



Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería en Ciencias y Sistemas

**DATAWAREHOUSE ELECTORAL COMO HERRAMIENTA DE APOYO A LA
INVESTIGACIÓN SOCIOPOLÍTICA**

Aixa Odilia Rodríguez Salán

Claudia Patricia Mejía Méndez

Asesorado por el Ing. Gabriel Fernando Castillo Contreras

Guatemala, febrero de 2009

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**DATAWAREHOUSE ELECTORAL COMO HERRAMIENTA DE APOYO A LA
INVESTIGACIÓN SOCIOPOLÍTICA**

TRABAJO DE GRADUACIÓN

**PRESENTADO A JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA**

POR

AIXA ODILIA RODRÍGUEZ SALÁN

CLAUDIA PATRICIA MEJÍA MÉNDEZ

ASESORADO POR EL ING. GABRIEL FERNANDO CASTILLO CONTRERAS

AL CONFERÍRSELE EL TÍTULO DE

INGENIERA EN CIENCIAS Y SISTEMAS

GUATEMALA, FEBRERO DE 2009

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. Murphy Olympo Paiz Recinos
VOCAL I	Inga. Glenda Patricia García Soria
VOCAL II	Inga. Alba Maritza Guerrero de López
VOCAL III	Ing. Miguel Ángel Dávila Calderón
VOCAL IV	Br. José Milton De León Bran
VOCAL V	Br. Isaac Sultán Mejía
SECRETARIA	Inga. Marcia Ivónne Véliz Varga

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PÚBLICO

DECANO	Ing. Murphy Olympo Paiz Recinos
EXAMINADOR	Ing. Marlon Antonio Pérez Turck
EXAMINADOR	Inga. Floriza Felipa Ávila Pesquera de Medinilla
EXAMINADOR	Inga. Sonia Yolanda Castañeda
SECRETARIA	Inga. Marcia Ivónne Véliz Varga

HONORABLE TRIBUNAL EXAMINADOR

Cumpliendo con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presentamos a su consideración nuestro trabajo de graduación titulado:

**DATAWAREHOUSE ELECTORAL COMO HERRAMIENTA DE APOYO A LA
INVESTIGACIÓN SOCIOPOLÍTICA,**

tema que fuera asignado por la Dirección de la Escuela de Ingeniería en Ciencias y Sistemas, en febrero de 2008.

AIXA ODILIA RODRÍGUEZ SALÁN

CLAUDIA PATRICIA MEJÍA MÉNDEZ

ACTO QUE DEDICO A

- Mi madre** Por la enseñanza que el sacrificio y la perseverancia son las únicas armas que te llevan al triunfo, gracias por tu ejemplo de lucha.
- Mi padre** Porque la distancia ha sido dura, pero tu coraje ha sido fuerte, gracias por los años de paciencia y confianza.
- Mi hermana** Porque siempre has estado allí para ayudarme, darme la mano y levantarme, porque siempre has sido mi soporte, gracias por las noches de desvelo en conjunto.
- Mi sobrina** Gracias por tu alegría y tu sonrisa con la que iluminas nuestras vidas.
- Rolando** Gracias por la sabiduría de tus consejos, por la paciencia y las largas conversaciones que dejan su fruto.
- A mi familia** Por convertir esta familia en un oasis de descanso donde puedo ser yo misma y ser esta loca que me fascina ser.

Joel y Claudia

Por los años de aguante, de apoyo, de desvelo, de cariño, por acompañarme en las buenas y en las malas, por ser los hermanos de vida y de siempre.

Luis Fernando

Por el apoyo, por los momentos compartidos y por los años de cariño desinteresado.

A Ti

Por enseñarme a amar, por compartir tu vida y permitirme compartir la mía, por los años venideros y los futuros triunfos juntos.

Mis amigos

A ustedes que hoy comparten esta alegría conmigo y aquellos que no están a mi lado, por los años de soporte, risas, enojos, pero sobre todo, por quererme tal y como soy y aceptarme sin complejos ni limitaciones.

AIXA RODRÍGUEZ

ACTO QUE DEDICO A

- Mi madre** Por ser la razón de mi vida, por el cariño incondicional, por apoyarme siempre y estar en todo lo que constituye mi mundo.
- Mi padre** Por ser el maestro más grande que he tenido, intelectual y prácticamente, en los estudios y en el camino de la vida, porque siempre has sido la mente que más he admirado.
- Mis hermanos** Gustavo, Vero, Sheny, Lucky, Vicky, Gaby. Por apoyarme y aguantarme incondicionalmente, por ser mis amigas y confidentes de toda la vida, porque siempre han estado allí, esto es principal y fundamentalmente por ustedes.
- Mis sobrinos** Por alegrar con juventud e inocencia nuestras vidas.
- A mi tía, prima y sobrinos** Por confiar en que algún día este logro sería posible.
- Joely Aixa** Por ser mi reflejo, complemento y apoyo durante todos estos años. Por la experiencia vivida y no vivida, porque con ustedes tuve consciencia de muchos aspectos por primera vez.

**Mis amigos de
la Universidad**

Luis Alberto, Gianni, Ludwing, Fernando, Nery y Eddy. Por ser mis acompañantes en esos años de estudio y enseñarme a ser una mejor persona.

**Mis amigos de
CCIE**

Por aceptarme tal y como soy, por enseñarme que las diferencias ideológicas más que barreras son aprendizajes, gracias por el apoyo y cariño que me han brindado.

CLAUDIA MEÍA

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES	V
GLOSARIO	XIII
RESUMEN	XVII
OBJETIVOS.....	XIX°
INTRODUCCIÓN.....	XXI
1. DEMOCRACIA Y TICS	1
1.1. Introducción.....	1
1.2. Democracia y sociedad	3
1.2.1. Sociedad	3
1.2.2. Democracia.....	4
1.2.3. Sistemas electorales.....	5
1.2.4. Sistema político electoral guatemalteco	7
1.2.4.1. Introducción.....	7
1.2.4.3. Sistema electoral.....	9
1.3. Sociedades de la información.....	10
2. INTELIGENCIA DE NEGOCIOS.....	15
2.1. Sistemas de información.....	15
2.1.1. TPS – Sistema de procesamiento de transacciones.....	21
2.1.2. MIS – Sistema de información gerencial.....	21
2.1.3. DSS – Sistema de soporte de toma de decisiones.....	21
2.1.4. ESS – Sistema de soporte ejecutivo	22
2.2. Sistemas Gestores de Bases de Datos.....	24
2.3. Online Transaction Process y Online Analytical Processing.....	28

2.4. Datawarehousing.....	35
2.5. Datamining	43
3. DATAWAREHOUSE ELECTORAL.....	47
3.1. Introducción.....	47
3.2. Arquitectura.....	47
3.3. Modelo de datos.....	49
3.3.1. Descripción general.....	49
3.3.2. Dimensiones.....	50
3.3.2.1. Voto.....	50
3.3.2.2. Región.....	51
3.3.2.3. Elección.....	53
3.3.2.4. Organización.....	53
3.3.2.5. Tiempo	55
3.1.3. Tabla de hechos.....	55
3.1.3.1. Fact votos.....	56
3.1.4. Mapeo conceptual.....	57
4. HERRAMIENTAS DE BI.....	61
4.1. Introducción.....	61
4.2. Características funcionales.....	62
4.3. Selección de la herramienta.....	65
4.3.1. Cuadrante mágico.....	65
4.3.2. Metodología.....	67
4.3.3. Resultados.....	69
4.4. Entorno	71
5. IMPLEMENTACIÓN.....	73
5.1. Metodología.....	73
5.1.1. Esquema fisico	75
5.1.2. Esquema de datos.....	79

5.1.2.1. ETL de datos	82
5.1.2.1.1. Carga inicial.....	82
5.1.2.1.2. Carga incremental.....	89
5.1.2. Esquema lógico.....	92
5.1.2.1. Crear proyecto.....	93
5.1.2.2. Definir datasource.....	94
5.1.2.3. Establecer datasource view.....	99
5.1.2.4. Generar cubo	103
5.1.2.5. Definir dimensiones y jerarquías.....	111
5.1.2.6. Construir y procesar.....	118
5.1.3. Esquema de análisis.....	125
5.2. Reportes.....	128
5.2.1. Votos por Organización Política en la 1ra Vuelta Presidencial.....	128
5.2.2. Votos emitidos en los distintos procesos electorales en municipios del departamento de Guatemala	131
5.2.3. Votos de elecciones de diputación nacional en el año 1995, por departamento según el tipo de voto.....	133
5.2.4. Votos emitidos por organización política en 2da vuelta presidencial del año 2003.....	135
CONCLUSIONES	137
RECOMENDACIONES	139
REFERENCIAS BIBLIOGRÁFICAS	141
BIBLIOGRAFÍA.....	147
APÉNDICE.....	151

