**2016**

**Victor Carreño**

**Magister en Tecnologías de Informacion**

29-8-2016

**Huertos Acmes Chile S.A**

Creación de un Datawarehouse.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| UNIVERSIDAD TECNOLOGICA DE CHILE | | | | |
| Programa | **INGENIERIA EN INFORMATICA** | **Número Documento** | 01 / 02 |
| Materia | TI1232 -Sección XXX – >Tecnologia de Informacion y Comunicación II | **Versión 0.1** | Final septiembre 2018 |
| Profesor | Manuel Rojas – manuel.rojas28@inacapmail.cl | Autores  Sebastián Cáceres  Victor Carreño  David Navarro | - |

Contenido

[**1.**](#_gjdgxs) **INTRODUCCIÓN** 3

[**2.**](#_30j0zll) **ANÁLISIS DE REQUERIMIENTOS** 4

[**2.1.**](#_1fob9te) **Captar los requisitos de información de los distintos grupos de usuarios.** 4

[**2.2.**](#_3znysh7) **Información sobre el uso que se piensa dar a la BD.** 4

[**3.**](#_3dy6vkm) **DISEÑO** 5

[**3.1.**](#_1t3h5sf) **Conceptual** 5

[**3.1.1.**](#_4d34og8) **Enfoque centralizado (Top-Down).** 5

[**3.1.2.**](#_17dp8vu) **Enfoque de integración de vistas (Bottom-Up)** 5

[**3.1.3.**](#_3rdcrjn) **Integración De Vistas (Esquemas)** 5

[**3.2.**](#_26in1rg) **Lógico** 6

[**3.3.**](#_lnxbz9) **Físico** 6

[**BIBLIOGRAFIA** 7](#_35nkun2)

[**Indicaciones** 8](#_1ksv4uv)

1. **INTRODUCCIÓN**

Durante el transcurso del tiempo la informática ha incrementado el poder con el que se desenvuelve dentro de las empresas actuales bien consolidadas, por lo que la incorporación de nuevas tecnologías de información TIC’s es cada vez más apreciada por las mismas.

Así mismo, las empresas a través de la participación en el mundo de las TIC’s han ido evolucionando su productividad con nuevas tecnologías de sistemas operacionales, permitiendo a los usuarios realizar labores en periodos más corto de tiempo e incluso en forma paralela, obteniendo así un mejor desempeño en los procesos operacionales de una organización.

Debido a esto las grandes empresas se han visto en la obligación de mantener actualizada la información sobre sus operaciones, la cual ha sido respaldada en bases de datos u otras fuentes poco funcionales para los requerimientos actuales de una empresa.

Con el objetivo de ayudar en la toma de decisiones, se implementarán bases de datos más complejas, que a su vez entreguen, mediante un análisis de datos, información de mayor valor para así evaluar decisiones estratégicas y proyecciones sobre comportamientos futuros del negocio. En atención a esta necesidad surge la funcionalidad ofrecida por el DATAWAREHOUSE ya que, reuniendo la información de distintas fuentes, puede dar solución a interrogantes creadas con mayor complejidad y así satisfacer los requerimientos más específicos de los usuarios de la empresa.

A continuación, se abordará el tema de cómo implementar este sistema para nuestra empresa, señalando a su vez las variables que se encuentran involucradas.

1. **ANÁLISIS DE REQUERIMIENTOS**
   1. **Captar los requisitos de información de los distintos grupos de usuarios.**

Mediante entrevistas se identificarán las necesidades actuales de la empresa que se pretende dar solución con la implementación de un Data Warehouse en cuanto a resultados consistentes y utilizables. por lo que se realizará un análisis detallado de requisitos a través de los datos recopilados con las preguntas de la entrevista, identificando cuáles serán los indicadores y perspectivas que se utilizarán en la construcción del Data Warehouse. Para finalizar con la confección de un diagrama conceptual que permite visualizar el resultado obtenido.

* 1. **Información sobre el uso que se piensa dar a la BD.**

Para entregar un servicio óptimo se creará un Data Warehouse mediante una base de datos relacional donde se definirán las dimensiones de tiempo y hechos con el fin de capturar los datos requeridos por el usuario con el fin de tomar una decisión.

