado en nodos del

	sistema.
Diagrama de Casos de uso	Describe el alcance del sistema desde un punto de vista externo. Permite mostrar los requerimientos funcionales o necesidades de los usuarios.
Diagrama de Actividad	Describe el flujo del sistema. Muestra en orden la secuencia de actividades y las decisiones a tomar para llevar a cabo un proceso, el cual puede ser una actividad del negocio.
Diagrama de secuencia	Muestra la interacción entre objetos en un determinado caso de uso. Modela cómo interactúan los objetos basados en una secuencia de tiempo.
Diagrama de comunicación	Similar al diagrama de secuencia, es usado para modelar un alcance dinámico de un caso de uso. Este diagrama se focaliza más en mostrar la colaboración de objetos en vez de la secuencia de tiempo.
Diagrama de Tiempo	Muestra el alcance de los objetos en un periodo de tiempo determinado.
Diagrama de iteración (Iteraction overview)	Se centra en el flujo de control de las iteraciones. Describe el intercambio de mensajes.

Apéndice #4: Diagrama de Actividades

Las figura 34 muestra la secuencia de actividades que se realiza cuando se realiza una solicitud de estudio, la figura 35 muestra la secuencia de actividades realizada cuando se solicita un material en el almacén de inventario.

Solicitud de Estudios.

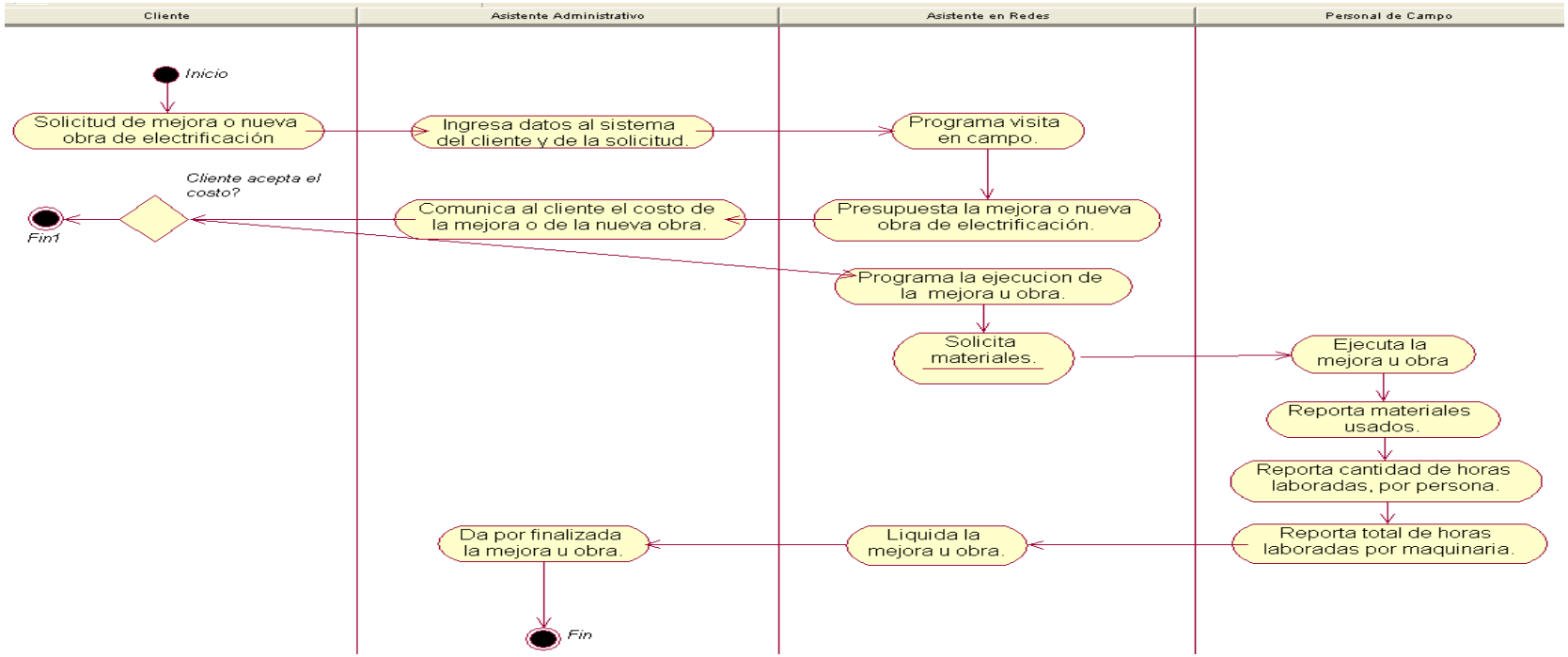


Figura 34. Diagrama de Actividades. Solicitud de Estudios (Elaboración Propia).

Movimiento del almacén.

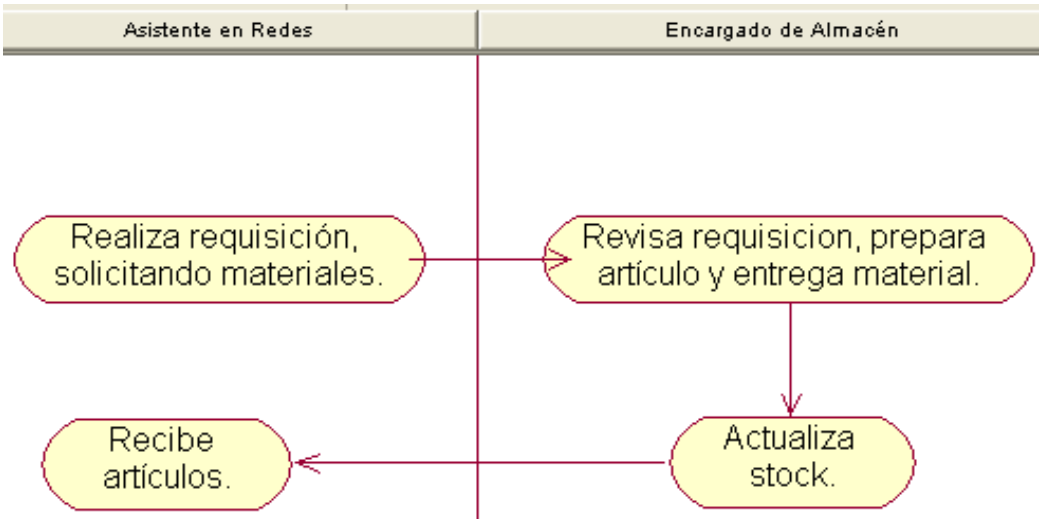


Figura 35. Diagrama de Actividades. Movimiento del Almacén (Elaboración Propia).

Apéndice #5: Documento de Análisis de Requerimientos

1. Objetivo

Este documento tiene por objetivo definir en forma clara y precisa los requerimientos del diseño del Datamart, el cual está orientado a satisfacer las necesidades de los usuarios del proceso de Planificar y Desarrollar la Red de la UEN de Energía.

2. Alcance

Se definirán las necesidades de información que requiere el proceso de Planificar y Desarrollar la Red para una adecuada toma de decisiones. El proyecto contempla solamente el diseño del Datamat, no así su implementación ni la herramienta de Inteligencia de Negocios.

3. Requerimientos

Las necesidades de información se pueden dividir en dos grandes temas:

- *Solicitudes de Estudio*. Una buena planificación de la red eléctrica permitirá cubrir todas las necesidades de crecimiento en una forma adecuada, con tiempos razonables y convenientes para los ciudadanos del área servida por JASEC, atender potenciales beneficiarios que no cuentan con servicio eléctrico, abastecer a clientes que se encuentran en comunidades organizadas en la periferia del sistema, todo ello permitirá que JASEC siga siendo una empresa competitiva. Por esto es importante analizar información referente a las solicitudes de estudio realizado por los clientes actuales y potenciales. Es de mucha importancia analizar información registrada en las solicitudes de estudios, ya que con dicha información se puede responder a preguntas tales como:

- ¿Cuál es el cantón con mayor crecimiento?
- ¿Cuál es el ingreso económico por concepto de nuevas obras?
- ¿Cuál es la duración promedio entre la solicitud y la formalización de una solicitud?
- ¿Cuál es el tipo de estudio más solicitado?
- ¿Cuánto le cuesta a JASEC el desarrollo de una nueva obra?

- *Materiales*. El proceso de Planificar y Desarrollar la Red trabaja en conjunto con el proceso de Control de Materiales, el cual le provee todo tipo de material necesario para desarrollo de nuevas obras de electrificación. Por ello es de gran importancia contar con información de consumo de material con el objetivo de mantener todo tipo de material necesario para el desarrollo oportuno de la red eléctrica. Con este tipo de información se puede responder a preguntas tales como:
  - ¿Cuál es el artículo de más consumo?
  - ¿Cuál es el artículo de consumo de mayor costo?

4. Análisis de Requerimientos

a. Diagrama de contexto

En este apartado definiremos a grandes rasgos la solución que permita satisfacer dichos requerimientos, la figura 36 muestra que toda la solución se compone de la implementación del Datamart y del sistema de Inteligencia de Negocio que permita a los usuarios explotar la información almacenada en el Datamart.

Es importante recalcar que el trabajo a realizar tiene un alcance de diseño del Datamart, no así su implementación.

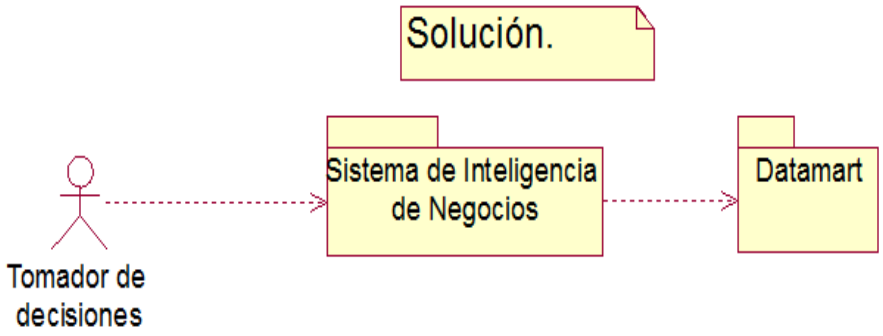


Figura 36, Diagrama de contexto (Elaboración Propia).

b. Análisis de requerimientos de los usuarios

En el apartado de requerimientos se ha indicado que las necesidades de información de los usuarios se pueden dividir en dos grandes temas (Solicitudes de Estudios y Materiales), por lo que a podríamos agruparlo en un único paquete, como lo muestra la figura 37.

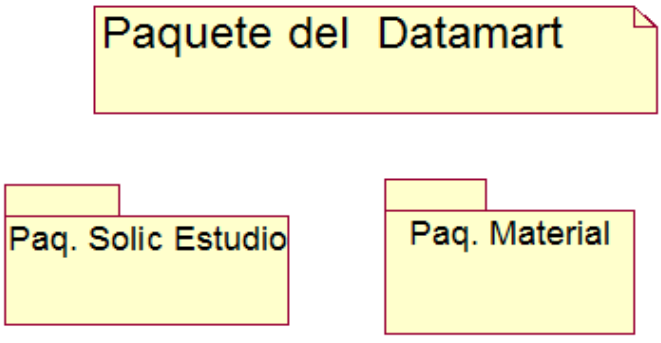


Figura 37. Paquete Datamart (Elaboración Propia).

b.1 Paquete de Solicitud de Estudios

El paquete cubre las necesidades de información relacionadas con las Solicitudes de Estudios. La figura 38 muestra los casos de uso que agrupa la necesidad de información relacionado a las solicitudes.

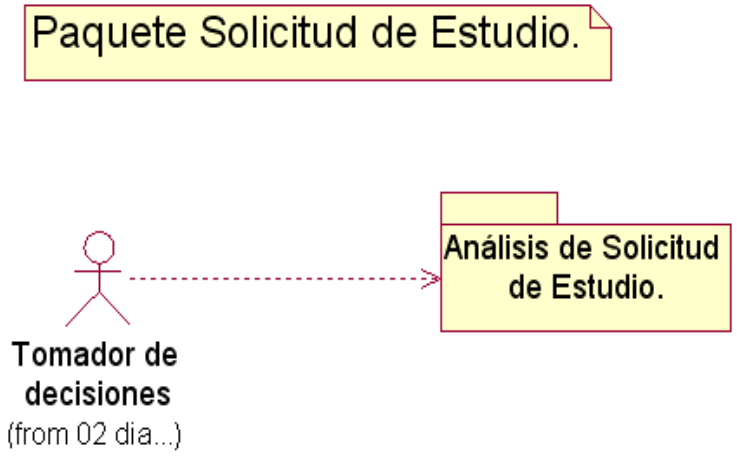


Figura 38. Paquete de Análisis de Solicitud de Estudios (Elaboración Propia).

Caso de uso

Caso de uso	Análisis de Solicitud de Estudio.
Frecuencia de actualización	Diaria
Dimensiones de análisis de la información	• Cantón • Tipo de servicio. • Categoría de la obra. • Tipo de cliente (Jurídico o físico). • Cliente. • Estado de la solicitud. • Estado de la solicitud (Aprobada, notificada, etc.). • Fecha.
Fórmulas de cálculo	No aplica.
Granularidad	Detallada
Otros	Subtotales agrupados ya sea por fecha, cliente, tipo de servicio u alguna otra dimensión.

b.2 Paquete de Materiales

El paquete cubre las necesidades de información relacionadas con los materiales utilizados en las diferentes Solicitudes de Estudios. La figura 39 muestra los casos de uso que agrupa la necesidad de información relacionada con las solicitudes.

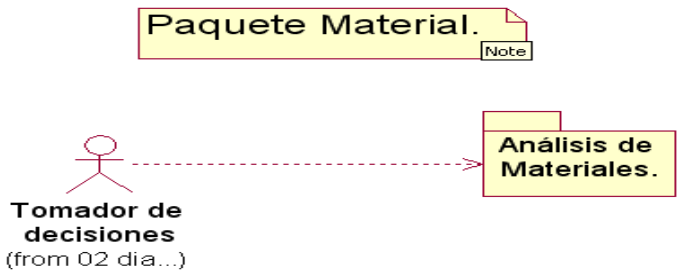


Figura 39. Paquete de Análisis de Materiales (Elaboración Propia).

Caso de uso

Caso de uso	Análisis de Materiales.
Frecuencia de actualización	Diaria
Dimensiones de análisis de la información	• Nombre de artículo. • Solicitud de estudio. • Número de requisición. • Tipo de servicio. • Categoría de la obra. • Cantón. • Fecha.
Fórmulas de cálculo	No aplica.
Granularidad	Detallada
Otros	Subtotales agrupados ya sea por fecha, cliente, tipo de servicio u alguna otra dimensión.

Realizado por:	Osvaldo Navarro Navarro.	
Aprobado por:	Ing. Roger Carrillo. Coordinador Planificar y Desarrollar la Red.	

Apéndice #6. Diagrama de Casos de Uso

En el documento de Análisis de Requerimientos (Apéndice #5) se determinó dos grandes casos de uso:

- Análisis de Solicitud de Estudio.
- Análisis de Materiales.

De acuerdo con la metodología DWEP, procederemos a presentar las especificaciones de los casos de uso.

1. Paquete de Análisis de Estudio

a. Caso de uso de análisis de estudio

Caso de Uso:	Análisis de Estudio.
ID:	CU001
Precondición:	Carga diaria de datos relacionados a los estudios.

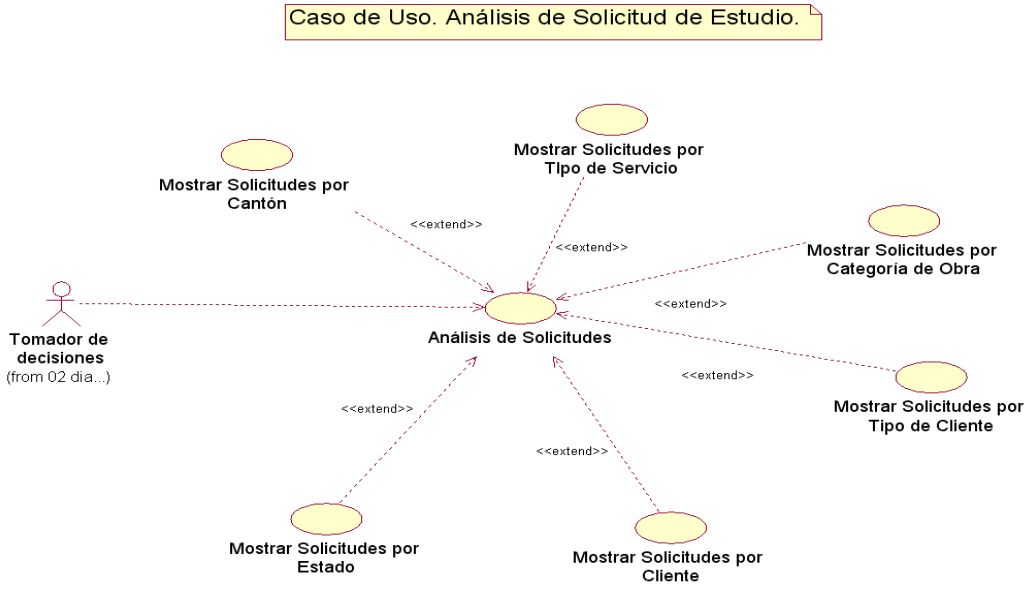


Figura 40. Diagrama Caso de Uso Análisis de Solicitudes (Elaboración Propia).