ÍNDICE DE ILUSTRACIONES

FIGURAS

1.	Elementos de un sistema de información.....	16
2.	Actividades de los sistemas de información.....	17
3.	Niveles de gestión organizacional y su relación con los sistemas de información.....	20
4.	Relación e intercambio de información entre sistemas de IT	23
5.	Arquitectura del DBMS.....	24
6.	OLAP - OLTP según objetivos de construcción.....	29
7.	OLAP - OLTP según perfil de usuarios.....	29
8.	OLAP - OLTP según integración de los datos.....	30
9.	OLAP - OLTP según frecuencia de acceso.....	30
10.	OLAP - OLTP según orientación de los datos.....	31
11.	OLAP - OLTP según manipulación de datos.....	31
12.	OLAP - OLTP según historia.....	32
13.	Esquema del Datawarehouse.....	35
14.	Esquema estructural del Datawarehouse	36
15.	Ejemplo de niveles de esquematización.....	37
16.	Integración Datamarts-Datawarehouse.....	39
17.	Cubos de información OLAP.....	40
18.	Esquemas físicos multidimensionales	42
19.	Elementos de BI.....	43
20.	Arquitectura del Datawarehouse Electoral.....	48

21.	Modelo general de Datamart Electoral.....	49
22.	Atributos de la dimensión voto	50
23.	Atributos de la dimensión región.....	51
24.	Jerarquía de la dimensión región.....	52
25.	Atributos de la dimensión elección	53
26.	Atributos de la dimensión organización.....	54
27.	Jerarquía de la dimensión organización.....	54
28.	Atributos de la dimensión tiempo.....	55
29.	Atributos de la tabla de hechos detalle	56
30.	Modelo estrella.....	57
31.	Cuadrante mágico de plataformas de BI.....	66
32.	Proceso de implementación de Datawarehouse Electoral.....	74
33.	Crear nueva base de datos.....	75
34.	Definición de estructura de la base de datos.....	76
35.	Esquema de tablas de base de datos dbElectoral.....	79
36.	Archivos de datos electorales (1985-2003).....	80
37.	Esquema de archivo excel de datos electorales.....	80
38.	Datos de dimensión elección.....	81
39.	Valores de la dimensión organización.....	81
40.	Parámetros de la dimensión región	81
41.	Datos de la dimensión votos.....	82
42.	Datos cargados en la dimensión voto.....	83
43.	Datos cargados en la dimensión región.....	84
44.	Valores cargados a la dimensión elección.....	84
45.	Valores cargados a la dimensión organización.....	85
46.	Datos establecidos en la dimensión tiempo.....	86
47.	Carga de datos a la tabla Stg_FactVotos.....	87
48.	Datos cargados a la tabla de hechos Fact_Votos.....	88

49.	Ejemplo de nuevo valor en la dimensión voto.....	89
50.	Ejemplo de inserción de nuevo dato en dimensión región.....	90
51.	Ejemplo de nuevo dato en dimensión elección	90
52.	Ejemplo de ingreso de nuevo dato en dimensión organización.....	91
53.	Ejemplo de nuevo dato en dimensión tiempo.....	91
54.	Proceso de generación de esquema lógico.....	93
55.	Nuevo proyecto	93
56.	Selección de proyecto de analysis services.....	94
57.	Crear nuevo data source.....	95
58.	Crear conexión a base de datos ROLAP.....	95
59.	Parámetros de conexión a base de datos ROLAP.....	96
60.	Selección de conexión a base de datos.....	97
61.	Parámetro de credenciales.....	98
62.	Finalización de creación de data source.....	98
63.	Creación del data source view	99
64.	Selección del data source.....	100
65.	Mapeo de relaciones.....	100
66.	Tablas del data source que forman el cubo.....	101
67.	Finalización de la creación de data source view.....	102
68.	Diseño del data source view.....	102
69.	Creación de nuevo cubo.....	103
70.	Selección de fuente de datos.....	104
71.	Indicar data source view del cubo.....	104
72.	Definición de dimensiones, tabla de hechos y tiempo	105
73.	Diseño general de modelo OLAP	106
74.	Establecimiento de dimensión tiempo.....	107
75.	Definición de métricas	108
76.	Selección de valores de dimensiones.....	109

77.	Finalización de la creación del cubo	110
78.	Diseño del cubo dwElectoral.....	110
79.	Propiedades de los atributos dimensionales.....	111
80.	Propiedad NameColumn.....	112
81.	Creación de nueva jerarquía.....	112
82.	Mantenimiento de dimensiones.....	113
83.	Configuración de dimensión elección.....	113
84.	Configuración de dimensión tiempo.....	114
85.	Parámetros de la dimensión voto	114
86.	Configuración de dimensión organización (atributo IdTipo).....	115
87.	Configuración de dimensión organización (atributo OrganizacionSk)	115
88.	Jerarquías de la dimensión organización.....	116
89.	Configuración de dimensión región (atributo RegionSk).....	116
90.	Parámetros de dimensión región (atributo IdDepartamento).....	117
91.	Configuración de dimensión región (atributo IdRegion).....	117
92.	Jerarquía de dimensión región.....	118
93.	Creación del esquema físico del modelo OLAP.....	119
94.	Resultados de la construcción del cubo.....	119
95.	Ingreso a SQL Server Analysis Services.....	120
96.	Estructura física del cubo dwElectoral.....	120
97.	Procesamiento del cubo OLAP.....	121
98.	Inicio del procesamiento del cubo OLAP	122
99.	Procesamiento del cubo OLAP.....	123
100.	Herramienta browse de Analysis Services.....	124
101.	Resultados de consulta.....	124
102.	Creación de nueva librería de Proclarity.....	125
103.	Asignación de nombre y descripción a la librería	126
104.	Librería creada en Proclarity.....	126

105. Escoger el servidor OLAP.....	127
106. Selección de cubo dwElectoral.....	127
107. Resultados en Excel de votos por organización en 1ra vuelta presidencial del año 2003.....	128
108. Votos por organización política 1ra vuelta presidencial año 2003 vistos en Proclarity.....	129
109. Gráfica de resultados electorales del año 2003, por organización política en 1ra vuelta presidencial.....	130
110. Emisión de voto en municipios de Guatemala.....	131
111. Análisis por árbol de descomposición	132
112. Total de votos emitidos, según el tipo en elección de 1995.....	133
113. Gráfico de selección por un departamento específico	134
114. Votos por Partido Político en 2da vuelta presidencial 2003 por región y departamento.....	135
115. Mapa de comportamiento por organización política, departamento y región	136
116. Pantalla inicial de instalación.....	152
117. Aceptación de licencia.....	152
118. Instalación de prerrequisitos	153
119. Validación de instalaciones previas.....	154
120. Ingreso de valores de licencia	154
121. Componentes de instalación.....	155
122. Instancia de base de datos.....	156
123. Método de autenticación.....	157
124. Opciones de Report Server	157
125. Parámetros de instalación.....	158
126. Pantalla de progreso.....	158
127. Pantalla inicial de instalador de Proclarity.....	160

128. Acuerdo de licencia.....	160
129. Definición del tipo de instalación	161
130. Definición del directorio virtual de trabajo	162
131. Localización física del servidor	162
132. Indicar servidor de SQL	163
133. Proporcionar usuario y clave del administrador de Proclarity	163
134. Resumen.....	164
135. Finalización de la instalación.....	164
136. Herramienta de administración del servidor.....	165
137. Asignación de permisos de usuario.....	165
138. Búsqueda de actualizaciones.....	166
139. Ventana de actualización de componentes.....	167
140. Ícono para Proclarity Professional	167
141. Entorno de Proclarity Professional.....	168

TABLAS

I.	Cuadro comparativo de tecnologías OLAP y OLTP.....	29
II.	Arquitecturas OLAP.....	33
III.	Técnicas de minería de datos según área de especialización.....	44
IV.	Entidades de la dimensión voto	51
V.	Entidades de la dimensión región.....	52
VI.	Entidades de la dimensión elección.....	53
VII.	Entidades de la dimensión organización.....	54
VIII.	Entidades de la dimensión tiempo.....	55
IX.	Script de creación.....	58
X.	Elementos funcionales de las plataformas de BI.....	63
XI.	Criterios de evaluación de primer nivel.....	68
XII.	Criterios de evaluación de segundo nivel	68
XIII.	Resultados obtenidos por herramienta en criterios de Integración.....	69
XIV.	Resultados de criterios de presentación de la información	69
XV.	Resultados para criterios de análisis	70
XVI.	Resultados para criterios de soporte	70
XVII.	Resultados para criterios de otros factores.....	70
XVIII.	Valores obtenidos para criterios de primer nivel.....	70
XIX.	Resultados de TOC.....	71
XX.	Tabla descriptiva de herramientas BI de la solución.....	72

GLOSARIO

- BI** Acrónimo de *Business Intelligence* (Inteligencia de Negocios). Son el conjunto de herramientas y estrategias que permiten la extracción de conocimiento de los datos históricos de una organización.
- CRM** Siglas del término en inglés *Customer Relationship Management* (Administración de la Relación con los Clientes) que define al modelo estratégico de gestión centrada en el cliente.
- Datamart** Subconjunto de datos del datawarehouse enfocados a un área específica del negocio y creada con el objetivo de dar soporte para la toma de decisiones.
- Datamining** La minería de datos consiste en la extracción de información de valor, a través de la aplicación de técnicas estadísticas e inteligencia artificial, para la identificación de patrones y características ocultas en los datos.