Un usuario que accederá al Data Warehouse corresponde al Servicio Agrícola Ganadero (SAG), ya que es el órgano oficial del Estado de Chile que tiene a su cargo el comercio de la ganadería y la silvicultura de ese país.

El uso que pretende es mejorar y optimizar la información recopilada desde las distintas fuentes para una mayor eficacia en su toma de decisiones a efectuar y realizarlas con mayor responsabilidad.

1. **DISEÑO** 
   1. **Conceptual**

.

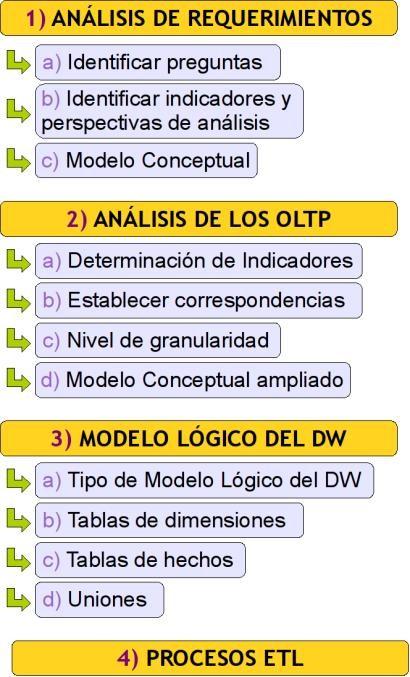
Para llevar a cabo este proyecto se llevarán a cabo las siguientes actividades:

* Se revisará bibliografía referente a los sistemas de datawarehouse
* Se realizará reunión con usuarios claves de la compañía para la captación de requerimientos.
* Se estructurarán los requerimientos captados a los usuarios de la empresa.
* Se realizará una revisión de la data contenida en los archivos Excel entregada por la empresa.
* Se realizará la consolidación de los archivos Excel y se normalizará la data de los mismos.
* Se modelará la Base de datos relacional.
* Se realizarán los testeos para la población de la base de datos.
  + 1. **Enfoque centralizado (Top-Down).**

Este modelo fue implementado por Bill Innon, el enfoque Top-down se utiliza cuando la tecnología y los problemas del negocio alcanzando los objetivos buscados. Se trata de un método sistémico que minimiza los problemas de integración, pero es costoso, debido a la gran cantidad de datos y su poca flexibilidad.

Además, en este método se formula un resumen del sistema sin especificar detalles. Cada parte del sistema se refina diseñándola con mayor detalle. Después cada parte nueva se vuelve a refinar cada vez con mayor detalle, hasta que la especificación completa es lo suficientemente detallada para validar el modelo. Este modelo se diseña con frecuencia con ayuda de la “caja negra” que hace más fácil cumplir requerimientos, aunque estas cajas negras no expliquen en detalle los componentes instalados.

El enfoque Top Down considera que el almacén de datos debe responder a las necesidades de todos los usuarios en la organización, y no solo de un determinado grupo.

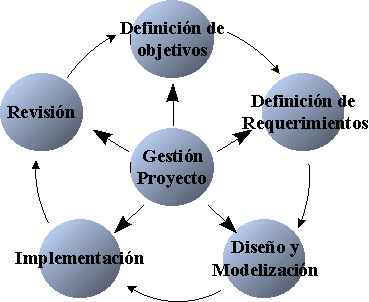


* + 1. **Enfoque de integración de vistas (Bottom-Up)**

Es un modelo que se caracteriza por ser rápido, ya que se basa en experimentos y prototipos propuestas por Ralph Kimball, siendo este un modelo flexible que permite a la empresa ir más lejos con los menores costos. La idea sería construir un DataMart independiente para evaluar las ventajas del nuevo sistema que avanzamos.

Las estrategias basadas en los flujos Bottom-UP pueden ser potencialmente necesarias y suficientes, porque se basan en el conocimiento de todas las variables que pueden afectar a los elementos del sistema.

Las metodologías más utilizadas dentro del modelo Bottom-up son Rapid Warehousing Methodology (RWN) y Kimball Ciclo de vida.