2. Paquete Análisis de Materiales

a. Caso de uso Análisis de Materiales

Caso de Uso:	Análisis de Materiales.
ID:	CU002
Precondición:	Carga diaria de datos relacionados de materiales relacionados con los estudios.

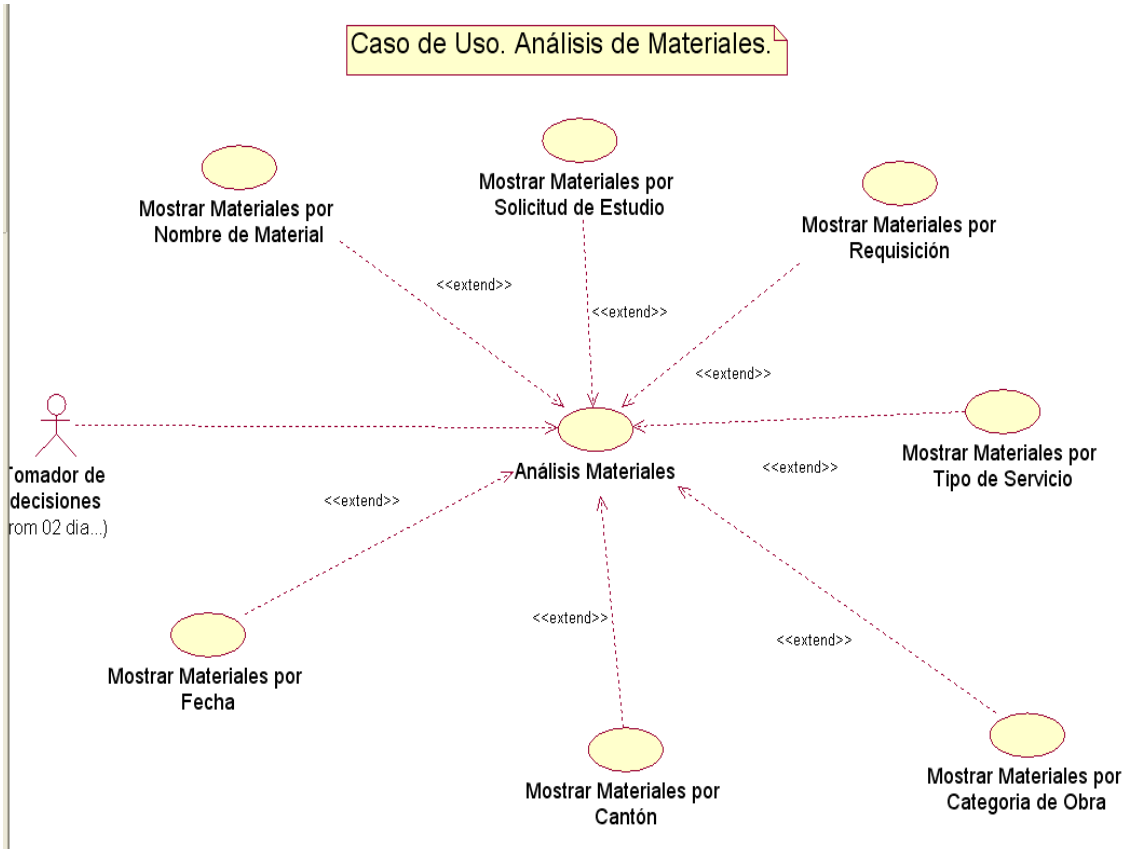


Figura 41. Diagrama Caso de Uso Análisis de Materiales (Elaboración Propia).

Apéndice #7. Esquema Conceptual/Análisis Dimensional.

ESQUEMA CONCEPTUAL

Apegado a la metodología DWEP, el esquema conceptual consta de tres niveles. La figura 42 nos muestra los niveles, donde:

Nivel 1: Definición del modelo. Diagrama básico, el cual muestra todos los modelos estrella que comprende el Datamart. Cada paquete representa un esquema estrella del modelo multidimensional. Se puede presentar dependencia entre dos paquetes, el cual indica que una o varias dimensiones es compartida por varias esquemas estrellas.

Nivel 2: Definición del esquema estrella. Muestra las diferentes dimensiones y tabla de hecho de un determinado modelo estrella.

Nivel 3: Es la definición de las tablas de hecho y dimensión, que muestra sus atributos.

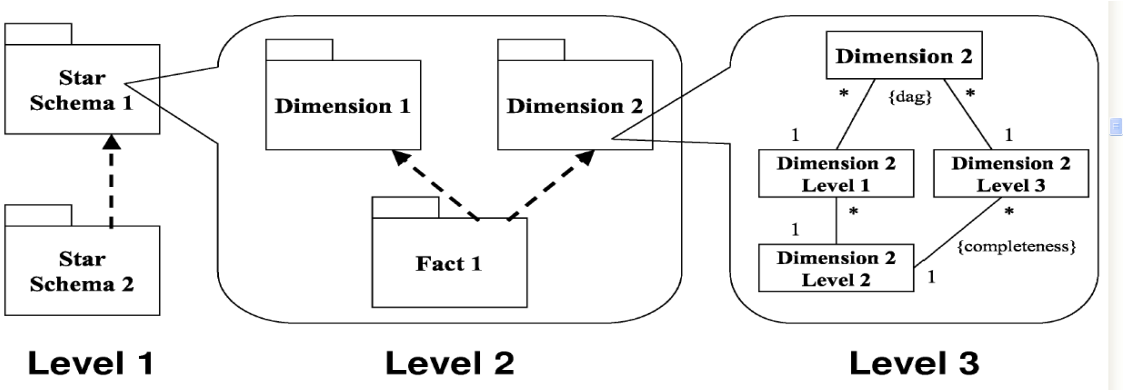


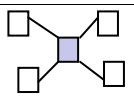
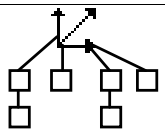
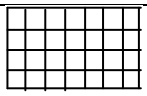
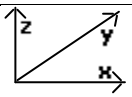
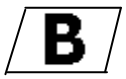
Figura 42. Relación entre los tres niveles (Luján, 2005).

Como podemos observar, el modelado comienza desde un alto nivel de abstracción y en los siguientes niveles se incrementa el nivel de detalle.

Para el modelamiento se utiliza el estereotipo desarrollado por el Dr. Sergio Luján Mora, el cual permite utilizar el UML Profile para un Modelamiento Multidimensional.

La tabla 9 muestra los estereotipos a utilizar.

Tabla 9. Estereotipos (Luján, 2005).

Estereotipo	Descripción	Ícono
StarPackage	Un paquete representa una definición del esquema estrella. Se emplea en el nivel 1 y se compone de hechos y dimensiones.	
FactPackage	Un paquete representa un hecho. Se emplea en el nivel 2 y se compone de medidas y dimensiones.	
DimensionPackage	Un paquete representa una dimensión. Se emplea en el nivel 2 y se compone de jerarquías.	
Fact	Es una clase que representa una tabla de hecho. Se emplea en el nivel 3 y se compone de medidas.	
Dimension	Es una clase que representa a una tabla dimensión. Se emplean en el nivel 3 y se compone de jerarquías.	
Base	Es una clase que representa el nivel de jerarquía de una dimensión. Se emplea en el nivel 3.	

7.1. Nivel 1. Esquema Conceptual del Datamart

La figura 43 muestra los diferentes esquemas estrella propuestos para la solución, así como su relación. Se muestra que el esquema de Inventario se relaciona con el esquema de Solicitud Estudio debido a que de estudio comparte una dimensión con el esquema de Inventario, la cual es la dimensión de Fechas.

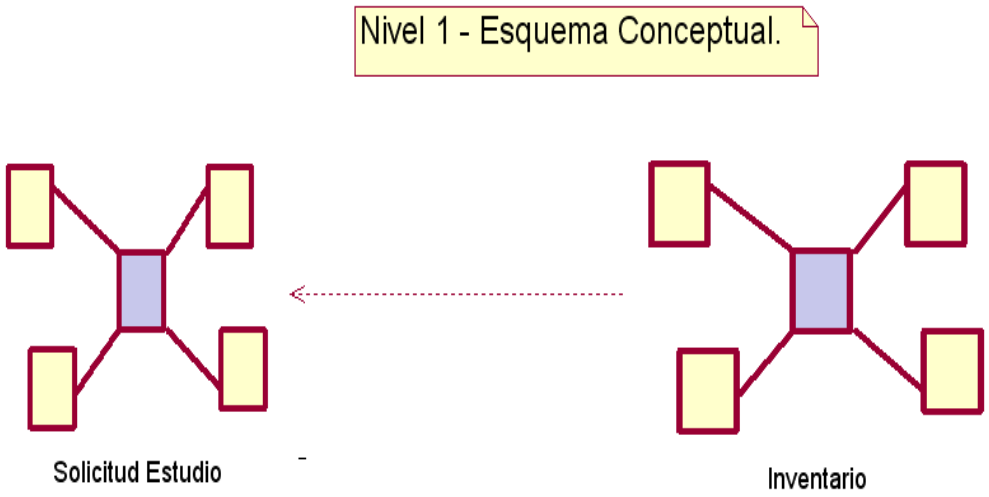


Figura 43. Esquema conceptual del Datamart. Nivel 1 (Elaboración Propia).

7.2. Nivel 2. Esquema Conceptual del Datamart

En el nivel 2 es un nivel con más detalle que el nivel 1. En las figuras 44 y 45 se muestra las diferentes dimensiones y hecho de cada esquema.

7.2.1. Esquema Solicitud Estudio

Consta de una tabla de hechos principal que contiene información referente a las solicitudes y permite satisfacer las diferentes necesidades de información, mencionados en el caso de uso CU001 Análisis de Estudio.

La figura 44 muestra es nivel 2 del esquema estrella de Solicitud Estudio.

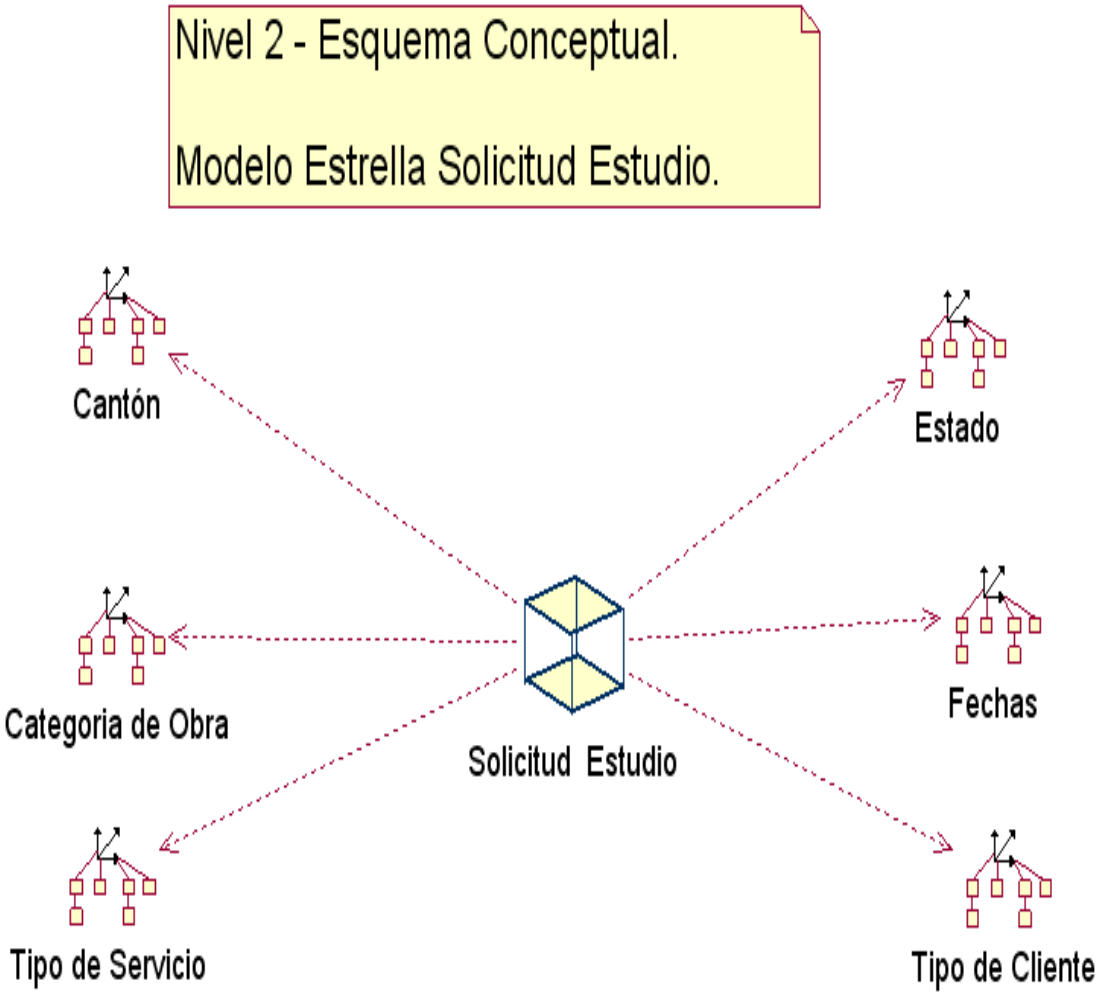


Figura 44. Esquema estrella de Solicitud de Estudio (Elaboración Propia).

7.2.2. Esquema Inventario

Consta de una tabla de hechos principal, la cual contiene información referente a los diferentes materiales utilizados en las diferentes solicitudes de estudios. Esta permite satisfacer las diferentes necesidades de información, mencionados en el caso de uso CU002 Análisis de Materiales.

La figura 45 muestra es nivel 2 del esquema estrella de Inventario.

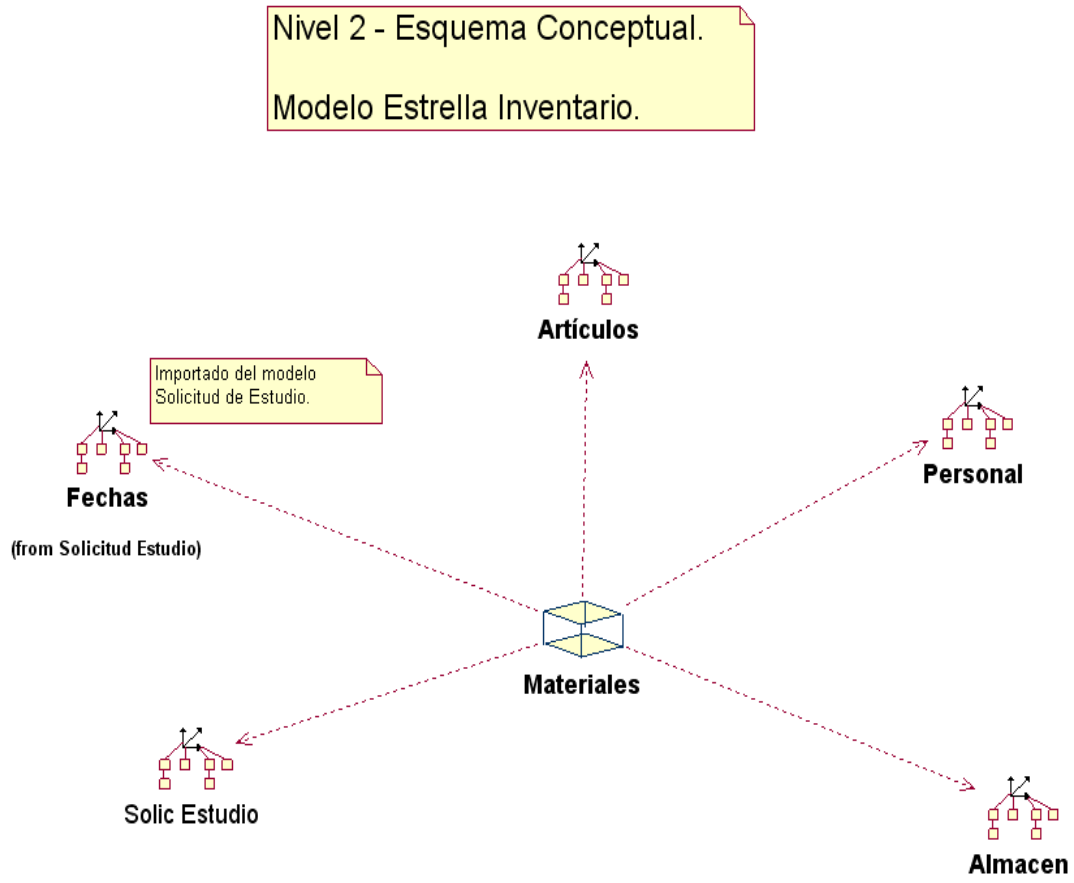


Figura 45. Esquema estrella de Materiales (Elaboración Propia).