Datawarehouse	Un almacén de datos es el conjunto de datos históricos, integrados y orientados a determinado ámbito que permiten el soporte a la toma de decisiones.
DBMS	Abreviatura del inglés <i>DataBase Management System</i> (Sistema de Gestión de Bases de Datos – SGBD). Elementos de software que proveen la interfaz para la administración interactiva entre usuarios, datos almacenados y aplicaciones.
ERP	<i>Enterprise Resource Planning</i> (Planificación de Recursos Empresariales) es un sistema que permite la integración y manejo de las operaciones de producción, distribución y manufactura de las organizaciones orientadas a la generación de bienes y servicios.
ETL	Siglas en inglés de Extracción, Transformación y Carga (<i>Extract, Transform, Load</i>) y es el proceso que permite a las organizaciones la integración de datos de distintas fuentes y formatos para que sean cargados al datawarehouse.
Firmware	Bloque de instrucciones con objetivos específicos y que establece la estructura lógica de funcionamiento de más bajo nivel que permite el control de los circuitos electrónicos.

Hardware Conjunto de partes físicas y tangibles de una computadora (componentes mecánicos, eléctricos y electromecánicos).

IT *Information Technology* (Tecnología de la Información) es el estudio, diseño, soporte, implementación, mantenimiento y desarrollo del software, firmware y hardware de los sistemas computacionales.

Software Son el conjunto de elementos lógicos y de soporte de una computadora que permiten la realización de una tarea específica.

RESUMEN

Es de vital importancia aceptar que para el fortalecimiento y consolidación de los procesos electorales en Guatemala y la construcción de una sociedad incluyente, es necesaria la investigación de los elementos y factores que amenazan la participación social y que van en detrimento del bienestar colectivo. Debe reconocerse la responsabilidad que como sancarlistas y guatemaltecos tenemos de aportar nuestros conocimientos en el cumplimiento de los preceptos universitarios de compromiso social e investigación.

Con el fin de proveer al sector de investigación sociopolítica de la Universidad de San Carlos de Guatemala una herramienta que permita la identificación de fortalezas, debilidades, patrones y mecanismos que caracterizan los procesos electorales guatemaltecos, se crea el presente sistema de inteligencia de negocios orientado a procesos electorales.

Inicialmente se establecen los parámetros teóricos y conceptuales que definen el ámbito del sistema, instaurando los elementos de interpretación que guiarán la construcción e implantación del mismo. Se elabora el modelo de datos que soporta la lógica funcional del datawarehouse, a través del establecimiento de dimensiones, jerarquías, tablas de hechos y métricas. Previo a la implementación se define la herramienta de reporte a utilizar, a través de la aplicación de una metodología de evaluación de proyectos de IT. Por último, se realiza la construcción del sistema y de reportes personalizados para la interpretación de los datos e identificación de elementos claves de participación ciudadana en procesos electorales.

OBJETIVOS

GENERAL

Utilizar las herramientas de IT para la implementación de un sistema de inteligencia de negocios orientado a procesos electorales, que proporcione un instrumento para el fortalecimiento de la investigación sociopolítica.

ESPECÍFICOS

1. Definir los principales elementos de una democracia y la importancia de la tecnología de la información.
2. Establecer el marco teórico de inteligencia de negocios que proporcione la base metodológica de construcción.
3. Construir el datawarehouse electoral a través del diseño del modelo lógico conceptual que soporte el almacenamiento, manipulación y acceso a los datos electorales.
4. Cargar los datos obtenidos al repositorio de datos luego de haber realizado las transformaciones necesarias para que se adapten a la estructura del modelo de metadatos OLAP.
5. Crear reportes personalizados que permitan la interpretación, análisis, comparación, observación e identificación de elementos representativos de participación electoral.

INTRODUCCIÓN

La Universidad de San Carlos de Guatemala tiene como mandato constitucional y compromiso con el pueblo guatemalteco, la promoción de la investigación en todas las áreas del conocimiento, así como la búsqueda de soluciones a los problemas de interés nacional. Cumpliendo entonces con la responsabilidad inherente que tenemos hacia con nuestra Alma Mater y nuestro país, se busca fusionar dos áreas del conocimiento aparentemente opuestas, como lo son la Ingeniería en Ciencias y Sistemas y la Ciencia Política para fomentar, promover y fortalecer la investigación sobre procesos electorales a través de la aplicación de tecnologías de la información.

La inteligencia de negocio ha demostrado su capacidad para la recopilación, estructuración y comprensión de datos organizacionales transformados en información de utilidad, para el mejoramiento y aseguramiento de los procesos de la misma. Aplicar esta ventaja en la interpretación de la información desde el punto de vista investigativo se convierte en un nuevo paradigma donde la investigación y la tecnología se tienden la mano para avanzar más allá de los límites conocidos.

Es entonces el propósito fundamental del proyecto de graduación elaborar una bodega de datos multidimensional que sirva de herramienta de apoyo a la investigación científica en materia electoral y permita a la Universidad de San Carlos de Guatemala, presentar e identificar deficiencias y fortalezas en los procesos electorales guatemaltecos, en aras de la construcción de una sociedad más participativa y un Estado democrático de derecho.

1. DEMOCRACIA Y TICs

1.1. Introducción

La democracia es uno de los sistemas políticos de gobierno más extendido del planeta desde hace varias décadas, aspecto que un teórico de la política como Samuel Huntington llamo la “*tercera ola*” de la democratización en el mundo. En ese sentido, la preocupación por la calidad y la profundidad de la democracia es un tema central para muchas sociedades, tal es el caso de Guatemala, ya que su antecedente inmediato es una larga historia de guerra civil y violaciones a los derechos humanos, que puso a nuestra sociedad en la mira de las naciones más desarrolladas del planeta.

En ese contexto político y social, las tecnologías de la información y comunicación (TIC's) han representado un cambio significativo en las tradicionales formas de relacionamiento social, político y económico, ya que han expandido las posibilidades de interacción, han favorecido el involucramiento en tiempo real de una gran cantidad de personas situadas en diferentes partes del globo, y han fortalecido la capacidad de la ciudadanía de monitorear el desarrollo de los fenómenos políticos, paso por paso. De hecho, esa superación del tiempo y la distancia mediante las TIC's, que es una de las características principales de nuestra época, han llevado a la conciencia de que el globo terráqueo súbitamente se volvió pequeño.

Resulta interesante cómo los avances tecnológicos han significado cambios también en la forma de concebir la actividad académica y profesional, ya que ahora se cuenta con poderosas herramientas computacionales, paquetes estadísticos y programas especialmente diseñados para la investigación, que han permitido la exploración de los fenómenos políticos y sociales de forma que nunca un investigador hubiera imaginado jamás.

Uno de esos campos ha sido la exploración de las dinámicas espaciales y temporales del voto en los procesos democráticos de elección de gobernantes, hecho político fundamental que sustenta las democracias modernas. Las herramientas estadísticas disponibles permiten generar y analizar bases de datos que busquen correlaciones, fenómenos escondidos en la cantidad de datos disponibles y la capacidad de relacionar múltiples realidades al comparar sucesivamente varios procesos electorales de una sola vez, por ejemplo, o aglutinar diversas realidades municipales con características similares, permitiendo de este modo a los investigadores de los fenómenos políticos tener herramientas para la comprensión de la realidad.