Definición de los objetivos:

En esta fase se define el grupo de proyecto que debe estar compuesto por representantes del departamento informático y de los departamentos usuarios del Data Warehouse, además de la figura de jefe de proyecto. se definirá el alcance del sistema y cuáles son las funciones que el Data Warehouse realizará como suministrador de información de negocio estratégica para la empresa. Se definirá, asimismo, los parámetros que permitan evaluar el éxito del proyecto.

Definición de los requerimientos de información:

Durante esta fase se mantendrán sucesivas entrevistas entre los representantes del departamento usuario final y los representantes del departamento de informática. Se realizará el estudio de los sistemas de información existentes, que ayudarán a comprender las carencias actuales y futuras que deben ser resueltas en el diseño del Data Warehouse. Asimismo, en esta fase el equipo de proyecto debe ser capaz de validar el proceso de entrevistas y reforzar la orientación de negocio del proyecto. Al finalizar esta fase se obtendrá el documento de definición de requerimientos en el que se reflejarán no sólo las necesidades de información de los usuarios, sino cuál será la estrategia y arquitectura de implantación del Data Warehouse.

Diseño y modelización:

Los requerimientos de información identificados durante la anterior fase proporcionarán las bases para realizar el diseño y modelización del Data Warehouse. En esta fase se identificarán las fuentes de los datos (sistema operacional, fuentes externas, etc.) y las transformaciones necesarias para, a partir de dichas fuentes, obtener el modelo lógico de datos del Data Warehouse. Este modelo estará formado por entidades y relaciones que permitirán resolver las necesidades de negocio de la organización.

Implementación:

La implantación de un Data Warehouse lleva implícitos los siguientes pasos:

* Extracción de los datos del sistema operacional y transformación de estos.
* Carga de los datos validados en el Data Warehouse. Esta carga deberá ser planificada con una periocidad que se adaptará a las necesidades de refresco detectadas durante la fase de diseño del nuevo sistema.
* Explotación del Data Warehouse mediante diversas técnicas dependiendo del tipo de aplicación que se dé a los datos.

Revisión:

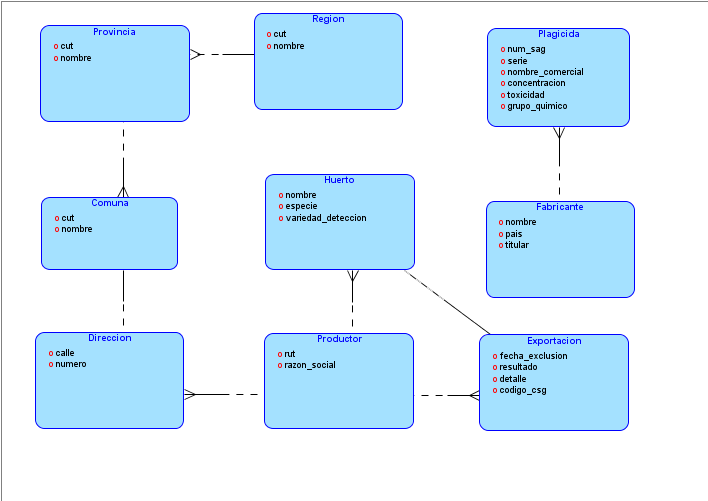
La construcción del Data Warehouse no finaliza con la implantación de este, sino que es una tarea iterativa en la que se trata de incrementar su alcance aprendiendo de las experiencias anteriores. Después de implantarse, debería realizarse una revisión del Data Warehouse planteando preguntas que permitan, después de los seis o nueve meses posteriores a su puesta en marcha, definir cuáles serían los aspectos para mejorar o potenciar en función de la utilización que se haga del nuevo sistema.