7.3. Nivel 3. Esquema Conceptual del Datamart

Detallamos las dimensiones y tablas de hechos mostradas en el nivel 2, el cual permite definir los sus atributos y sus posibles jerarquías. Los diagramas 46 al 45 muestran las diferentes dimensiones con sus respectivos atributos.

7.3.1. Dimensión Cantón

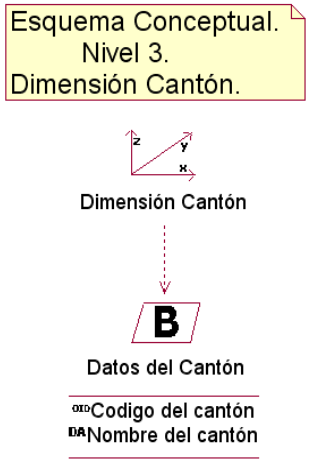


Figura 46. Dimensión Cantón (Elaboración Propia).

7.3.2. Dimensión Estado

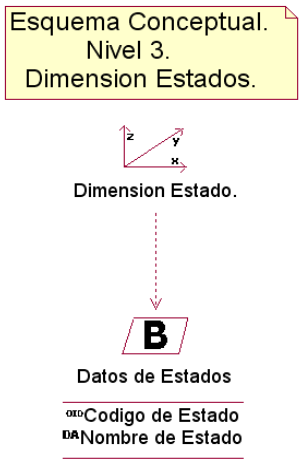


Figura 47. Dimensión Estado (Elaboración Propia).

7.3.3. Dimensión Tipo de Cliente

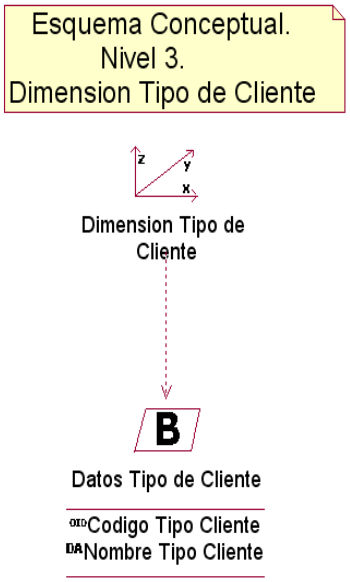


Figura 48. Dimensión Tipo de Cliente (Elaboración Propia).

7.3.4. Dimensión Fechas

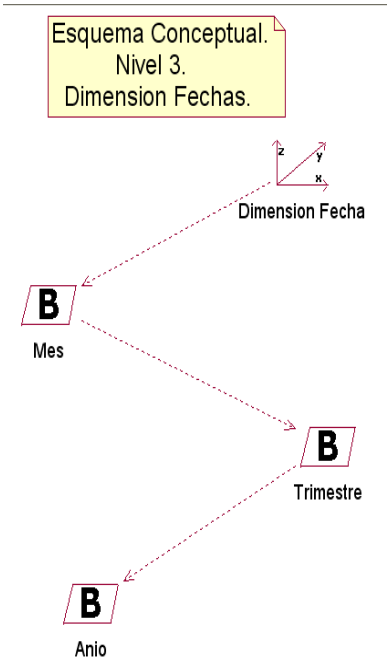


Figura 49. Dimensión Fechas (Elaboración Propia).

7.3.5. Dimensión Tipo de Servicio

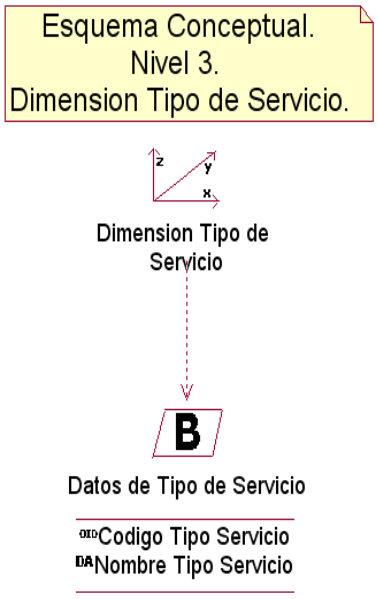


Figura 50. Dimensión Tipo de Servicio (Elaboración Propia).

7.3.6. Dimensión Categoría de Obra

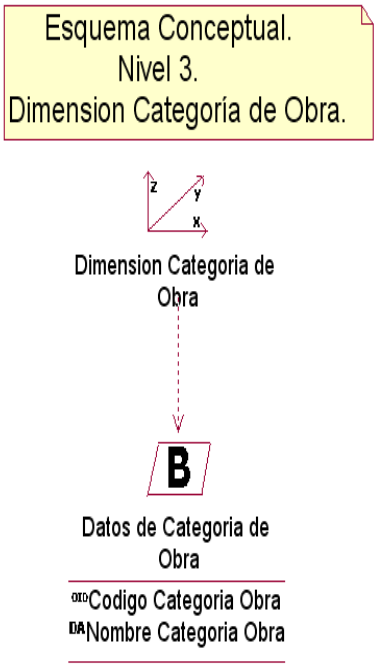


Figura 51. Dimensión Categoría de Obra (Elaboración Propia).

7.3.7. Dimensión de Artículos

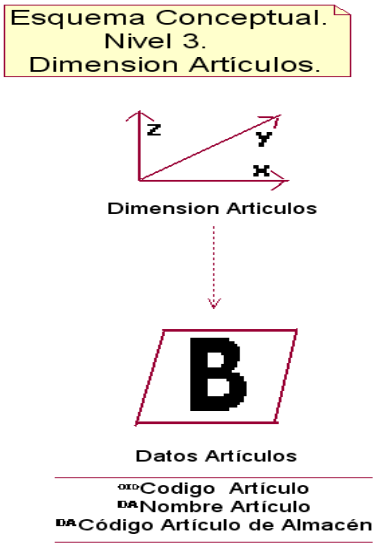


Figura 52. Dimensión Artículos (Elaboración Propia).

7.3.8. Dimensión Personas

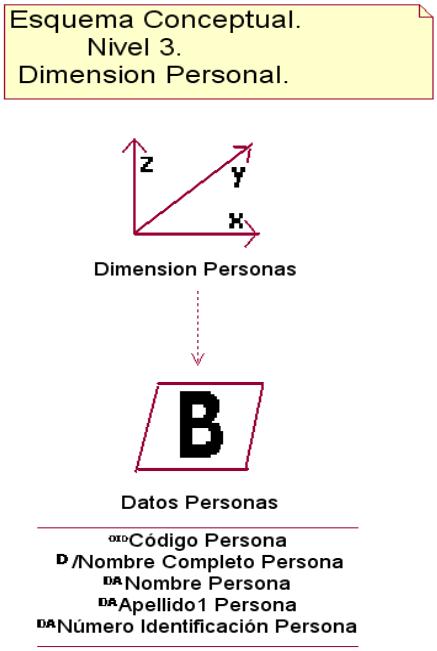


Figura 53. Dimensión Personas (Elaboración Propia).

7.3.9. Dimensión Almacén

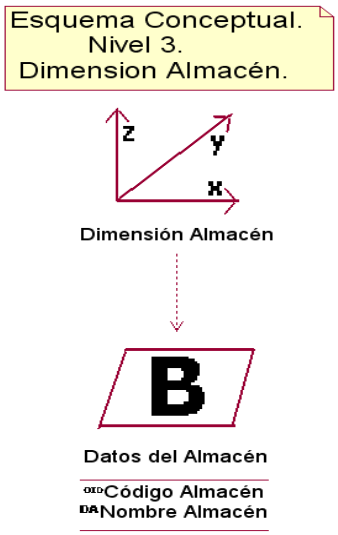


Figura 54. Dimensión Almacén (Elaboración Propia).

7.3.10. Dimensión Solic. Estudio

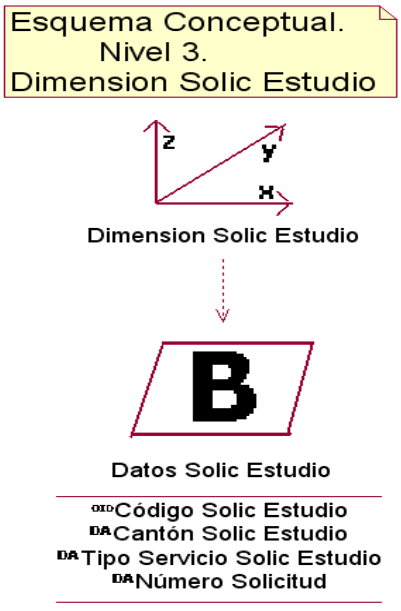


Figura 55. Dimensión Solic. Estudio (Elaboración Propia).

Los diagramas 56 y 57 muestra la relación de la tabla de hechos que se da entre la tabla de hechos con las diferentes dimensiones para los modelos Solicitud Estudio e Inventario.

7.3.11. Hecho Fact_SolicEstudio

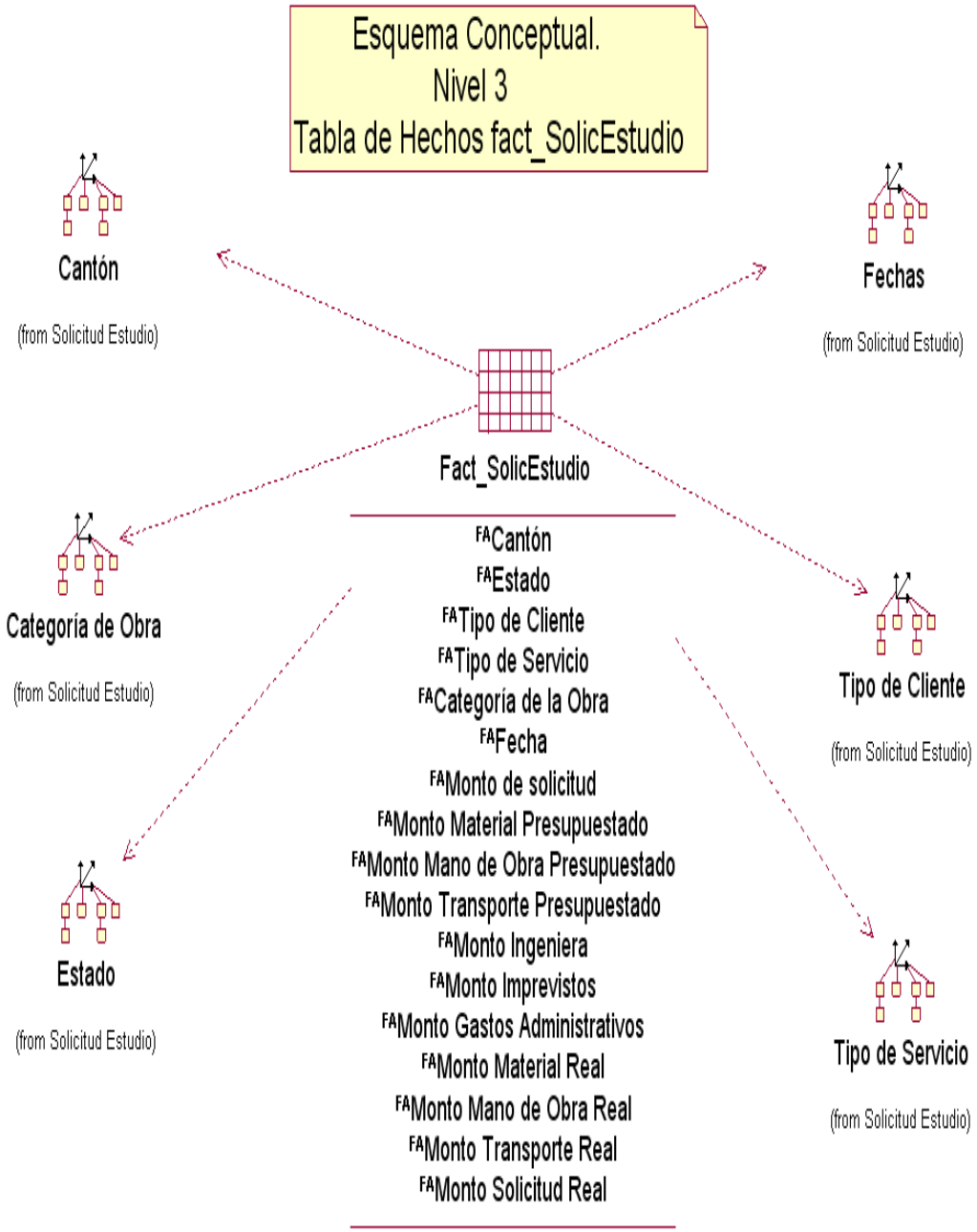


Figura 56. Esquema Conceptual de la Tabla de Hechos Fact_SolicEstudio (Elaboración Propia).

7.3.12. Hecho Fact_Material

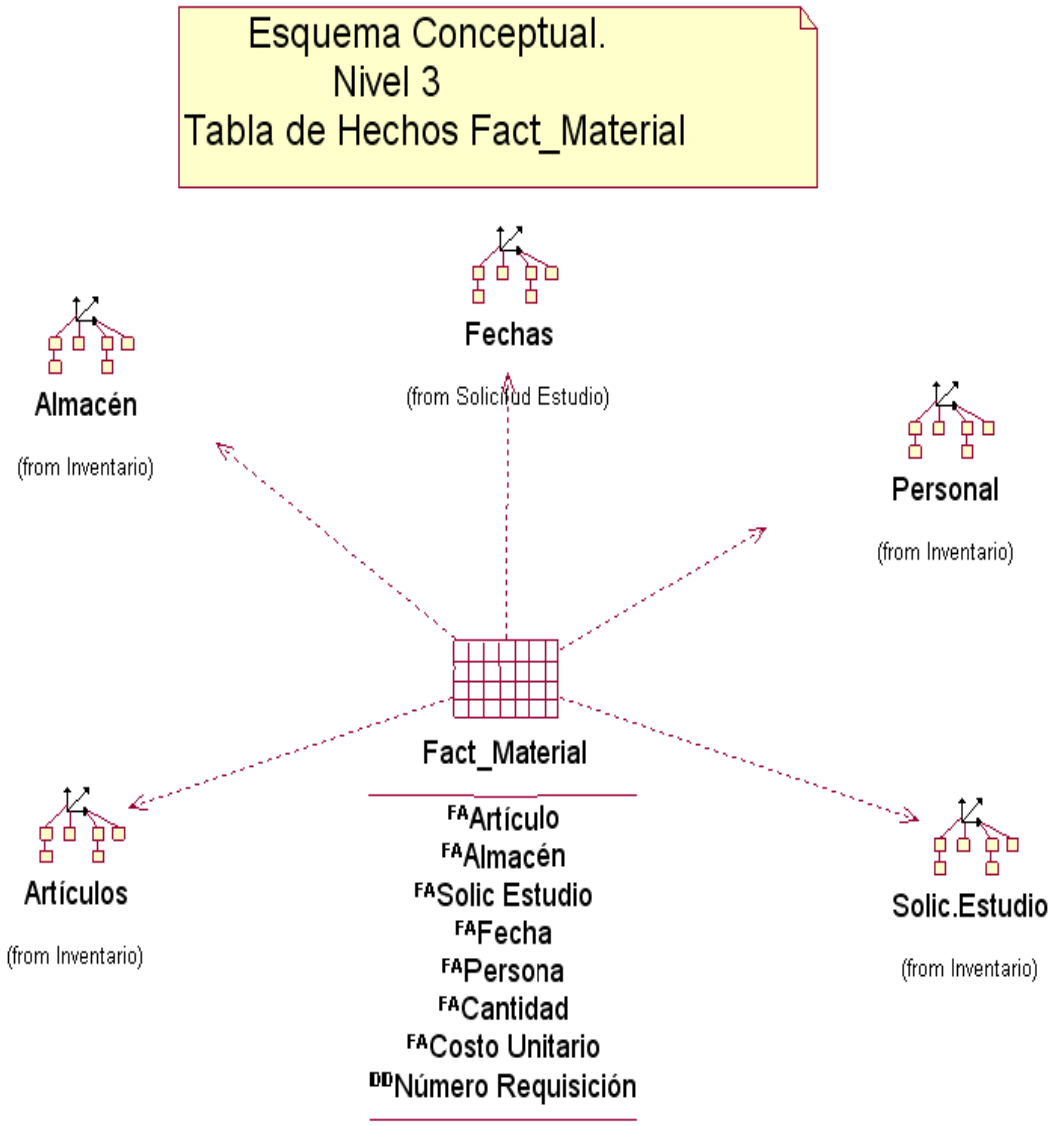


Figura 57. Esquema Conceptual de la Tabla de Hechos Fact_Material (Elaboración Propia).

ANÁLISIS DIMENSIONAL

7.4. Análisis de Temas

El Datamart se compone de dos grandes temas:

- Solicitud de Estudios.
- Materiales.

7.4.1. Solicitud de Estudios

Se debe analizar las solicitudes de estudio que se generan diariamente, el objetivo buscado es conocer el detalle de las solicitudes de estudio realizadas y solicitadas al procedo de Planificar y Desarrollar la Red.