De hecho, las TIC's aplicadas a la investigación, especialmente en el tema de los procesos electorales y de voto, permiten superar el principal problema de la investigación: *los procesos electorales son un fenómeno eminentemente individual, pese a que ocurre en la confluencia masiva de muchos electores, ya que la decisión que lleva posteriormente al triunfo o no de una opción partidaria, pasa indefectiblemente por la opción que cada elector asume en lo individual.*

Por eso, el investigador que intenta comprender las dinámicas de voto, tiene que superar la enorme dificultad de comprender muchos fenómenos de decisión aparentemente similares, pero que al ocurrir en la simultaneidad de muchos individuos dispersos por un gran territorio, se corre el riesgo de homogeneizar tendencias, –se sabe que las estadísticas dicen tanto de los fenómenos, como de las características que esconden-, que invisibilizan procesos y fenómenos atípicos, locales o regionales, que son sumamente importantes para comprender la democracia electoral, especialmente si lo que se busca es comprender la realidad.

1.2. Democracia y sociedad

1.2.1. Sociedad

El ser humano ha tenido siempre la necesidad de establecer relaciones con individuos que compartan sus mismas inquietudes y necesidades. Desde tiempos de la prehistoria, la organización a través de comunidades con objetivos establecidos y patrones diversos de comportamiento, ha sido una práctica común. Entidades donde cada integrante cumple con un rol específico que le permite a la colectividad funcionar de forma integral y que han evolucionado, hasta convertirse en sociedades cuyo precepto es la búsqueda del bien común. Aristóteles afirma que el hombre es un ser social por naturaleza, esto propone que, la necesidad de interacción a través de estructuras sociales proviene de las características inherentes del ser humano.

El término sociedad proviene del latín *socius* que significa compañía y varios autores coinciden en que sociedad es un sistema integrado por un conjunto de personas, que se relacionan entre sí por medio de leyes jurídicas y consuetudinarias, que comparten una misma cultura y que a través de la cooperación logran alcanzar los fines de su existencia. Según Joseph Fichter *“comunidad es un grupo territorial de personas con relaciones recíprocas, que sirven de medios comunes para lograr fines comunes”*.

Estado es la forma de organización política de las sociedades, formada por un conjunto de instituciones que permite regular el entorno social en que se desenvuelven. Las formas de gobierno, sistemas de gobierno o modelos políticos son conceptos que se refieren a la forma organizativa del Estado adoptada por las sociedades según las formas de poder asociadas, para Hegel las formas de gobierno son la estructura política de las sociedades bien determinadas. Por su parte Aristóteles identifica tres formas de gobiernos puras: Monarquía (gobierno de uno solo), Oligarquía (gobierno de la minoría selecta - aristocracia) y Democracia (gobierno de todos).

1.2.2. Democracia

Según publicación de Norberto Bobbio en 1978, *“desde la edad clásica hasta hoy, el término democracia ha sido siempre utilizado para designar una de las formas de gobierno, o sea uno de los diferentes modos en que puede ser ejercido el poder político. En particular designa aquella forma de gobierno en la cual el poder político es ejercido por el pueblo...”*. Por su parte Aristóteles plantea que *“la democracia surgió de este pensar de los hombres: que si son iguales en cualquier aspecto, son iguales en todos”*.

El término democracia proviene etimológicamente del griego “*demos*” cuyo significado es pueblo y “*kratos*” que se traduce en gobierno, pudiendo denotarse de este modo como: “*gobierno del pueblo*”. Puede establecerse entonces que democracia, es la organización del Estado, donde las decisiones colectivas definen la forma de gobierno, mediante mecanismos de participación que permiten la legitimación de los representantes y donde todos los habitantes son iguales. Según Elwyn White “*la democracia se basa en la suposición recurrente de que más de la mitad de las personas tienen la razón en más de la mitad de las ocasiones*”.

La democracia no puede ser definida como un conjunto de reglas, parámetros o instituciones, cada sociedad asume su visión democrática de forma particular y está determinada por factores históricos y culturales, así como circunstancias políticas, económicas y sociales. Se fundamenta en valores, actitudes, principios y prácticas que buscan la convivencia en armonía de los ciudadanos.

Existen tres formas principales de democracia: representativa, directa y participativa. Se entiende por democracia representativa a la delegación de la toma de decisiones en representantes electos, democracia directa aquella en la que las decisiones son adoptadas por los ciudadanos y democracia participativa donde la representación está dada por organizaciones y asociaciones ciudadanas.

1.2.3. Sistemas Electorales

El principio de constitución de una sociedad democrática es la convivencia de sus miembros como iguales y en libertad, con capacidad de toma de decisiones sobre los procesos y organización del Estado, siendo el proceso electoral el pilar fundamental sobre con el cual se asegura la inclusión participativa ciudadana.

Proceso o sistema electoral, no es más que un conjunto de reglas, principios, procedimientos y normas establecidos legalmente, por medio de los cuales los ciudadanos ejercen su voluntad política. Según el Instituto Federal Electoral (IFE) de México sistema electoral *“es el conjunto de medios a través de los cuales la voluntad de los ciudadanos se transforma en órganos de gobierno o de representación política....es una estructura intermedia del proceso a través de la cuál una sociedad democrática elige a sus gobernantes...”*. Por su parte Valles y Bosch establecen que un sistema electoral es *“el conjunto de elementos normativos y sociopolíticos que configura el proceso de designación de titulares de poder, cuando este proceso se basa en preferencias expresadas por los ciudadanos de una determinada comunidad política”*.

Existen tres esquemas principales de sistemas electorales, los cuales son aplicados con variaciones, según las características de cada país:

- **Mayoría relativa y absoluta:** En la elección por medio de mayoría relativa, cada elector tiene un voto y es seleccionado ganador el candidato que mayor número de votos obtenga, sin que esto signifique una mayoría representativa. Debido a que en casos de alta fragmentación del voto, el candidato seleccionado por mayoría relativa representaría a una minoría, se utiliza el sistema de mayoría absoluta, esto significa que el ganador deberá obtener al menos un 50% de los votos emitidos.
- **Representación proporcional:** El sistema de representación proporcional busca eliminar los problemas de sobre y sub representación, asignando a cada partido político el número de representantes proporcional a su fuerza electoral, para lo que se utiliza el sistema de representación D'Hondt.
- **Sistemas mixtos:** Este tipo de sistema electoral implica la combinación de mayorías absolutas con representación proporcional.

1.2.4. Sistema Político Electoral Guatemalteco

1.2.4.1. Introducción

Guatemala es una sociedad que apenas hace 23 años estrenó sistema democrático participativo, mediante la instauración de procesos electorales de elección popular y la asociación del pueblo en partidos políticos, permitiendo así el libre ejercicio de derechos y deberes ciudadanos. Al igual que otros países de América Latina, sufre del deterioro progresivo del sistema político, ocasionado por la discrepancia entre el funcionamiento de las instituciones electorales y las necesidades sociales. A pesar de la firma de los acuerdos de paz, como impulsor de las libertades públicas y derechos humanos, se encuentran grandes deficiencias en la gestión de los procesos electorales, generando esto desconfianza por parte de los electores. Anomalías tales como padrones electorales incompletos, inscripción ilegal de candidatos, duplicación de cédula y voto de fallecidos, irregularidades que generan una disminución significativa en la credibilidad que los ciudadanos tienen de los organismos electorales y su participación durante el proceso electoral.

En la clasificación internacional sobre participación electoral, en el 2001, Guatemala se encontraba en el nivel 154 de 169 países con un promedio de participación del 51.6%, reflejándose de este modo el desinterés generalizado por el ejercicio y atención de las demandas de la ciudadanía. Los esfuerzos en varios sectores poblacionales para la inclusión política, se deben a que hace poco más de 40 años no eran ciudadanos alrededor del 75% de los habitantes y no ejercían el voto la población pobre, analfabeta o mujer.

1.2.4.2. Antecedentes

Durante 36 años Guatemala fue víctima de conflicto armado interno, caracterizado por la profunda violencia, militarización de la sociedad, división de la población y gobiernos militares de facto que dejaron un saldo de 200,000 guatemaltecos muertos, en su mayoría indígenas. En 1985 la Constitución Política emite el Decreto 10-85 Ley Electoral y de Partidos Políticos convirtiéndose en el marco legal en materia electoral y entrando en vigencia el 14 de enero de 1986, fecha que marca una nueva etapa para Guatemala y señala el fin de una larga lista de gobiernos militares y el restablecimiento del derecho constitucional. No es sino hasta noviembre y diciembre de 1985 que se logra la transición hacia un gobierno civil, dando origen a procesos y luchas políticas que conllevaran a la finalización del enfrentamiento armado. Durante más de una década se llevaron a cabo las negociaciones entre los elementos en conflicto hasta que el 29 diciembre de 1996 se firman los Acuerdos de Paz.