Gestión del proyecto:

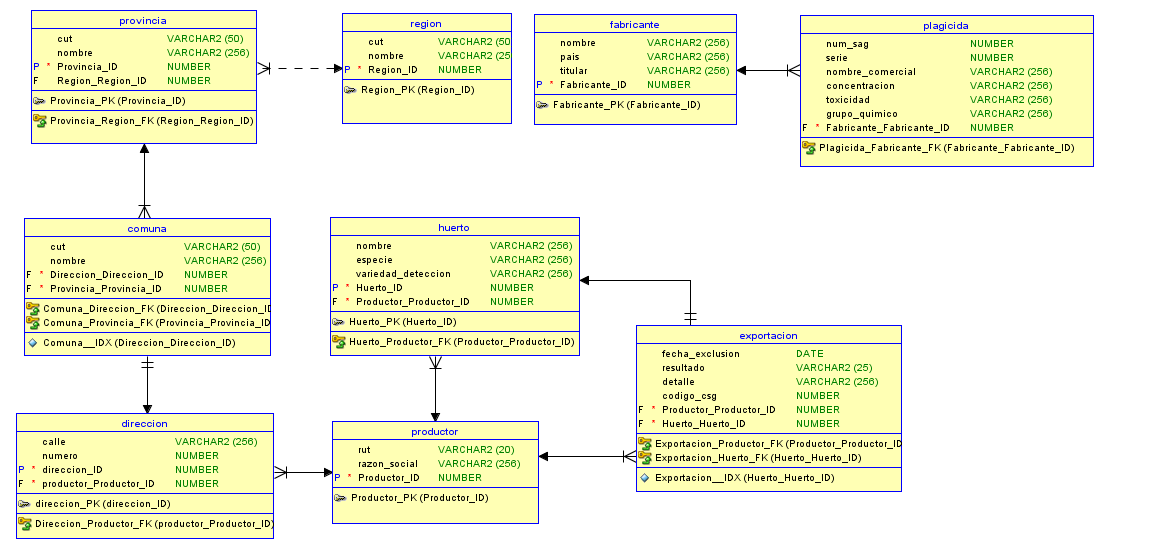
La gestión del proyecto debe encargarse de la coordinación y ejecución de las distintas fases que conforman la construcción e implantación de un Data Warehouse. Este proceso se tiene que apoyar en una metodología específica para este tipo de trabajos, si bien es más importante que la elección de la mejor de las metodologías, el realizar un control para asegurar el seguimiento de esta. En las fases que se establezcan es fundamental incluir una fase de formación en la herramienta utilizada, para un máximo aprovechamiento de la aplicación. Seguir los pasos de la metodología y comenzar el Data Warehouse por un área específica de la empresa permitirá obtener resultados tangibles en un corto espacio de tiempo.

* + 1. **Integración De Vistas (Esquemas)**
* Resolución de conflictos.
* Análisis de redundancias.
* Conflictos de nombres
  + Homónimos
  + Sinónimos
    - Diccionario de datos.
    - Cambiar el nombre a la entidad ó interrelación.
* Conflictos entre tipos de objetos
  + Se transformará el que más convenga.
* Conflictos de dominios
* Conflicto entre entidades.
  + Una entidad es un subconjunto de otra. La solución es introducir un subtipo.
  + Una entidad disjunta con otra y atributos comunes. La solución es crear un super tipo.
* Conflicto de cardinalidades en interrelaciones.
  + ¿Las dos interrelaciones son la misma?
  1. **Lógico**

Diagrama Lógico de BD



* 1. **Físico**

Diagrama Físico de BD

# **BIBLIOGRAFIA**

# 2005 Building The Data Warehouse ( 4th Edition) ( William H. Inmon)

1996 The Data Warehouse Toolkit Ralph Kimball

# 

A continuación, se detallan aspectos de formato y redacción para la elaboración del informe.

**Aspectos de formato y redacción**

1. Considere las siguientes instrucciones de redacción y formato:

* Redactar en tercera persona singular tiempo presente. En lo posible, se debe cuidar que los párrafos completos estén en una misma hoja, evitando las líneas viudas o huérfanas. Esto permite captar mejor la idea que encierra el párrafo completo.
* Formato de entrega: documento Word y PDF.
* Papel tamaño carta. Márgenes estándar (superior e inferior de 2,5 cm. izquierdo y derecho de 2,5 cm).
* Párrafos: alineación justificada, interlineado sencillo.
* Tipo de letra o fuente: Arial o Calibri (Cuerpo).
* Tamaño títulos: 14 y en negrita
* Tamaño subtítulos: 12 y en negrita.
* Tamaño textos: 11 normal.