El resultado es medido por:

- Cantón donde se debe ejecutar la obra.
- Tipo de servicio solicitado.
- Categoría de la obra a construir.
- Tipo de cliente que solicita la obra.
- Estado de la solicitud.
- Mes, trimestre, año.
- Monto de la solicitud.

7.4.2. Materiales

Se debe analizar el consumo de material con el objetivo de mantener todo tipo de material necesario para el desarrollo oportuno de la red eléctrica.

El resultado es medido por:

- Artículo que es requisado.
- Tipo de servicio en que se utiliza el artículo.
- Cantón donde se utiliza el artículo.
- Cantidad y costo unitario del artículo.

7.5. Dimensiones

El Datamart se conforma de las siguientes dimensiones:

- Cantón.
- Estado.
- Tipo de Cliente.
- Fecha.
- Tipo de Servicio.
- Categoría de Obra.
- Artículos.
- Solic. Estudio.
- Almacén.
- Personal.

7.5.1. Cantón

Descripción

Cada registro representa datos de los cantones donde se realizan las obras.

Atributos

Atributo	Descripción	Formato
Código del cantón.	Código del cantón.	0000000000
Cantón	Nombre del cantón.	Mayúsculas.

7.5.2. Estado

Descripción

Cada registro representa el estado de la solicitud, tal como registrado, ejecutado o aprobado, entre otros.

Atributos.

Atributo	Descripción	Formato
Código de Estado.	Código de estado	00
Estado	Nombre del estado.	Mayúsculas.

7.5.3. Tipo de Cliente

Descripción

Cada registro representa el tipo de cliente que realiza una solicitud de estudio, tal como cliente físico, jurídico o nacional, entre otros.

Atributos

Atributo	Descripción	Formato
Código del tipo de cliente.	Código del tipo de cliente	00
Tipo de cliente	Nombre del tipo de cliente.	Mayúsculas.

7.5.4. Fecha

Descripción

Cada registro representa una fecha.

Jerarquía:

Mes -> Trimestre ->Año.

Vemos que la jerarquía se compone de 4 niveles:

Nivel	Atributos
1	Fecha Día. Día de la semana.
2	Mes. Número de mes. Año y mes.
3	Trimestre. Año y trimestre.
4	Año

Atributos.

Atributo	Descripción	Formato
Fecha	Fecha	DD/MM/YYYY
Número del Día	Día del mes.	00
Día de la semana	Día de la semana: Lunes,	Mayúscula

	Martes, …Domingo	
Número del Mes	Número de mes del año 1..12	00
Nombre del mes	Nombre del mes: Enero..Diciembre	Mayúscula
Año y mes	Concatenación del año con el mes, por ejemplo 2013-10	yyyy-mm
Trimestre	Número del trimestre del año. TRIM01, TRIM02, TRIM03	TRIM00
Año y trimestre	Concatenación del año con el trimestre, por ejemplo, 2013-TRIM03	yyyy-TRIM00
Año	Número de año.	yyyy

7.5.5. Tipo de servicio

Cada registro representa un el tipo de servicio solicitado, tal como: Residencial o Comercial, entre otros.

Atributos

Atributo	Descripción	Formato
Código del tipo de Servicio.	Código del tipo de Servicio.	00
Nombre del tipo de servicio.	Nombre del tipo de servicio.	Mayúsculas.

7.5.6. Categoría de la obra

Cada registro representa un la categoría de la obra, tal como: Mejora o Particular, entre otros.

Atributos

Atributo	Descripción	Formato
Código de categoría	Código de la categoría de la obra.	00
Nombre de categoría.	Nombre de la categoría de la obra.	Mayúsculas.

7.5.7. Artículos

Cada registro representa información de los artículos utilizados.

Atributos

Atributo	Descripción	Formato
Código de artículo.	Código de artículo.	0000000
Nombre	Nombre del artículo.	Mayúsculas.
Código Artículo de Almacén	Código del artículo con que se registra en el almacén.	0000000000000.00

7.5.8. Solic. Estudio

Cada registro representa información de la solicitud de estudio donde se utilizó el material.

Atributos

Atributo	Descripción	Formato
Código de solicitud	Código de solicitud de estudio.	0000000
Cantón	Nombre del Cantón.	Mayúsculas.
Tipo de servicio	Nombre del tipo de servicio	Mayúsculas.

7.5.9. Almacén

Cada registro representa información del almacén de donde se retiró el material.

Atributos

Atributo	Descripción	Formato
Código de almacén	Código de almacén	0000000
Nombre de almacén.	Nombre de almacén.	Mayúsculas.

7.5.10. Personal

Cada registro representa información del personal que realizó el retiro del material.

Atributos

Atributo	Descripción	Formato
Código de personal	Código de personal	0000000
Número de identificación	Número de cédula o cédula de residencia.	000000000
Nombre	Nombre de la persona	Mayúsculas.

7.6. Hechos (Facts)

7.6.1. Facts/Dimensiones

	Hechos	
	Solicitud de Estudio	Materiales
Cantones	X	
Estados	X	
Tipo de Cliente	X	
Fechas	X	X
Tipo de Servicio	X	
Categoría de la Obra	X	
Solicitud de Estudio		X
Artículos		X
Personal		X
Almacén		X

7.6.2 Hecho Solicitud de Estudio

Cada registro representa un detalle de la solicitud de estudio.

Granularidad

Dimensión	Descripción	Llave Primaria
Cantón	Código de cantón en que se realiza el trabajo asociado a una solicitud.	✓
Estado	Estado en que se encuentra la solicitud.	✓
Tipo de Cliente	Tipo de cliente que realiza la solicitud.	✓
Tipo de servicio.	Tipo de servicio solicitado.	✓
Categoría de la obra.	Categoría en que se clasifica la obra.	✓
Fecha	Fecha en que se solicita la obra.	✓

Medidas

Nombre	Descripción
Monto de solicitud	Precio cobrado por la solicitud de estudio.
Monto prima.	Monto de la prima realizada por el cliente.
Monto pagaré.	Monto total a financiar al cliente.
Monto del estudio.	Monto del estudio de campo, que indique el costo

	total de la obra para el cliente.
Monto construcción	Monto que le cuesta a JASEC construir la obra.

7.6.3 Hecho Inventario

Cada registro representa un detalle de cada requisición de materiales asociado a la solicitud de estudio.

Granularidad

Dimensión	Descripción	Llave Primaria
Artículo	Código de artículo requisado.	✓
Almacén	Código de almacén de donde requisa el artículo.	✓
Solic. Estudio	Número de solicitud en la cual se va a utilizar el artículo.	✓
Número de requisición	Número de requisición con que sale el artículo del almacén.	✓
Fecha	Fecha en que se solicita la requisición.	✓
Personal	Código de la persona que retira el material del almacén.	✓

Medidas

Nombre	Descripción
Cantidad	Cantidad requisada de un artículo.
Costo unitario	Costo unitario con que se requisa el artículo.

Apéndice #8. Datos Fuente. Esquema Conceptual y Lógico

Este apéndice muestra la estructura de la información fuente del Datamart.

Apegado a la metodología DWEP, se utilizará los diagramas Esquema Conceptual de Datos Fuente (SCS, figura 58) y Esquema Lógico de Datos Fuente (SLS, figura 59). El SCS muestra a nivel conceptual la forma en que se organiza la estructura donde está alojada la información fuente del Datamart. SLS muestra a nivel lógico la estructura donde está alojada la información fuente del Datamart.

Datos Fuente, Esquema Conceptual (SCS)

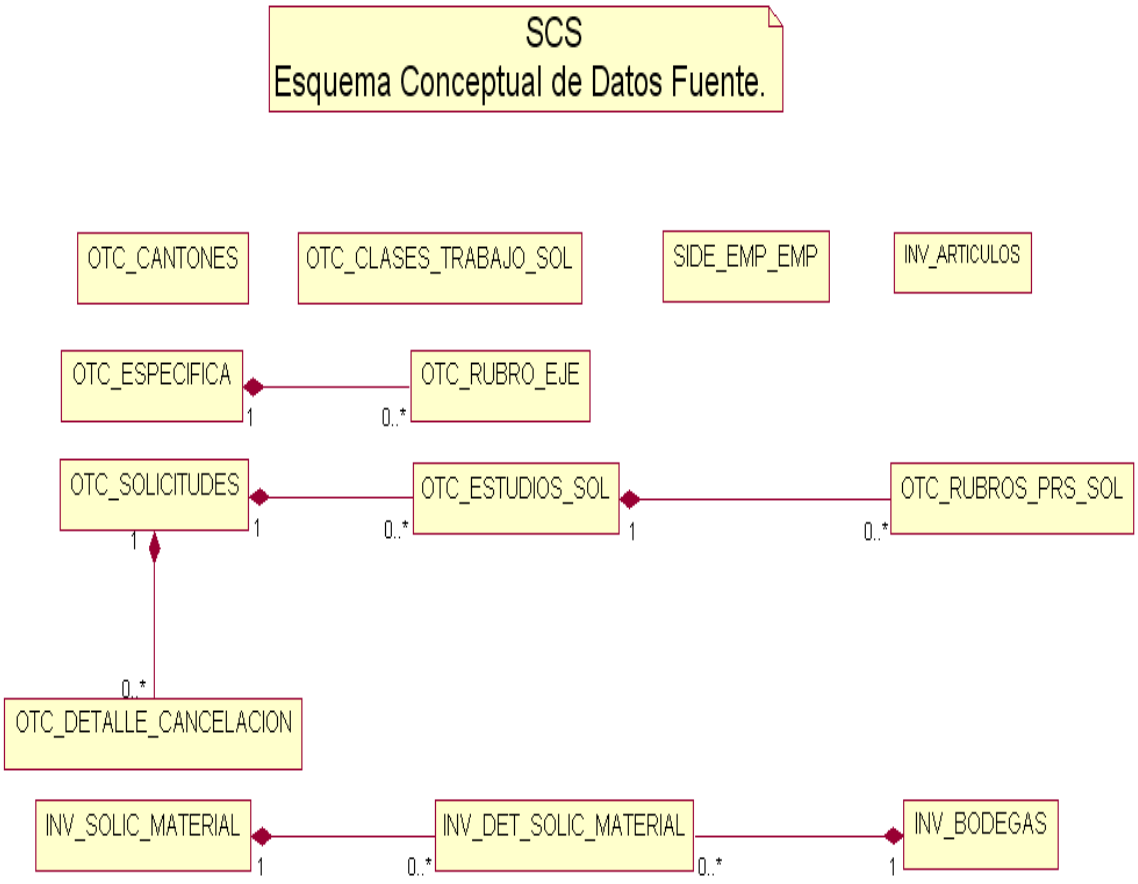


Figura 58. Diagrama SCS (Elaboración Propia).

Datos Fuente, Esquema Lógico (SLS)

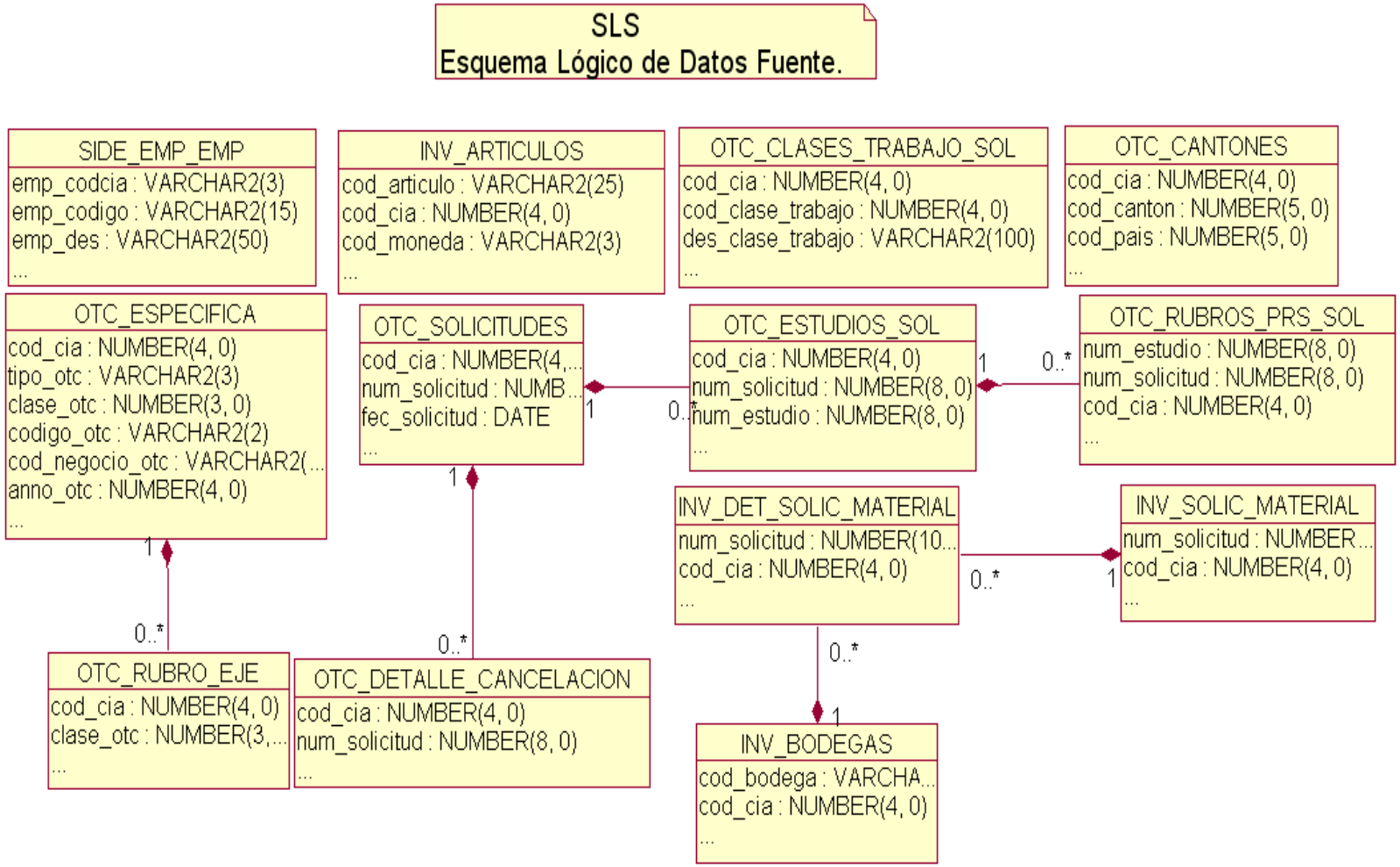


Figura 59. Diagrama SLS (Elaboración Propia).

Apéndice #9. Mapeo de Datos

En este apéndice se muestra el mapeo de información que se dará entre la fuente de datos y el Datamart. A través de la metodología DWEP, se utilizará un diagrama llamado Mapeo de Datos.

9.1. Diagrama de Mapeo de Datos

Al igual que los diagramas anteriores, se extiende el UML, y se centra en el flujo de datos y la interconexión que se da entre las bases de datos que contiene la información y fuente, y la base de datos que contendrá la información del Datamart. Se empleará la noción de mapeo para capturar la interconexión entre los elementos. El mapeo consta de tres componentes lógicos:

- Entidad Proveedora: Es el responsable de generar datos. Puede ser un esquema, tabla o atributo.
- Consumidor: Recibe el dato del proveedor.
- Intermediario: Indica cómo se debe realizar el mapeo, lo que incluye transformaciones y filtrados.

El diagrama da la flexibilidad de definir el modelo a varios niveles de detalle. En nuestro caso, lo definiremos a cuatro niveles, donde cada nivel es un nivel más detallado que el nivel anterior.

Nivel de base de datos (0): Cada esquema del DW o Datamart se representa como un paquete. Los mapeos entre los diferentes esquemas son modelados por un único paquete, el cual encapsula todos los detalles.

Nivel de flujo de datos (1): Permite describir las relaciones de datos en forma individual entre la fuente de datos y su destino.

Nivel de tabla (2): Describe las relaciones a nivel de tabla, y muestra detalles como el tipo de transformación.

Nivel de atributo (3): Describe el mapeo a nivel de atributos.

En la figura 60 muestra la relación entre los diferentes niveles. El nivel 0 muestra la relación existente entre el Modelo Conceptual del Dataware House (DWCS, apéndice 10) y Modelo Conceptual Fuente (SCS), el cual es realizado por el paquete Data Mapping. Se muestra que hay tres paquetes de mapeo, y se podría asumir que hay tres tablas en el Datamart (Nivel 1). En el nivel 2 es una vista más detallada de un paquete del nivel 1, el cual muestra la relación dada entre la fuente y el destino, con sus respectivas transformaciones, las cuales se representas por paquetes (paso 1 y paso 2), además de mostrar un almacén de datos intermedio que cual permite almacenar la transformación realizada. El nivel 3 muestra el mapeo realizado en un paquete de transformaciones del nivel 2, mostrando el mapeo a realizarse entre la fuente y el almacén de datos intermedio.