Durante la historia reciente de Guatemala, de 1985 a 1998 se efectuaron nueve procesos electorales, de los cuales tres fueron elecciones generales para elegir presidente y diputados, tres comicios para elegir corporaciones municipales, uno para la elección de diputados y dos consultas populares. Estos procesos electorales se amplían con tres adicionales, a partir de 1999 a 2007 con una única elección general y de manera simultánea para elegir el liderazgo del Ejecutivo, Legislativo y Municipal.

La Ley Electoral y de Partidos Políticos promulgada en 1985 ha sufrido dos bloques importantes de reformas en los años 2004 y 2006, teniendo como finalidad el fortalecimiento y superación de las deficiencias de los procesos electorales, modernizar el régimen electoral guatemalteco, aumentar la participación ciudadana, la descentralización organizativa, así como regular el funcionamiento, fiscalización y transparencia de los partidos políticos.

1.2.4.3. Sistema Electoral

Según la Constitución Política de 1985, Guatemala se constituye con una República presidencialista, democrática y representativa. El sistema político nacional está dividido en tres poderes: ejecutivo, legislativo y judicial. El Poder Ejecutivo está representado por el presidente y vicepresidente de la República, electos sobre una única fórmula partidista llamada binomio, a través del voto universal y directo por un período improrrogable de 4 años utilizando el sistema de mayoría absoluta con segunda vuelta. El Poder Legislativo está constituido por el Congreso de la República cuya asamblea unicameral cuenta con 158 representantes, electos por un período de 4 años a través del voto directo y utilizando el sistema de representación proporcional con fórmula de D'Hondt con al menos un representante por cada distrito. El Poder Judicial está integrado por la Corte Suprema de Justicia y la Corte de Constitucionalidad cuyos 13 representantes son elegidos por los integrantes del Congreso por un período de 5 años.

Los 22 departamentos están agrupados en 8 regiones y son presididos por un gobernador, nombrado por el presidente de la República y por un consejo departamental integrado por alcaldes municipales y representantes de los sectores de la sociedad civil, además se establecen gobiernos locales en los 331 municipios a través de la elección de alcaldes y concejales municipales.

El sistema político guatemalteco está caracterizado por la representación partidaria regulada a través de la Ley Electoral y de Partidos Políticos, donde se establece todo lo relativo al sufragio, derechos políticos, organizaciones políticas, órganos electorales, autoridades y proceso electoral.

El Tribunal Supremo Electoral es la institución encargada de la organización y realización de los procesos electorales, integrado por 5 magistrados titulares y 5 magistrados suplentes, electos por el Congreso de la República, de una nómina de 40 candidatos, quienes se mantienen en funciones durante 6 años. Cuenta según decreto ley con una asignación anual de no menos del 0.5% del Presupuesto General de Ingresos Ordinarios del Estado para cubrir gastos de funcionamiento, mientras que durante el año electoral esta cantidad es aumentada en base al presupuesto establecido por dicha institución.

1.3. Sociedades de la información

“La tecnología de la comunicación transforma todas las relaciones sociales y convierte al mundo en una aldea global, en la que el espacio y el tiempo son abolidos, y los hombres tienen que aprender a vivir en estrecha relación”. Este paradigma llamado *Aldea Global* fue acuñado en 1964, por Marshall McLuhan quien plantea el concepto de un mundo interrelacionado, con unificación de propósitos y objetivos sociales, políticos y económicos. Concibe la transformación de las sociedades, producto de los constantes y acelerados cambios de las comunicaciones, hacia una intercomunicación inmediata y fluida.

La idea de McLuhan, está representada y caracterizada en nuestros tiempos por las sociedades de la información, término utilizado a mediados de los años 70's por Daniel Bell, quien planteó el paso de una sociedad de bienes a una sociedad de conocimiento. El principal objetivo de las sociedades de la información es la transformación de la información en conocimiento de valor social y económico, a través del cambio en las formas de gestión del conocimiento. El concepto de sociedades de la información implica entonces, el establecimiento de nuevas estructuras sociales que generen cambios incidentes en todas las áreas, desde las formas de comercio, los métodos de aprendizaje hasta las prácticas de gobierno.

Según Fernando Giner y María de los Angeles Gil, una sociedad de la información es *“aquella sociedad que ordena, estructura su funcionamiento (modo de vida, forma de relaciones, modo de trabajo, etc.) en torno a las tecnologías de la información y la comunicación y convierte a la información en un factor de producción, intercambio y conocimiento”*. Es importante entonces reconocer el impacto y la trascendencia que tienen las tecnologías de la información para la construcción de sociedades de la información sobre todos los elementos que componen las sociedades: organizaciones, individuos, empresas y estados.

Las sociedades de la información sustentan su funcionamiento en tres elementos fundamentales: la tecnología, la información y el conocimiento que interactúan y retroalimentan de forma constante. La información representada en los datos de valor, la cual puede ser almacenada, manipulada, transmitida, depurada y consultada a través de las tecnologías de la información y comunicación, mismas que facilitan la realización de transformaciones para la obtención de conocimiento relevante que permita la generación de nuevos datos y por lo tanto de información.

Los conceptos de aldea global y sociedades de la información se han visto confirmados con el surgimiento de internet y el vertiginoso desarrollo que desde entonces, han tenido las tecnologías de la información y la forma en que estas han contribuido en el intercambio instantáneo de información. En el marco de esta globalización de la información y las comunicaciones, las nuevas tecnologías de la información, como canal de soporte, han revolucionado la forma en que los seres humanos interactúan entre sí y con su entorno. Se define como Tecnologías de la Información y Comunicación (TIC's), al conjunto de herramientas informáticas y computacionales que permiten el almacenamiento, procesamiento, acceso, recuperación y presentación de información, y que se han convertido en los nuevos medios de desarrollo.

Es importante reconocer entonces el alto impacto social que las tecnologías de la información y comunicación tienen dentro de las sociedades de la información en entornos tales como el empresarial, educativo, investigativo, gubernativo, entre otros.

En el ambiente empresarial las tecnologías de la información, más que una opción se han convertido en una necesidad, producto de la alta competitividad de los mercados y la agilización de los procesos productivos a través del uso de TICs. En este sentido podemos encontrar las variadas aplicaciones de sistemas de software y hardware que se desarrollan constantemente para la tecnificación y sistematización de los procesos y procedimientos empresariales, como son los ERPs, CRMs, Business Intelligence, wireless, sistemas de backups, virtualización, entre otros.

En el ámbito educativo, las TICs han facilitado la comunicación producto de los diversos formatos y métodos en que los contenidos pueden ser impartidos, además de proveer a los estudiantes de un alto grado de información sobre diversos temas. Aplicaciones como blogs, wikis, LMS, multimedia, CMS, podcast, plataformas de e-learning, entre otros, permiten el intercambio interactivo de conocimiento permitiendo a través del acceso a internet la gestión del capital humano.

La investigación también se ha visto impactada por las tecnologías de la información, el procesamiento de información y la realización de experimentos se ha convertido en tarea fácil, producto de las capacidades desarrolladas en los dispositivos que cada vez son más potentes y veloces. Además las sociedades científicas pueden comunicarse de forma más fluida y de este modo realizar intercambio instantáneo de resultados y datos relevantes, así como facilidad de acceso a la información contenida en espacios digitales y bases de datos documentales. Sobre este tema Barro Ameneiro plantea que las comunidades científicas en la era de las TICs son *“comunidades de conocimiento globales en las que los intercambios de resultados y descubrimientos se realizan casi de forma automática, favoreciendo el conocimiento y la difusión”*.

En el área gubernativa, la administración pública debe aplicar enfoques diferentes de servicio público al ciudadano, proveyendo a la población de nuevas formas de accesibilidad para la realización de trámites así como vías alternativas para la participación en la toma de decisiones. Esto puede verse reflejado desde elementos como el voto electrónico, hasta la generación de aplicaciones de tramitación vía internet.

2. INTELIGENCIA DE NEGOCIOS

2.1. Sistemas de Información

Los sistemas de información surgen producto de la necesidad del ser humano de controlar, mantener y medir su entorno a través del procesamiento de información y han evolucionado, hacia el soporte eficiente de los procesos operativos, estratégicos y organizacionales, a medida que aumentan las necesidades humanas, la información se vuelve más compleja y se desarrollan nuevas tecnologías y aplicaciones de software.

Están constituidos por cuatro elementos vitales que interactúan entre sí, procesando datos que convierten en información de utilidad y distribuyen de la forma más adecuada, para el cumplimiento de los objetivos estratégicos de la organización, abarcando todas las áreas del conocimiento según las características y necesidades específicas de cada una. En la siguiente figura se plasma la interacción entre personas-recursos-datos-reglas como unidades integrales de los sistemas de información.

Figura 1. Elementos de un sistema de información



Las bases teóricas de los sistemas de información provienen desde inicios del siglo XX, cuando el biólogo austriaco Ludwig Von Bertalanffy instaura el paradigma científico llamado Teoría General de Sistemas, y define un sistema como el “conjunto de elementos en interacción para lograr un objetivo común”.

Según Börje Langefors un sistema de información se trata de un “sistema generalmente incluido en otro mayor, que se ocupa de recibir, almacenar, procesar y distribuir información”, siendo su propósito la gestión del conocimiento. Por su parte James Senn establece que “la finalidad de un sistema es la razón de su existencia... las finalidades de los sistemas de información, como las de cualquier otro sistema dentro de una organización, son procesar entradas, mantener archivos de datos relacionados con la organización y producir información, reportes y otras salidas... los sistemas de información están formados por subsistemas que incluyen hardware, software, medios de almacenamiento de datos para archivos y bases de datos...”. Adicionalmente Kast y Rosenzweig establecen que la “tecnología es la organización y aplicación de conocimiento para el logro de fines prácticos. Incluye manifestaciones físicas como las máquinas y herramientas, pero también técnicas intelectuales y procesos utilizados para resolver problemas y obtener resultados deseados”.

Puede definirse entonces que, los sistemas de tecnología de la información son un subsistema de los sistemas de información y se constituyen en un recurso material, conformado por elementos de hardware, firmware y software, que facilitan y agilizan de forma contundente la transformación de datos en información, siendo sus objetivos principales el procesamiento de entradas, mantenimiento de los datos, producción de resultados, manejo de consultas, administración de transacciones diarias, generación de reportes y entradas para otros sistemas.

Los sistemas de información para cumplir con sus funciones realizan al menos 3 actividades fundamentales.

Figura 2. Actividades de los sistemas de información



- **Entrada:** El principal objetivo de esta actividad consiste en recopilar los datos requeridos para ser procesados, los cuales pueden provenir de fuentes diversas tales como teclados, cds, cintas magnéticas, códigos de barra, entre otros.
- **Transformación:** Esta etapa tiene como propósito efectuar cálculos, operaciones y variaciones de los datos en base a las reglas y estrategias del negocio para convertir los datos en información comprensible y de interés.
- **Salida:** Muestra los resultados obtenidos en los formatos y componentes requeridos como son impresoras, gráficos, archivos o bien otros módulos del sistema que utilizan esta información para poder realizar su proceso interno de recopilación, transformación y salida.

Contar con los datos es insuficiente para la toma de decisiones, la información cobra valor al ser transformada en conocimiento luego de ser conocida, interpretada y evaluada por seres humanos, con base a criterios y experiencias particulares y según mapas mentales individuales y organizacionales.

La evolución de los sistemas tradicionales de intercambio de información hacia nuevos paradigmas planteados por la era de la información, ha transformado las organizaciones en entidades de conocimiento que proveen la infraestructura necesaria para que sus integrantes compartan la información, dando lugar a nuevos conceptos, metodologías y ventajas competitivas. La combinación entre esta apertura y el desarrollo acelerado de las tecnologías plantea para las organizaciones, nuevas formas eficientes de interactuar con su entorno y gestionar el conocimiento, adoptando prácticas organizativas que respondan y se adapten a las necesidades cambiantes del entorno donde se desarrollan.

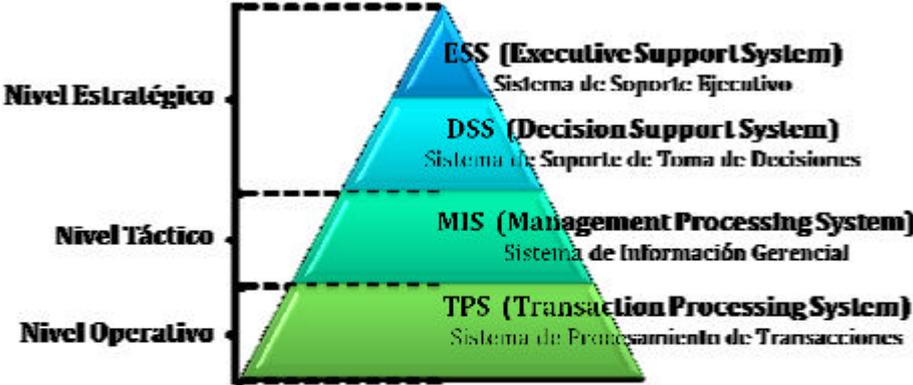
Es producto de estos cambios que las organizaciones deben establecer los procedimientos, medios y métodos que les permitan alcanzar sus objetivos a corto, mediano y largo plazo, esto se logra a través de la planeación. Planear según John Ivancevich significa que *“las decisiones que hoy se adopten producirán resultados útiles en alguna fecha futura, resultados que se desprenden de la finalidad y los objetivos de la organización”* y Mintzberg establece que *“la planeación estratégica no es más que el proceso de relacionar las metas de una organización, determinar las políticas y programas necesarios para alcanzar objetivos específicos en camino hacia esas metas y establecer los métodos necesarios para asegurar que las políticas y los programas sean ejecutados, o sea, es un proceso formulado de planeación a largo plazo que se utiliza para definir y alcanzar metas organizacionales”*.

Administrar organizaciones dentro de este contexto de planificación y adaptabilidad requiere del establecimiento de tres niveles de gestión con objetivos, necesidades, responsabilidades y características específicas.

- **Nivel Operativo:** El principal objetivo de este nivel de gerencia radica en la ejecución de las actividades diarias de la organización, está integrado por las jefaturas responsables de la programación, control y ejecución de las operaciones básicas, requiere de información relacionada con el seguimiento y planificación de las tareas cotidianas, referida al corto plazo y alto nivel de detalle.
- **Nivel Táctico:** Este nivel de administración tiene como propósito fundamental la instrumentación de lineamientos de corta duración que pueden irse ajustando y modificando a medida que se producen alteraciones en el entorno, requiere de información consolidada sobre el funcionamiento interno de los subsistemas que componen la organización, referida al mediano plazo.
- **Nivel Estratégico:** La finalidad de este nivel de gestión es la elaboración de planes que integren directrices, objetivos, políticas, metas y acciones, que tomando en cuenta factores externos predecibles e impredecibles, permitan a la organización establecer ventaja competitiva al adaptarse, transformar y aprovechar los cambios que se suscitan en su entorno, requiere de información externa de tendencias y cómo la organización puede verse afectada con estas, referida al largo plazo y bajo nivel de detalle.

Con el transcurso del tiempo y a medida que las necesidades e intereses de las organizaciones se han vuelto más complejos y específicos, cada nivel de decisión y gestión de la organización requiere de información orientada al apoyo de sus procesos, para el cumplimiento de sus funciones, los sistemas de información se han transformado y especializado para soportar efectivamente cada uno de estos niveles, esta relación puede observarse en la siguiente figura.

Figura 3. Niveles de gestión organizacional y su relación con los sistemas de información



Los distintos sistemas de información, como se muestra en la figura anterior, se encuentran escalonados en una pirámide jerárquica, donde cada uno de los niveles inferiores provee los datos a los niveles subsiguientes, por lo que existe una estrecha relación entre cada uno de ellos.

2.1.1. TPS – Sistema de Procesamiento de Transacciones

Estos sistemas de información tienen como objetivo la ejecución y registro de las transacciones diarias de la organización, representando el pilar sobre el cual se soporta el funcionamiento y estructuración de la misma, sirviendo como elemento de apoyo a la gerencia operativa. Generan ahorros significativos a través de la automatización de tareas, intercambio frecuente de información y manejo de grandes volúmenes de datos. Ejemplos de este tipo de sistemas están Facturación, Recursos Humanos, Sistemas de Ventas, Nómina, Bancos, entre otros.

2.1.2. MIS – Sistema de Información Gerencial

Proveen de los elementos necesarios para la gerencia administrativa de las organizaciones, a través de reportes de actividades semanales, mensuales o anuales orientados a la información interna, siendo un factor fundamental de apoyo a la planificación, gestión y control de componentes de la organización, tales como materiales, proyectos, personal, entre otros. Ejemplo de estos sistemas tenemos los ERP's (Enterprise Resource Planning) que permiten la administración y planificación de recursos, así como los CRM's (Customer Relationship Management) los cuales permiten gestionar la relación con los clientes.