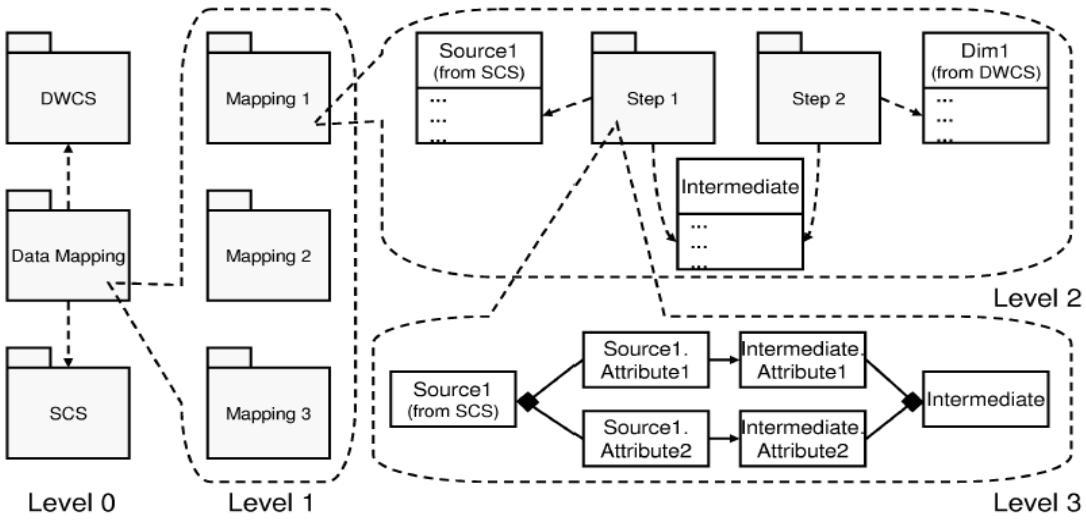


Figura 60. Diagrama de Mapeo. Niveles (Luján, 2005).

9.1.1. Nivel 0.

Es el nivel de base de datos, la figura 61 muestra la relación entre la base de datos que almacena la información fuente (SCS) y la base de datos del Datamart (DWCS – Estudio Solicitud y de Inventario).

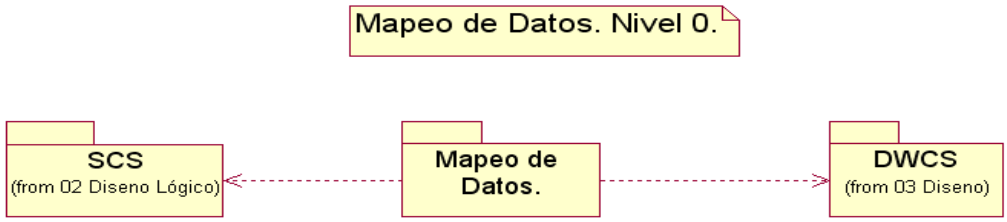


Figura 61. Mapeo de Datos. Nivel 0 (Elaboración Propia).

9.1.2. Nivel 1.

La propuesta de diez dimensiones, de las cuales se mapearan nueve, ya que la tabla de fechas no tiene su correspondiente tabla fuente, su contenido se cargara por medio de una función. La figura 62 muestra el diagrama de Mapeo de Datos a nivel 1.

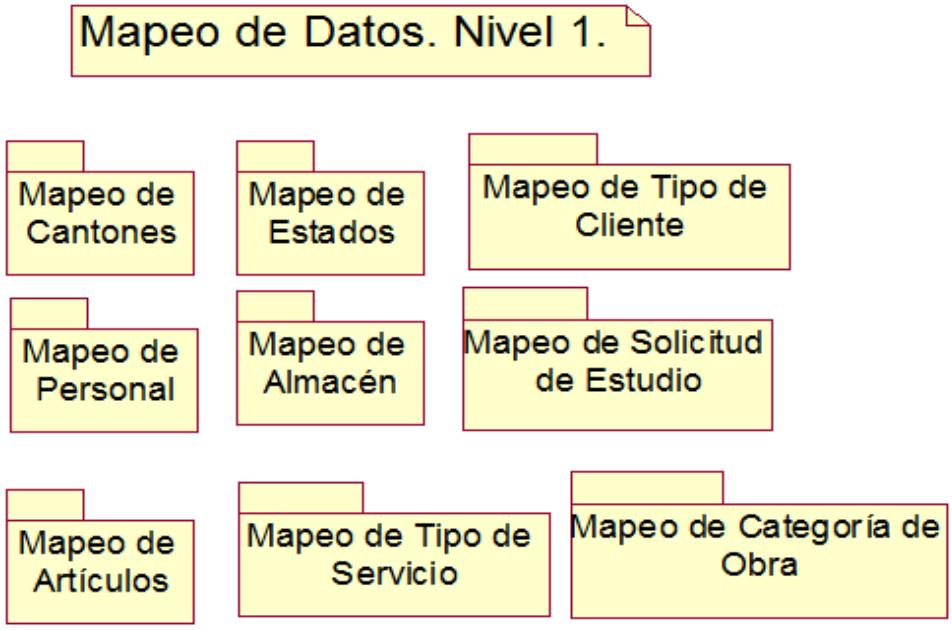


Figura 62. Mapeo de Datos. Nivel 1 (Elaboración Propia).

9.1.3. Nivel 2.

Este nivel detalla cada paquete de mapeo del nivel 1. En este nivel se muestra las tablas que intervienen (Fuente y Destino).

La figura 63 muestra el Mapeo de Cantones.

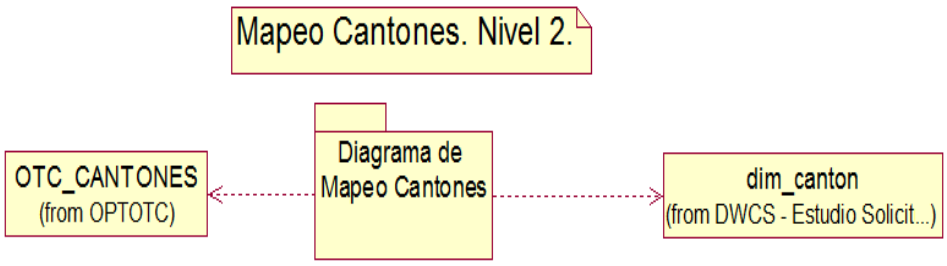


Figura 63. Mapeo de Datos. Nivel 2. Cantones (Elaboración Propia).

La figura 64 muestra el Mapeo de Estados.

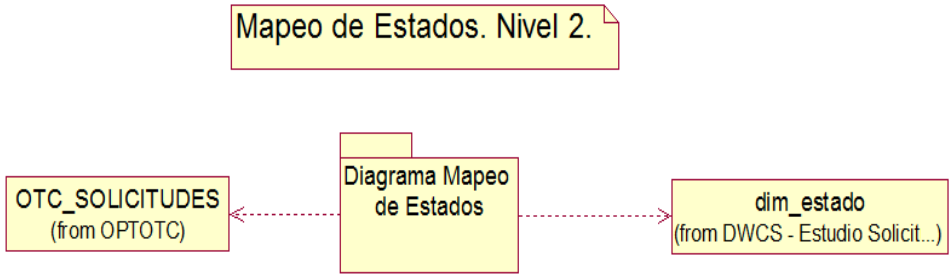


Figura 64. Mapeo de Datos. Nivel 2. Estados (Elaboración Propia).

La figura 65 muestra el Mapeo de Tipo de Cliente.

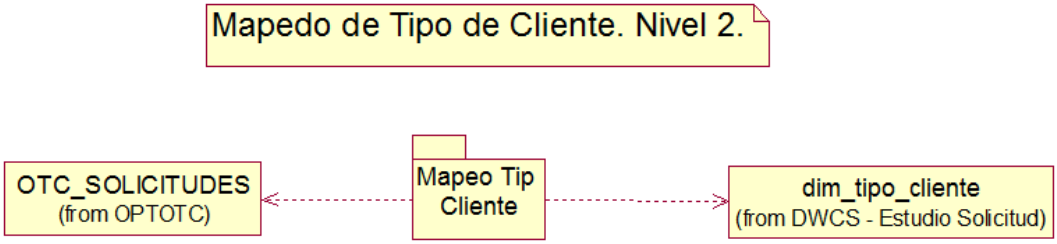


Figura 65. Mapeo de Datos. Nivel 2. Tipo de Cliente (Elaboración Propia).

La figura 66 muestra el Mapeo Personal.

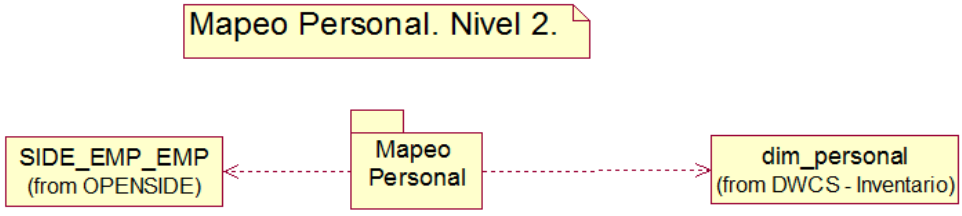


Figura 66. Mapeo de Datos. Nivel 2. Personal (Elaboración Propia).

La figura 67 muestra el Mapeo Almacén.

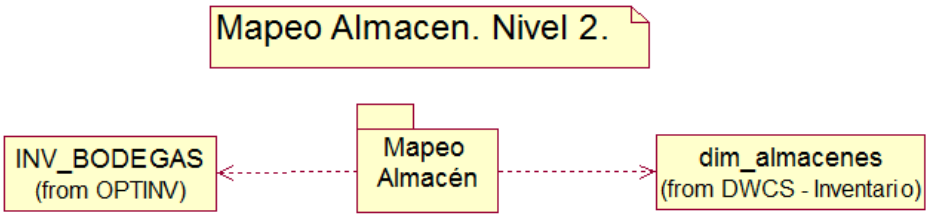


Figura 67. Mapeo de Datos. Nivel 2. Almacén (Elaboración Propia).

La figura 68 muestra el Mapeo Estudio Solicitud.

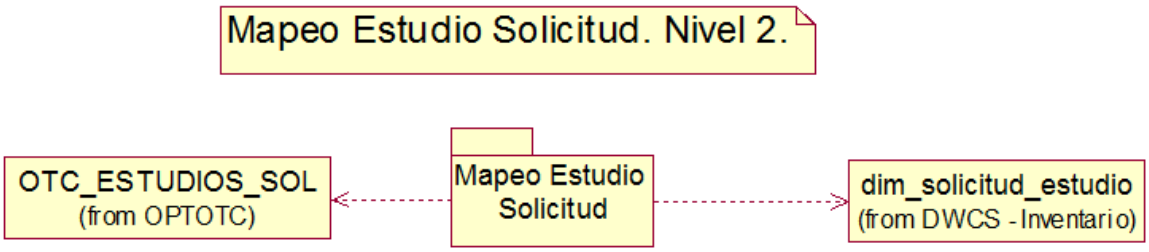


Figura 68. Mapeo de Datos. Nivel 2. Estudio Solicitud (Elaboración Propia).

La figura 69 muestra el Mapeo de Artículos.

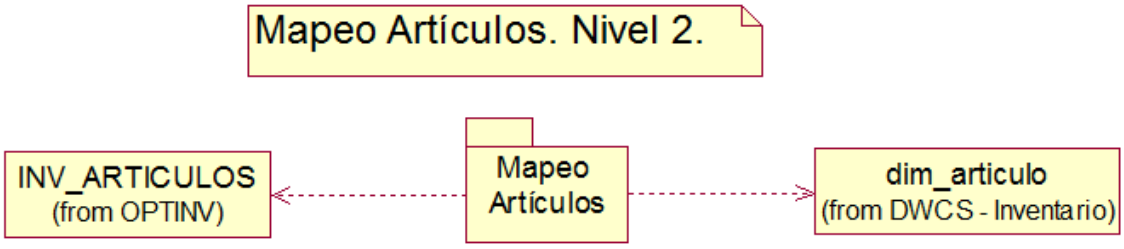


Figura 69. Mapeo de Datos. Nivel 2. Artículos (Elaboración Propia).

La figura 70 muestra el Mapeo de Tipo de Servicio.

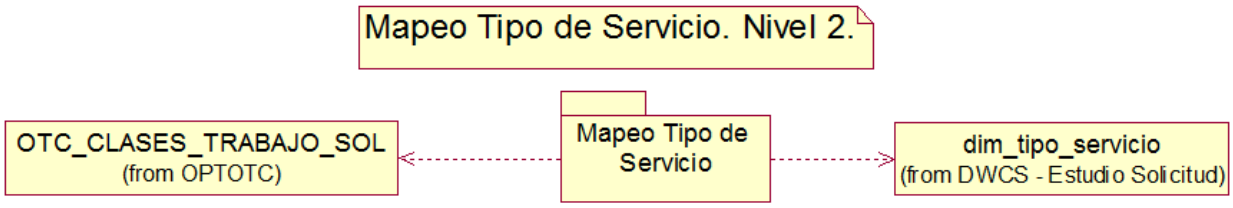


Figura 70. Mapeo de Datos. Nivel 2. Tipo de Servicio (Elaboración Propia).

La figura 71 muestra el Mapeo de Categoría de Obra.

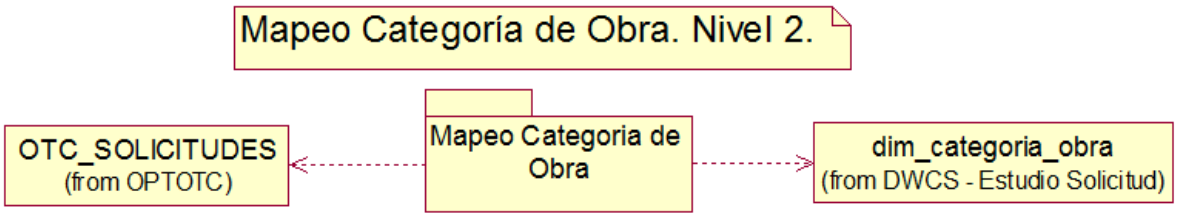


Figura 71. Mapeo de Datos. Nivel 2. Categoría de Obra (Elaboración Propia).

9.1.4. Nivel 3.

El mapeo se realiza a nivel de atributos, entre la base de datos fuente y la base de datos del Datamart. En el nivel 2 existen paquetes llamados Mapeo <nombre>, en el nivel 3 se detalla este tipo de paquete.

La figura 72 muestra el diagrama de mapeo de cantones, nivel 3.

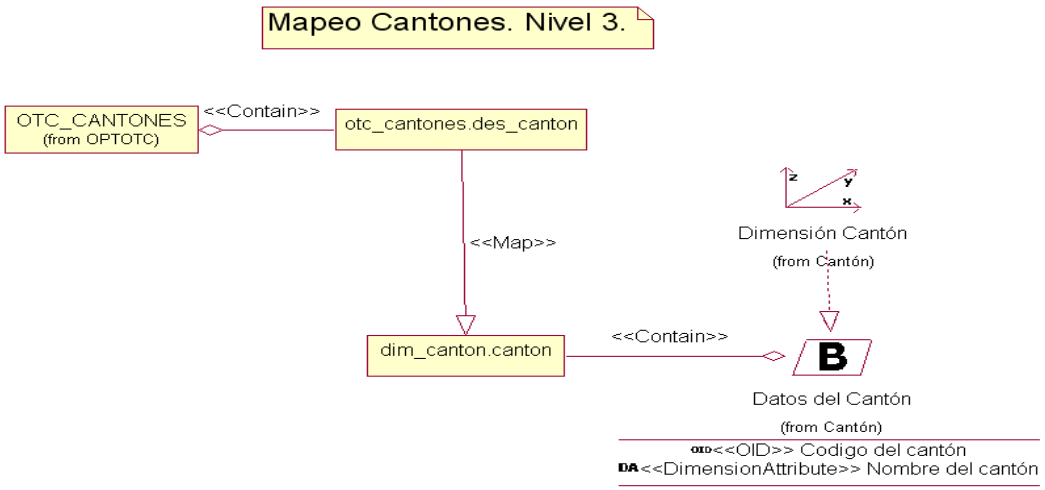


Figura 72. Mapeo de Datos. Cantones. Nivel 3 (Elaboración Propia).

La figura 73 muestra el diagrama de mapeo de estados, nivel 3.

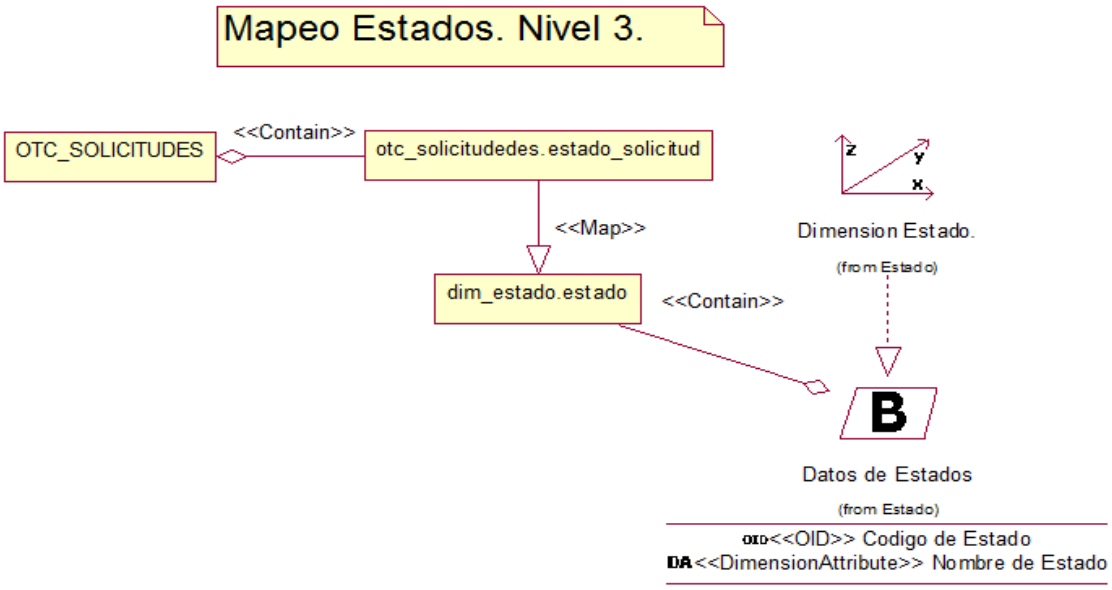


Figura 73. Mapeo de Datos. Estados. Nivel 3 (Elaboración Propia).

La figura 74 muestra el diagrama de mapeo de Tipo de Cliente, nivel 3.

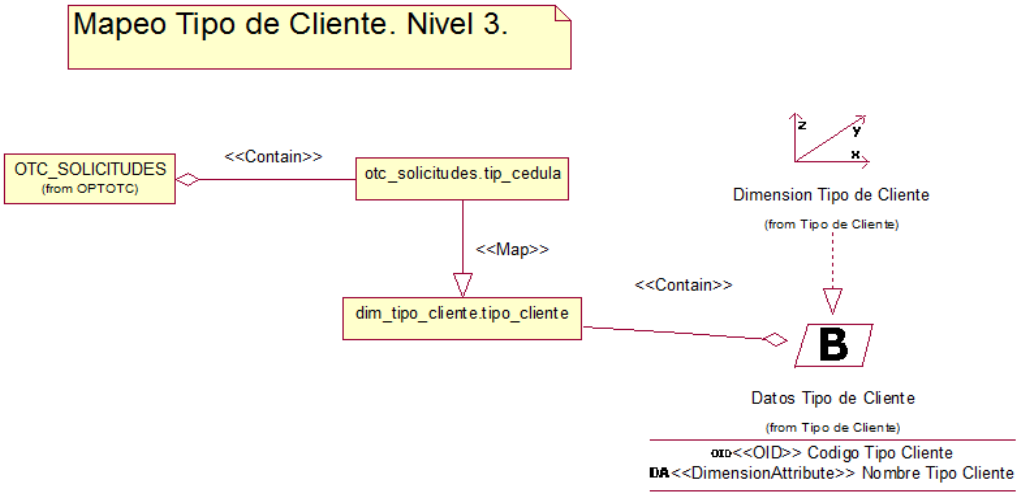


Figura 74. Mapeo de Datos. Tipo de Cliente. Nivel 3 (Elaboración Propia).

La figura 75 muestra el diagrama de mapeo de Personal, nivel 3.

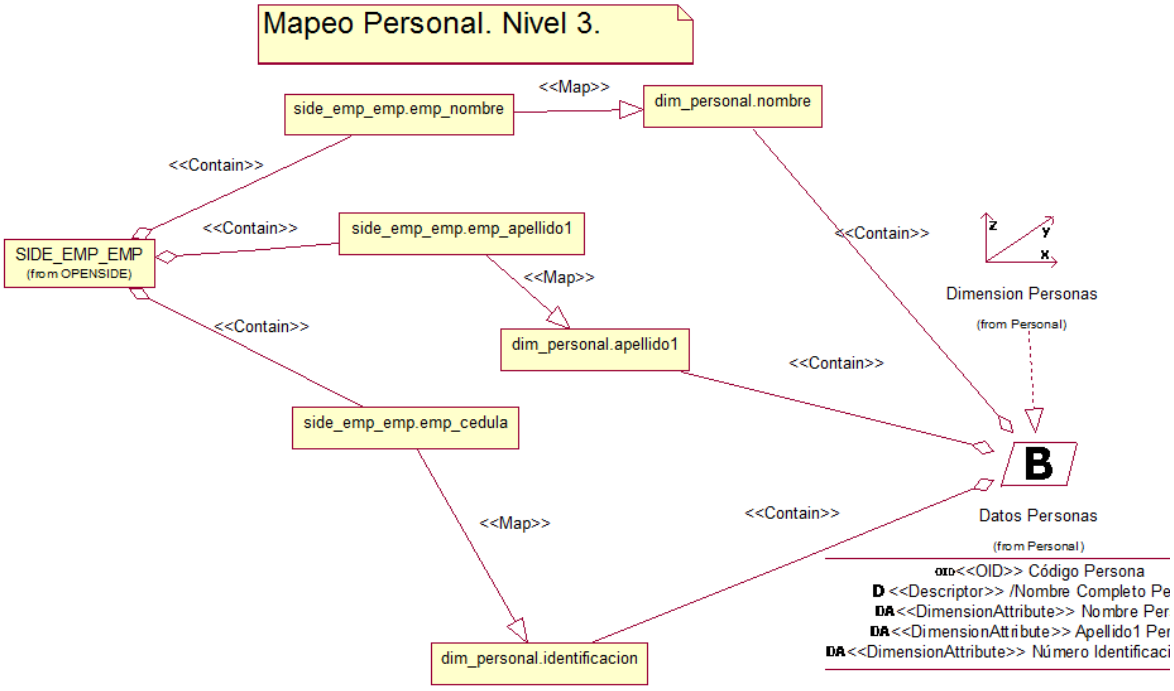


Figura 75. Mapeo de Datos. Personal. Nivel 3 (Elaboración Propia).

La figura 76 muestra el diagrama de mapeo de Almacén, nivel 3.

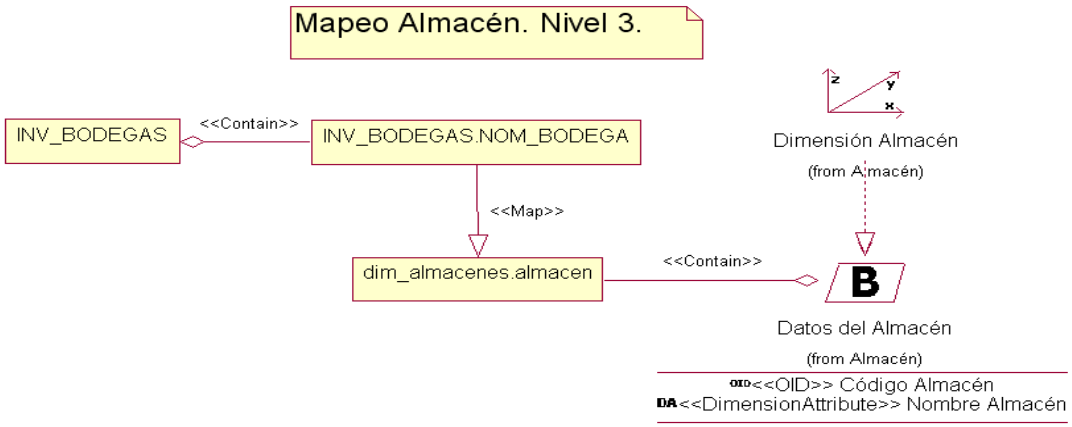


Figura 76. Mapeo de Datos. Almacén. Nivel 3 (Elaboración Propia).

La figura 77 muestra el diagrama de mapeo de Estudio Solicitud, nivel 3.

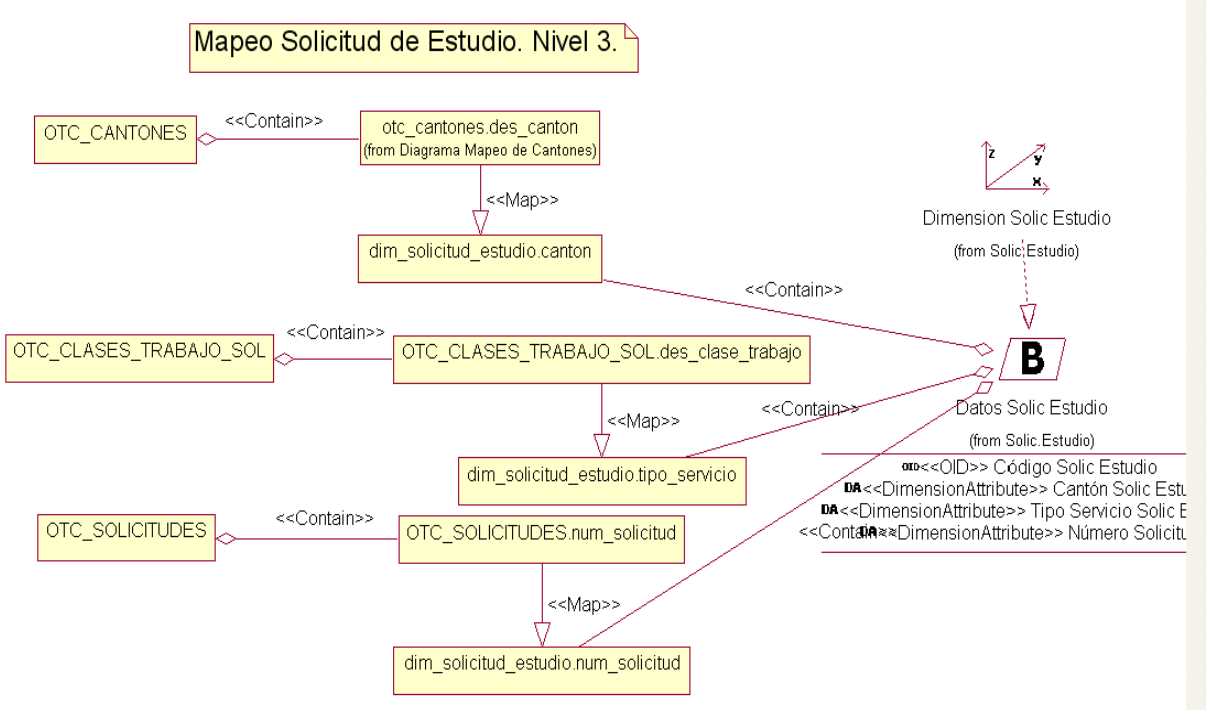


Figura 77. Mapeo de Datos. Estudio Solicitud. Nivel 3 (Elaboración Propia).

La figura 78 muestra el diagrama de mapeo de Artículos, nivel 3.

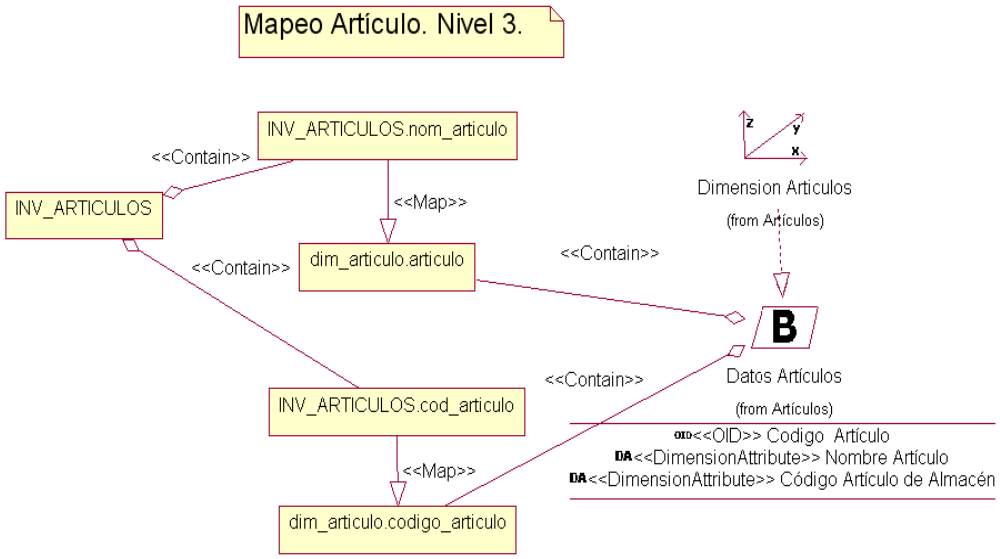


Figura 78. Mapeo de Datos. Artículos. Nivel 3 (Elaboración Propia).

La figura 79 muestra el diagrama de mapeo de Tipo de Servicio, nivel 3.

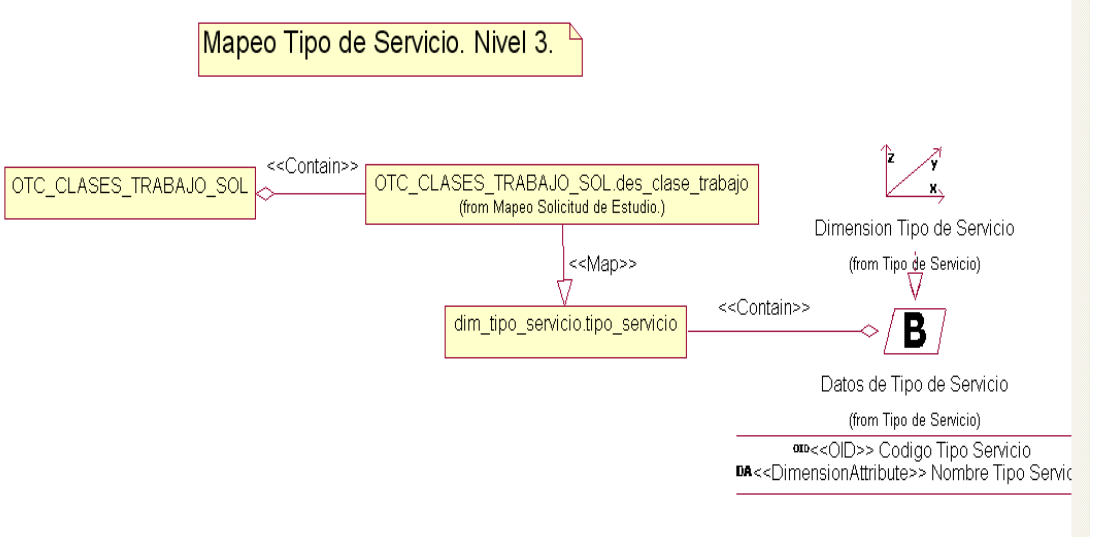


Figura 79. Mapeo de Datos. Tipo de Servicio. Nivel 3 (Elaboración Propia).

La figura 80 muestra el diagrama de mapeo de Categoría de Obra, nivel 3.

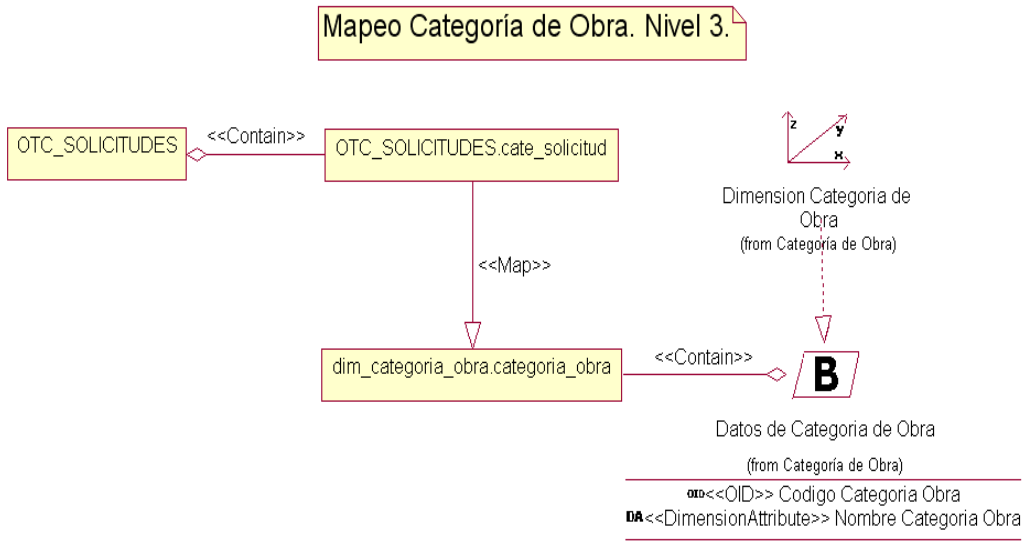


Figura 80. Mapeo de Datos. Categoría de Obra. Nivel 3 (Elaboración Propia).

Apéndice #10. Esquema Conceptual del Datamart (DWCS)/ Esquema Lógico del Datamart(DWLS).

10.1. Esquema Conceptual del Datamart (DWCS).

La figura 81 muestra el DWCS del esquema Solicitud de Estudio y de Inventario.