2.1.3. DSS – Sistema de Soporte de Toma de Decisiones

La toma de decisiones estratégicas de las organizaciones implica una alta intervención humana basada en elementos teóricos, factores externos y conocimiento tácito.

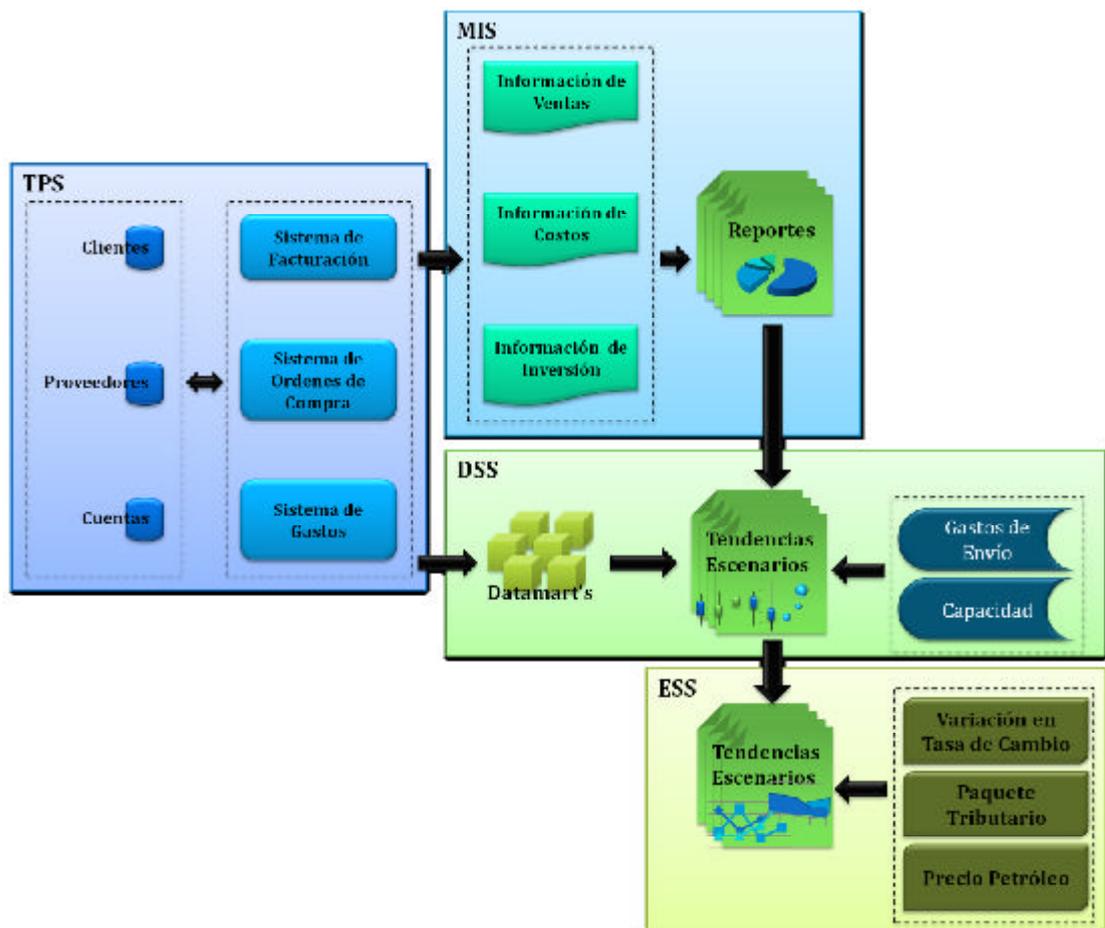
Los sistemas de soporte a toma de decisiones se convierten en un elemento más de apoyo, que facilita la toma de decisiones ante situaciones únicas, cambiantes e imprevistas, para lo cual hace uso de la información provista por los TPS y MIS. Su función principal es proveer escenarios de interpretación entre alternativas basadas en estimaciones generando así un marco de comparativas para las partes encargadas de la toma de decisiones, basados en el conocimiento histórico previo de las situaciones vividas por la organización y los elementos internos de la misma que permiten establecer variación en el comportamiento. Dentro de estos sistemas se encuentran las herramientas de BI (Business Intelligence), las cuales permiten establecer un conocimiento cognoscitivo basado en el comportamiento histórico de la organización.

2.1.4. ESS – Sistema de Soporte Ejecutivo

Este tipo de sistemas de información tienen como principal función presentar información sintetizada a gerencias y altos mandos, quienes tienen la responsabilidad de establecer la planificación estratégica y objetivos de la organización. A través de este tipo de sistemas los ejecutivos pueden monitorear el desempeño de la organización, seguimiento del comportamiento de los agentes externos, detectar problemas, identificar tendencias, no están orientados a resolver problemas específicos, sino más bien sirven para definir las acciones a largo plazo que debe tomar la organización para alcanzar sus objetivos. Este tipo de herramientas permiten la identificación de patrones, tendencias, temporalidades y realización de proyecciones que permitan definir el modelo organizacional a seguir para alcanzar propósitos específicos.

En la siguiente figura se muestra la forma de intercambio entre los sistemas de información, puede observarse cómo: los datos generados por los sistemas de procesamiento de transacciones se convierten en los reportes de información para los sistemas de información gerencial. Estos reportes, a su vez, en conjunto con modificaciones de información interna se convierten en los resultados necesarios para la toma de decisiones. Por último, estos escenarios incorporados a elementos externos permiten proveer los elementos para los sistemas de soporte ejecutivo.

Figura 4. Relación e intercambio de información entre sistemas de IT

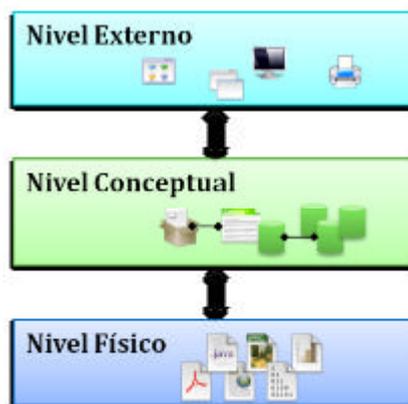


2.2. Sistemas Gestores de Bases de Datos

La aparición de las bases de datos a mediados de la década de los 70's y el asentamiento de la teoría relacional, suscitó el surgimiento de un nuevo paradigma sobre la gestión y almacenaje de los datos y la forma en que estos son manipulados. Se conoce como base de datos al conjunto de datos relacionados entre sí, almacenados de forma sistemática y que son de interés para la organización. Los Sistemas Gestores de Base de Datos (DBMS – Data Base Management System) son herramientas que permiten el almacenamiento, manipulación y estructuración de grandes cantidades de datos que son definidos una sola vez y que pueden ser accedidos de forma concurrente, sirviendo de interfaz de comunicación con los usuarios u otros programas o aplicaciones.

Los DBMS se definen a través de una arquitectura de 3 niveles, que desempeñan funciones específicas que les permiten complementarse para cumplir de forma óptima los objetivos del programa. Siendo responsabilidad del DBMS administrar los flujos de información, las transformaciones y correspondencias que sean necesarias entre los diferentes niveles.

Figura 5. Arquitectura del DBMS



- **Nivel físico:** Se constituye en el nivel más bajo de abstracción, estableciendo el esquema real de almacenamiento de datos a nivel de archivos, así como los métodos de acceso a los mismos.
- **Nivel conceptual:** Establece la taxonomía, relaciones y distribución de la información de la organización a través del modelo lógico de datos.
- **Nivel externo:** Define las diferentes porciones de información del nivel conceptual que son accedidas por los grupos de usuarios según el área de interés.

Además de la arquitectura física, los DBMS deben cumplir con un conjunto de características que le permiten gestionar la información y la forma en que los niveles se relacionan entre sí:

- **Abstracción** La forma en que los datos son almacenados y estructurados físicamente son transparentes para el usuario.
- **Independencia:** Capacidad de realizar modificaciones en los esquemas de los niveles lógicos y físicos sin que sea necesaria la realización de cambios en las aplicaciones que hace uso de estos.
- **Redundancia:** A través de un diseño conceptual correcto se evita al máximo la repetición o redundancia de información.
- **Consistencia:** Implica que los datos almacenados se mantengan actualizados de forma armónica y simultánea, a fin de mantener la misma información en todos los elementos.
- **Seguridad:** Garantizar que la información se encuentra protegida de usuarios malintencionados que intenten leer información que no les corresponde.