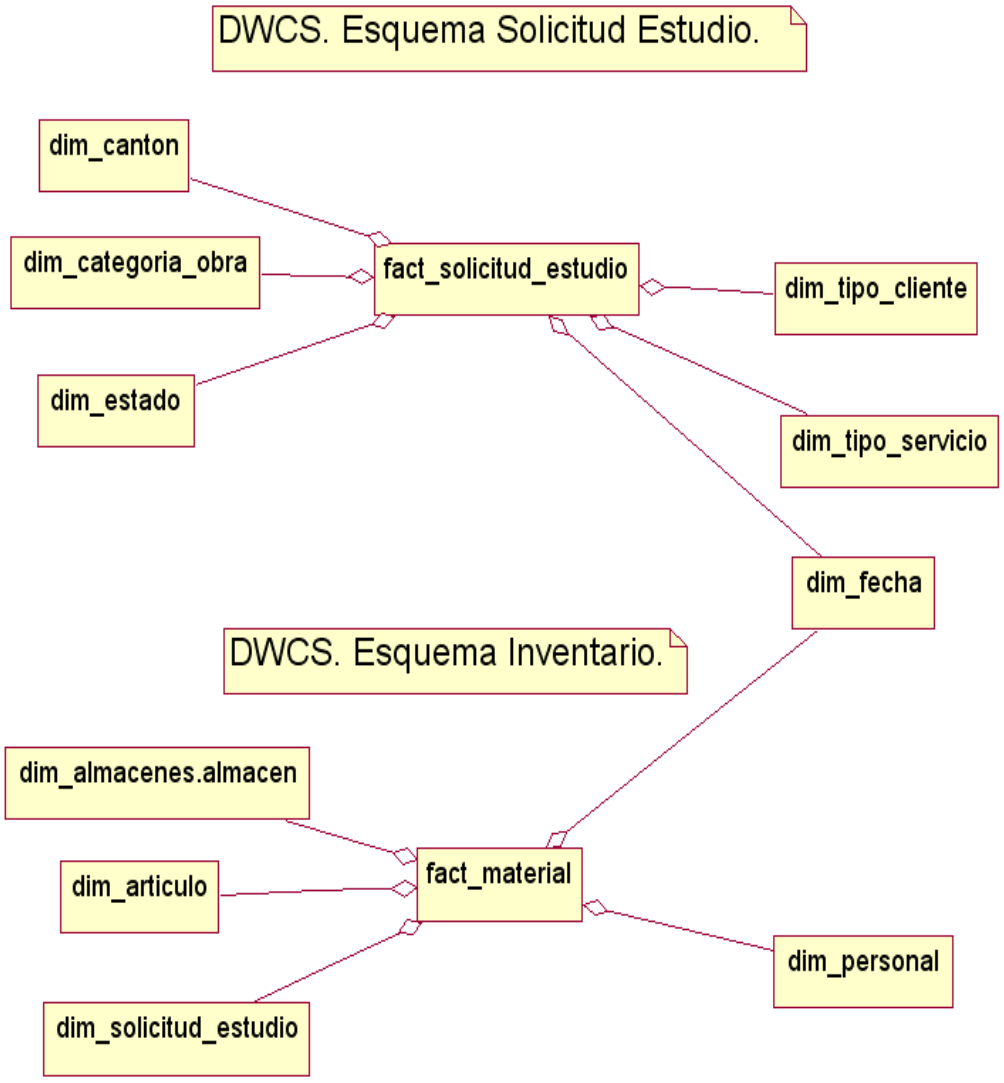


Figura 81. DWCS – Esquema Solicitud Estudio/Inventario (Elaboración Propia).

10.2. Esquema Lógico del Datamart (DWLS).

La figura 82 muestra el DWLS del esquema Solicitud Estudio.

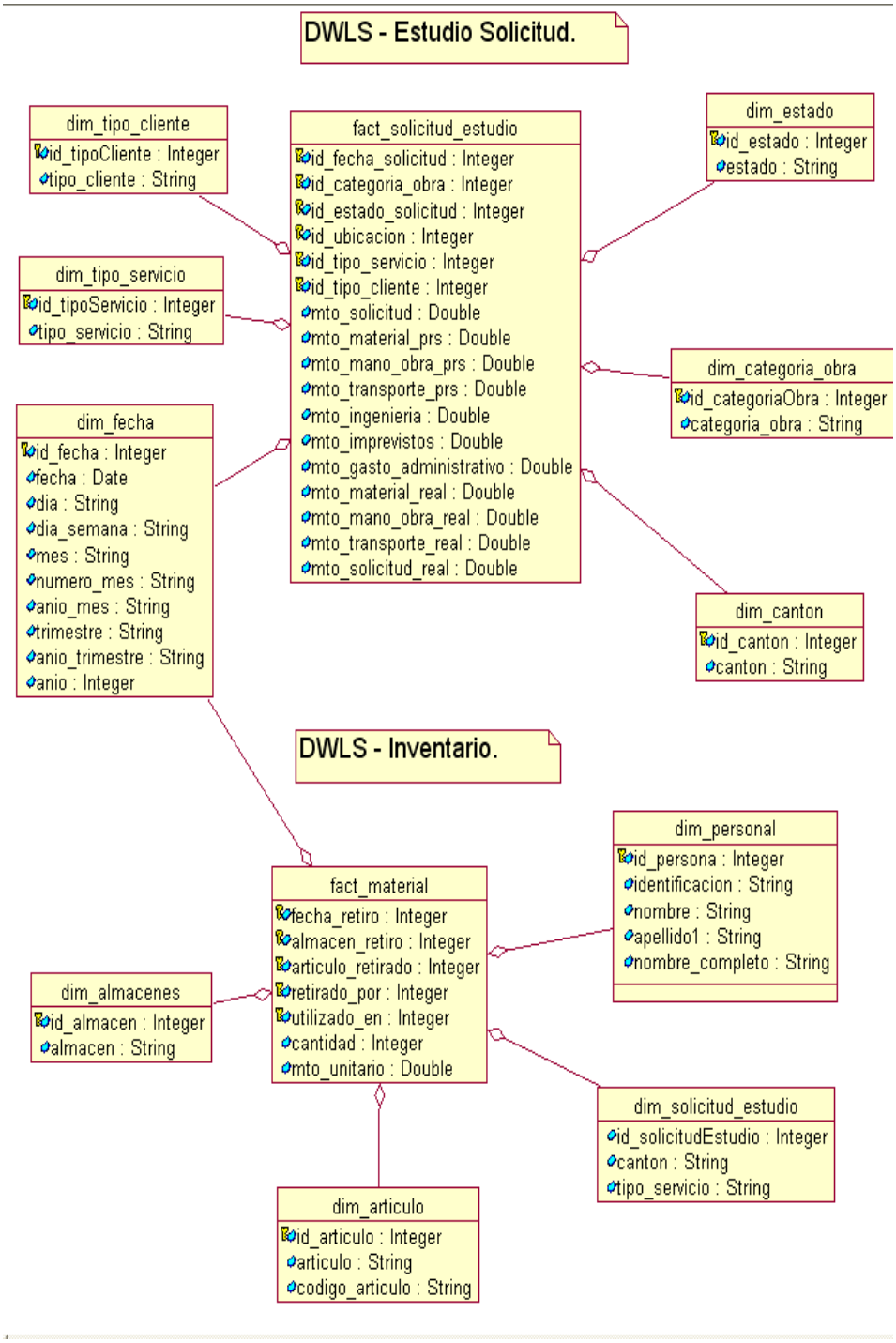


Figura 82. DWLS – Solicitud Estudio (Elaboración Propia).

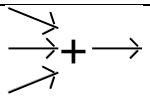
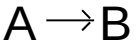
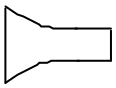
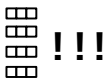
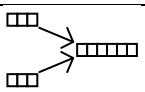
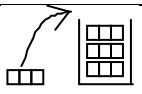
Apéndice #11: Procesos ETL.

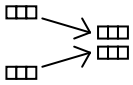
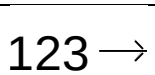
Este apéndice permite detallar el diseño de los procesos ETL, se utilizará la notación propuesta por la metodología DWEP, creándose diagramas de Proceso ETL.

11.1 Diagrama de Proceso ETL.

La tabla 10 muestra los estereotipos basados en UML a utilizar que permitirán representar las operaciones típicas de un proceso ETL, tal como la integración de fuentes de datos, la transformación, generación de claves entre otros.

Tabla 10. Estereotipos ETL (Luján, 2005).

Estereotipo	Descripción	Icono.
Agregación.	Realizar agregaciones, tal como, sumatorias, promedios, entre otros, con en base algún criterio.	
Conversión.	Realizar un cambio sobre los datos, ya sea, un cambio de tipo de dato, el formato o cálculo de nuevos datos a partir de datos existentes.	
Filtro.	Seleccionar (filtrar) datos por algún tipo de criterio.	
Incorrecto.	Si en el proceso hay datos incorrectos, se deben de re-direccionados a otro destino para su posterior verificación.	
Unión (Join).	Permite unir fuentes de datos relacionadas por medio de uno o varios atributos.	
Carga (Load).	Indica una carga de datos en las estructuras de datos destino.	

Log.	Bitácora de la actividad del proceso ETL, audita las transformaciones que se realizan.	
Combinar.	Diferentes conjuntos de datos se integran. Los conjuntos de datos deben de tener atributos en común.	
Delegado (Surrogate)	Permite generar una clave única.	

La figura 83 muestra el diagrama de proceso ETL para la dimensión de Cantón.

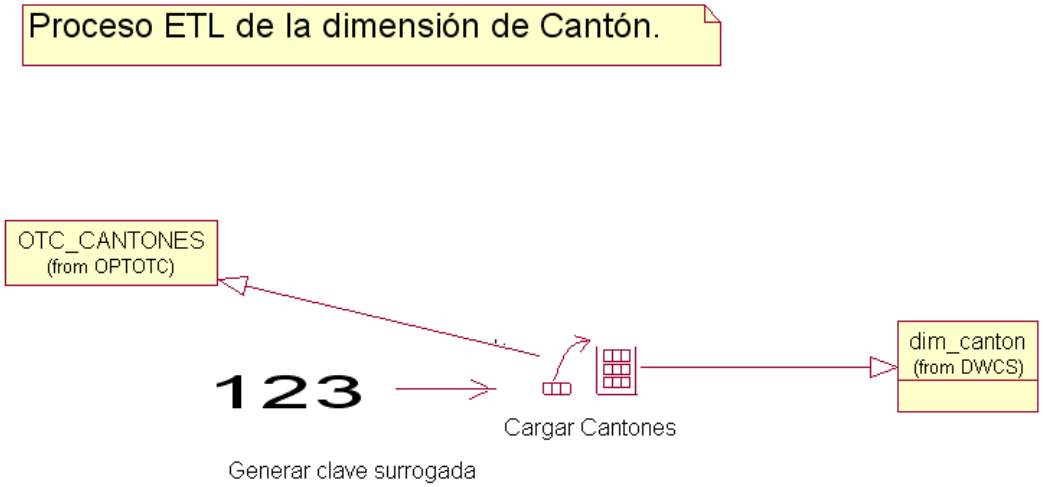


Figura 83. Diagrama de proceso ETL. Cantones (Elaboración Propia).

La figura 84 muestra el diagrama de procesos ETL para la dimensión de Estados.

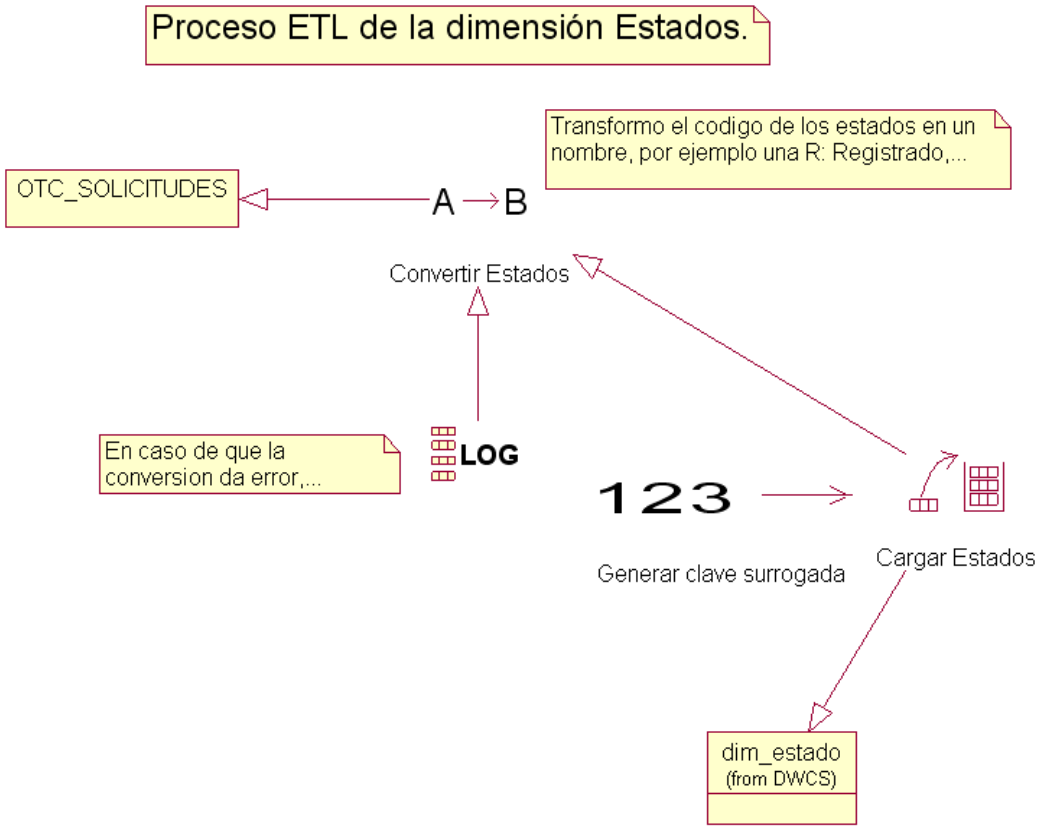


Figura 84. Diagrama de proceso ETL. Estados (Elaboración Propia).

La figura 85 muestra el diagrama de procesos ETL para la dimensión de Tipo de Cliente.

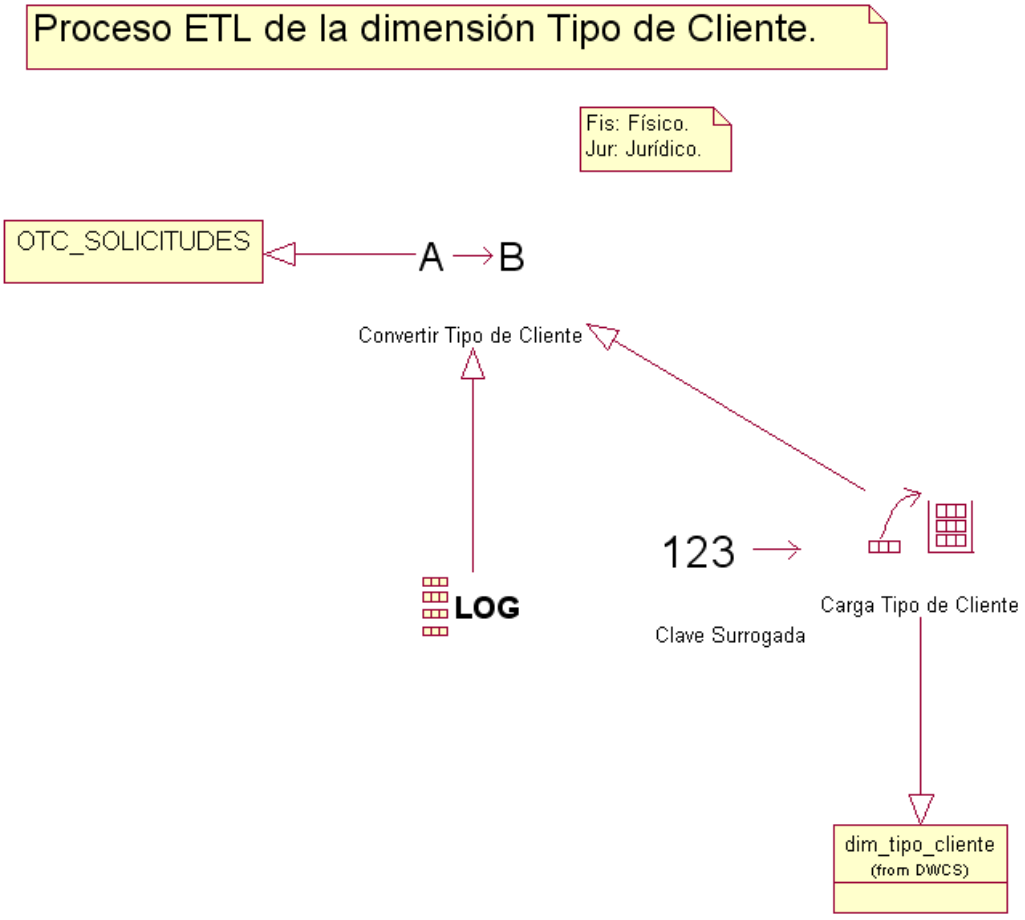


Figura 85. Diagrama de proceso ETL. Tipo de Cliente (Elaboración Propia).

La figura 86 muestra el diagrama de procesos ETL para la dimensión Personal.

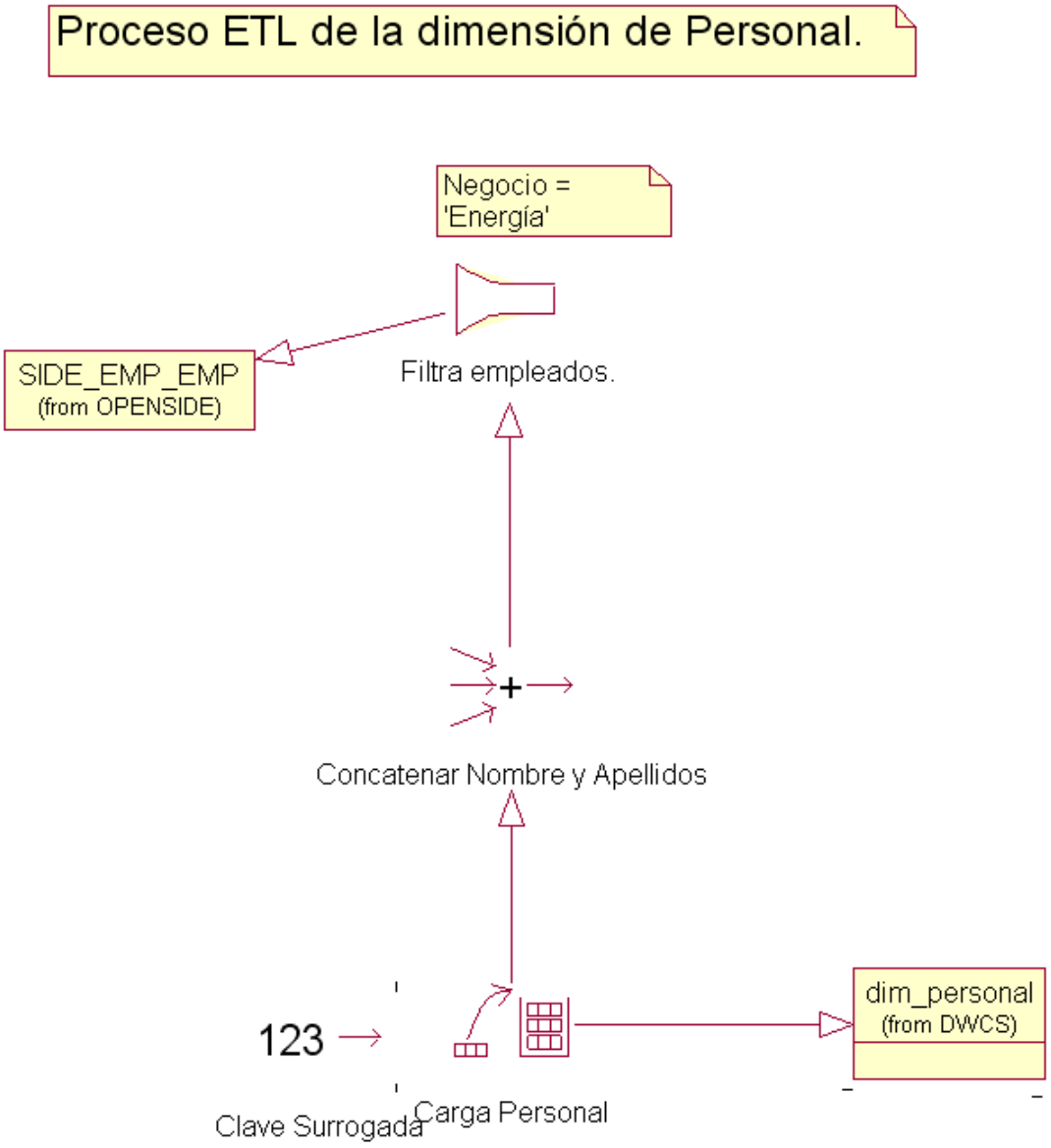


Figura 86. Diagrama de proceso ETL. Personal (Elaboración Propia).

La figura 87 muestra el diagrama de procesos ETL para la dimensión Almacén.

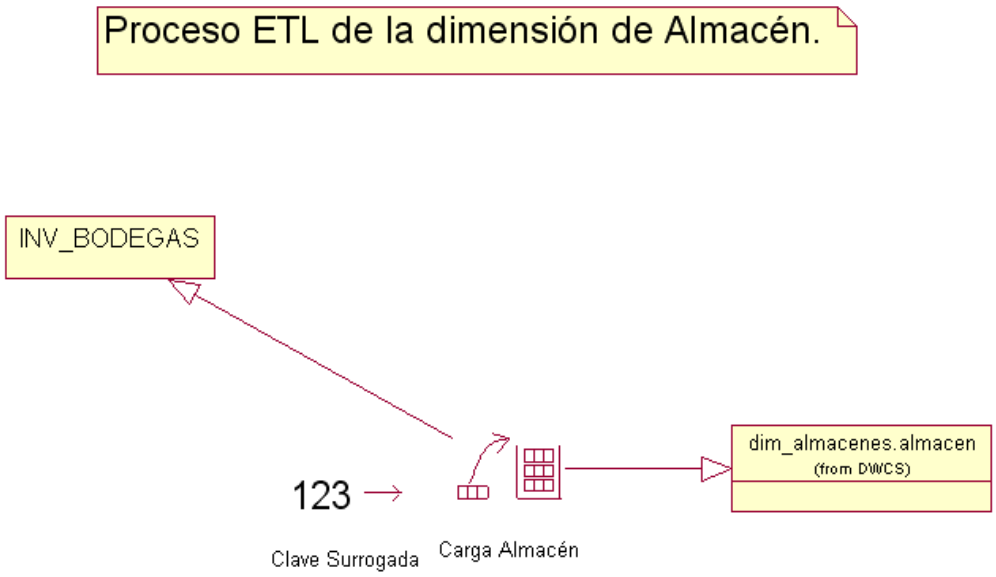


Figura 87. Diagrama de proceso ETL. Almacén (Elaboración Propia).

La figura 88 muestra el diagrama de procesos ETL para la dimensión Solicitud Estudio.

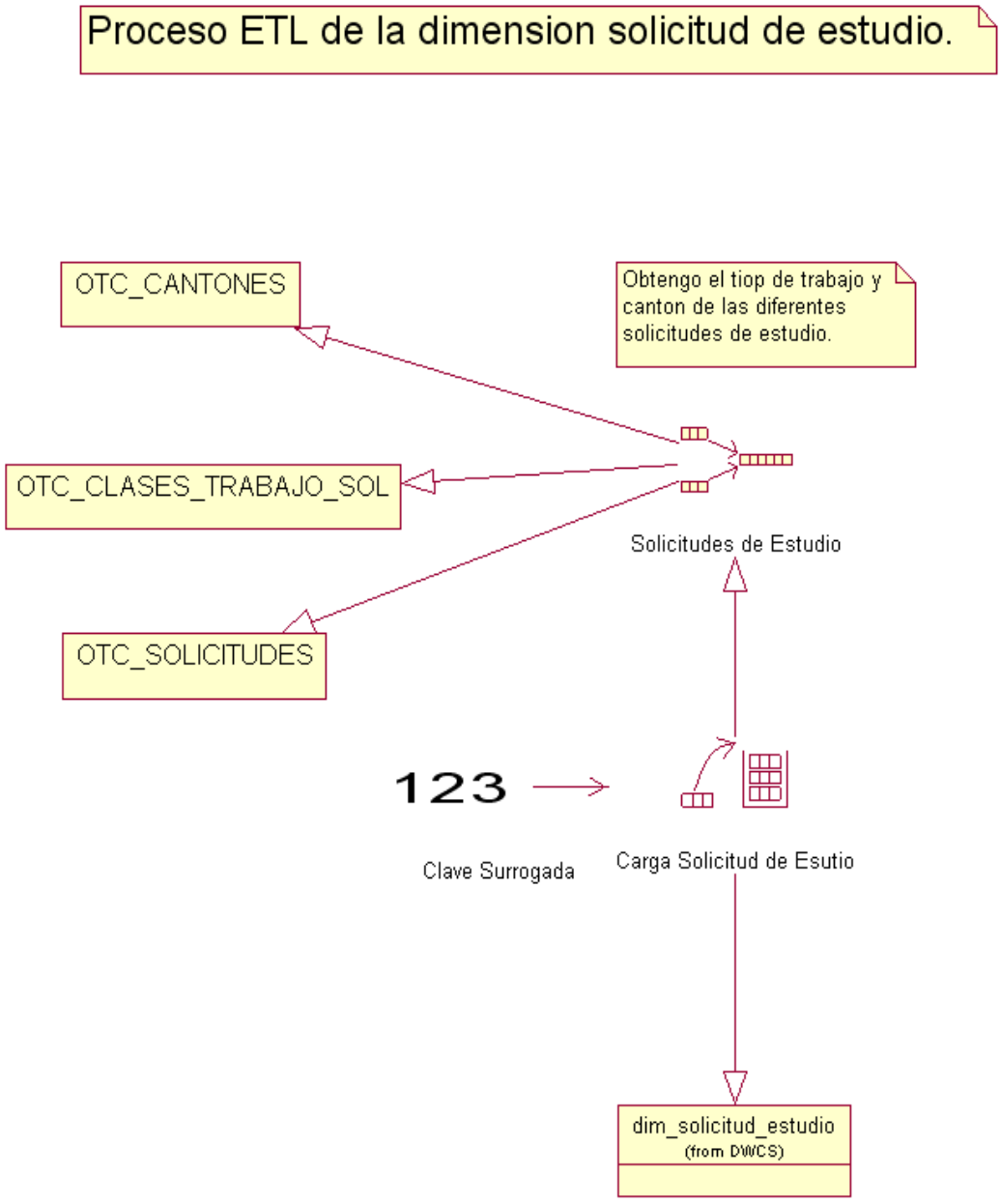


Figura 88. Diagrama de proceso ETL. Solicitud Estudio (Elaboración Propia).

La figura 89 muestra el diagrama de procesos ETL para la dimensión Almacén.

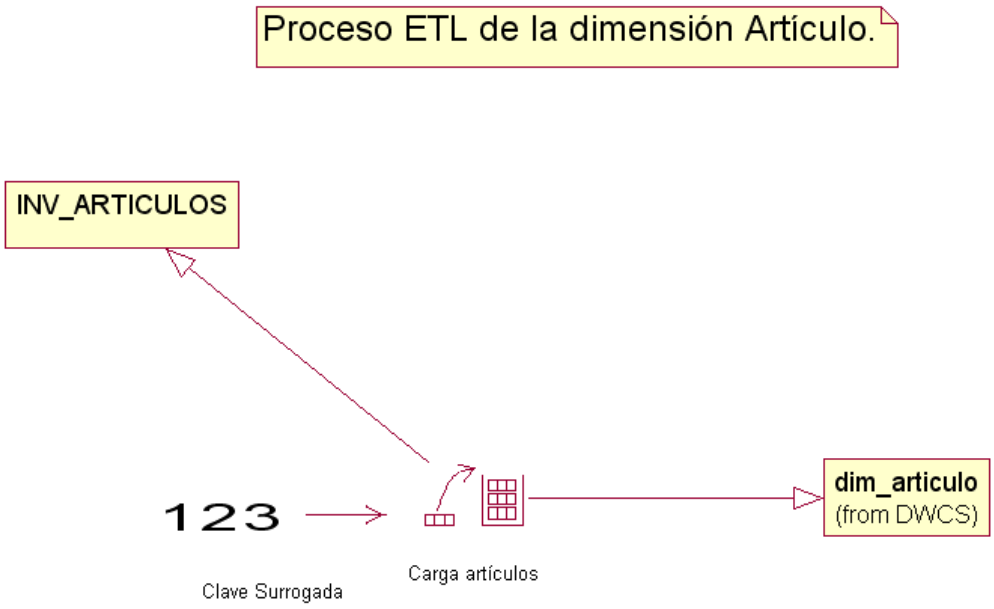


Figura 89. Diagrama de proceso ETL. Artículo (Elaboración Propia).

La figura 90 muestra el diagrama de procesos ETL para la dimensión Tipo de Servicio.

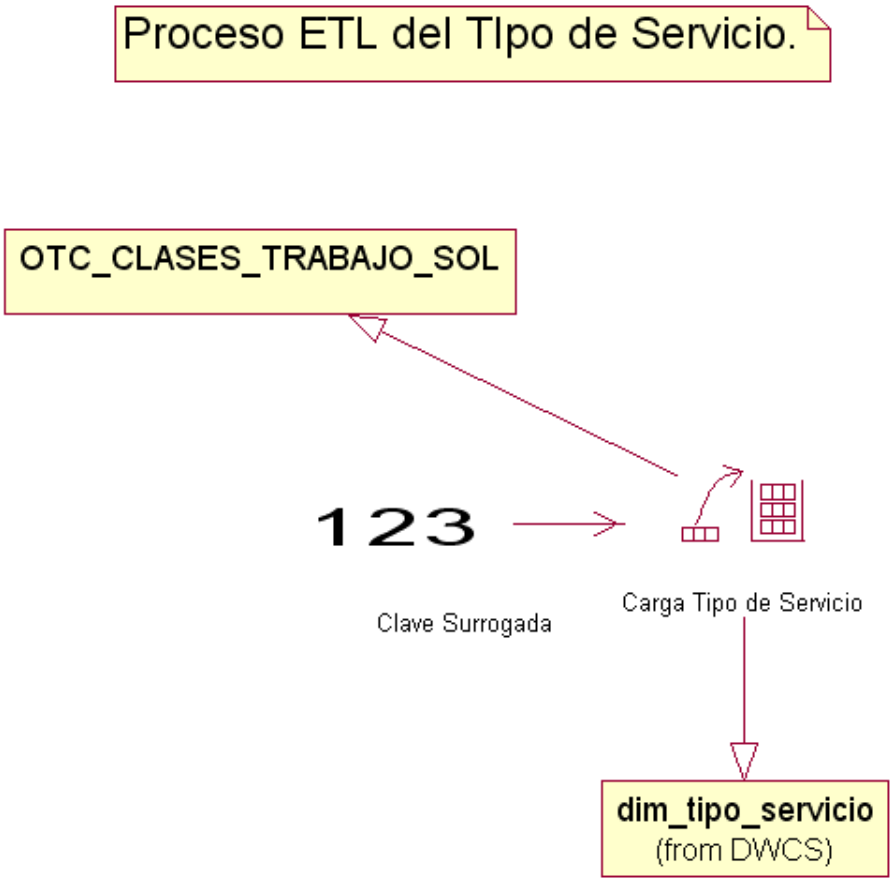


Figura 90. Diagrama de proceso ETL. Tipo de Servicio (Elaboración Propia).

La figura 91 muestra el diagrama de procesos ETL para la dimensión Categoría de Obra.

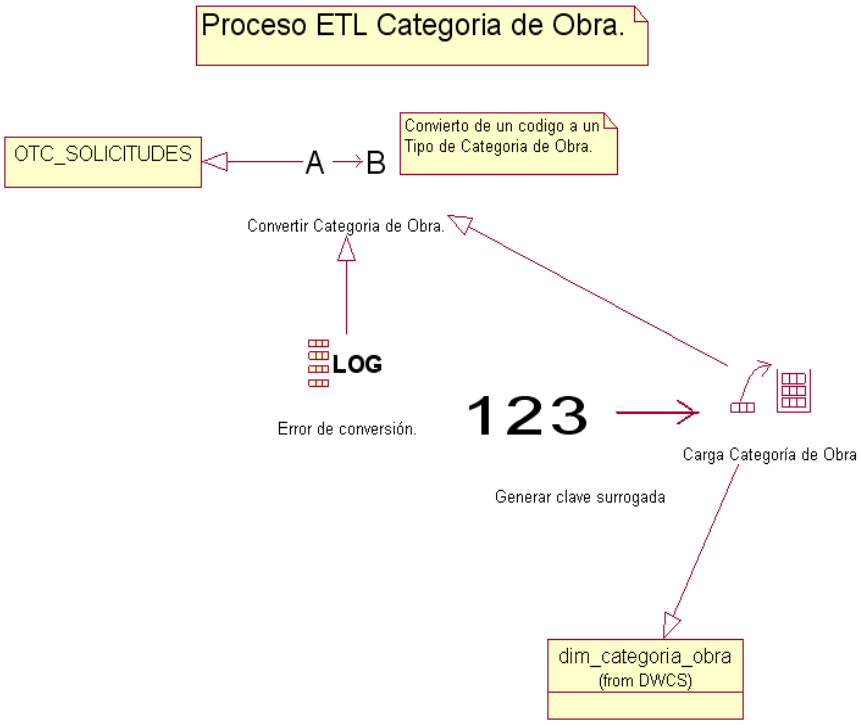


Figura 91. Diagrama de proceso ETL. Categoría de Obra (Elaboración Propia).

La figura 92 muestra el diagrama de procesos ETL para la dimensión Fechas.

Figura 92. Diagrama de proceso ETL. Fechas (Elaboración Propia).