- **Integridad:** Facultad de proteger los datos almacenados de fallos que puedan corromperlos.
- **Respaldo y recuperación:** Proveer de forma eficiente herramientas para respaldo y reinstalación de copias de seguridad que permitan establecer un punto de restauración.
- **Control de concurrencia:** Manejo de sesiones y accesos simultáneos a los datos contenidos en la base de datos.
- **Tiempo de respuesta:** Minimización del tiempo que tarda el DBMS en realizar una tarea y enviar la respuesta o almacenar la información.

Los sistemas gestores de bases de datos son clasificados, según el modelo abstracto mediante el cual almacenan y recuperan la información del contenedor de datos a través de algoritmos y conceptos matemáticos. En la actualidad, los tipos de gestores de bases de datos que se constituyen en los más utilizados para los procesos empresariales y organizacionales son los relacionales y multidimensionales, donde cada uno se enfoca en áreas específicas y soportan diferentes sistemas de información.

- **Bases de datos jerárquicas:** Almacena información en estructuras jerárquicas enlazando los registros en forma de estructura de árbol donde un nodo padre puede tener varios nodos hijos.
- **Bases de datos de red:** Los registros están representados a través de un diagrama de grafo de red dirigido y los nodos contienen información sobre las relaciones.

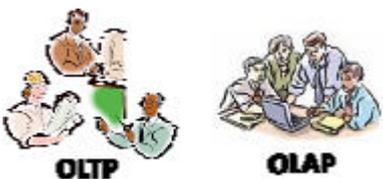
- **Bases de datos relacionales:** Definido a través de un modelo relacional que permite la realización de operaciones sobre estructuras de registros y campos almacenados en tablas que se vinculan entre sí. Se constituyen en el principal soporte para los TPS permitiendo el desarrollo de los procesos diarios empresariales a través del Procesamiento de Transacciones en Línea (OLTP – Online Transaction Process).
- **Bases de datos de objetos:** Los elementos están organizados a través de objetos y permite el manejo de conceptos de programación orientada a objetos como abstracción, herencia, generalización, encapsulamiento y poliformismo entre otros.
- **Bases de datos multidimensionales:** Manejado a través de cubos de información que conforman las dimensiones donde cada valor almacenado corresponde a una métrica y sirve de núcleo de los DSS para el manejo de grandes cantidades de información a través del Procesamiento Analítico en Línea (OLAP – Online Analytical Processing).

2.3. Online Transaction Process y Online Analytical Processing

Los primeros DBMS del mercado tenían como principal propósito establecer las condiciones para el procesamiento de datos a través de sistemas transaccionales, sin embargo, en los últimos años éstos han ido evolucionando para satisfacer las necesidades empresariales de análisis y toma de decisiones. Tanto el área transaccional como analítica son de suma importancia para el correcto funcionamiento de las empresas en el mundo globalizado actual, pero producto de las grandes diferencias y debido a la intrínseca relación y dependencia entre ambas, se han generado dos formas de soporte técnico que permita la generación de un entorno idóneo para su puesta en funcionamiento. Establecer las diferencias que existen tanto en aspectos de diseño como elementos de funcionalidad representa el elemento fundamental para comprender a fondo el propósito de cada uno y su ambiente de utilización.

- **Online Transaction Process (OLTP)** o Procesamiento en línea de transacciones, se define como el proceso mediante el cual las peticiones realizadas por los usuarios son atendidas de manera inmediata y registrada de manera individual.
- **Online Analytical Processing (OLAP)** o Procesamiento analítico en línea, se define como el proceso que permite analizar de forma dinámica la información a través de estructuras multidimensionales.

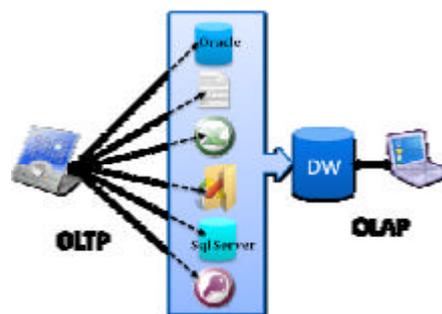
Tabla I Cuadro comparativo de tecnologías OLAP y OLTP

OLAP - OLTP
<p>Objetivos de construcción</p> <p>Los sistemas OLTP sustentan aplicaciones específicas y mantienen la integridad de los datos a través de la administración de las transacciones diarias de la organización, mientras que los sistemas OLAP están orientados a la realización de análisis del negocio, identificación de tendencias, pronósticos a través de la interpretación de datos históricos.</p> <p>Figura 6 OLAP - OLTP según objetivos de construcción</p>  <p>El diagrama ilustra la arquitectura de los sistemas OLTP y OLAP. A la izquierda, tres dispositivos móviles etiquetados como 'Vendedor' están conectados por líneas a un servidor etiquetado como 'Servidor', con el texto 'OLTP' debajo. A la derecha, un cilindro de datos etiquetado como 'DW' está conectado por una línea a un dispositivo etiquetado como 'Analista de Negocios', con el texto 'OLAP' debajo.</p>
<p>Perfil de usuarios</p> <p>En cuanto al perfil de los usuarios que utilizan los sistemas OLTP son los cuadros operativos de la organización que interactúan con las aplicaciones específicas según el área de competencia, en cambio los sistemas OLAP están orientados a las altas gerencias quienes están a cargo de la toma de decisiones.</p> <p>Figura 7. OLAP - OLTP según perfil de usuarios</p>  <p>El diagrama muestra dos grupos de personas. A la izquierda, un grupo de personas etiquetado como 'OLTP' que parecen ser cuadros operativos. A la derecha, un grupo de personas etiquetado como 'OLAP' que parecen ser miembros de alta gerencia.</p>

Integración de datos

Los datos en sistemas OLTP se encuentran distribuidos en diferentes tipos de archivos, formatos, plataformas, dbms según su función y propósito dentro de la organización. Los sistemas OLAP deben tener integrado en un único repositorio con la misma estructura proveniente de los diferentes datos operacionales OLTP.

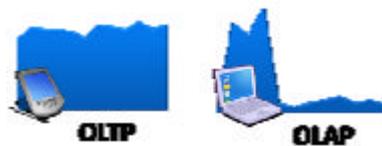
Figura 8. OLAP - OLTP según integración de los datos



Modelos de frecuencia de acceso

Los sistemas transaccionales OLTP generalmente mantienen un patrón constante de accesos requiriendo consumo de grandes cantidades de recursos y bajos tiempos de respuesta para la realización de cada transacción, por el contrario los sistemas OLAP presentan un patrón de uso sutil con crestas de usos eventuales producto del flujo de trabajo y la disponibilidad de los datos.

Figura 9. OLAP - OLTP según frecuencia de acceso



Orientación de los datos

En los sistemas transaccionales los datos están alineados según la aplicación, teniendo la estructura necesaria para cumplir con las necesidades específicas de la aplicación que alimentan, en cambio los datos de análisis se encuentran integrados y separados lógicamente según dimensión o área temática de la organización.

Figura 10. OLAP - OLTP según orientación de los datos



Manipulación de datos

El ingreso de datos registro a registro, la manipulación constante y el intercambio y modificación de la información almacenada es característica fundamental de los sistemas OLTP, en contraste, los sistemas OLAP solo tienen acceso para consulta de la información almacenada, la cual solo es alimentada ocasionalmente por procesos incrementales.

Figura 11. OLAP - OLTP según manipulación de datos



Historia

Los sistemas OLTP retienen datos para un período entre los 60 y 90 días producto que el tiempo no es un elemento clave para su correcto funcionamiento, al contrario de los sistemas OLAP que se alimentan de los períodos históricos relegados por los sistemas OLTP, información que sirve de base para la realización de análisis de tendencias.

Figura 12. OLAP - OLTP según historia



Las plataformas de implementación y almacenamiento de sistemas OLAP han ido evolucionando en base a capacidad de almacenamiento, estructuras internas, técnicas de recuperación de información, funcionalidad y rendimiento, acoplándose a las necesidades crecientes de los mercados actuales y las formas de manejo y nivel de detalle de la información requerida en las organizaciones.

Tabla II Arquitecturas OLAP

Arquitecturas OLAP	
MOLAP	<p>La arquitectura MOLAP (Multidimensional OLAP) es la primera en aparecer en la década de los 80's. Su principal característica es la utilización de estructuras multidimensionales (matrices y vectores) para el almacenamiento de los datos.</p> <p>Características principales:</p> <ul style="list-style-type: none">• Tiempos de respuestas imperceptibles• Pre-cálculo de intersecciones entre dimensiones• Alta compresión de datos• Limitada flexibilidad• Poca capacidad de escalabilidad• Escasa facultad de agregación en tiempo de ejecución• Baja capacidad de manejo de grandes volúmenes de datos

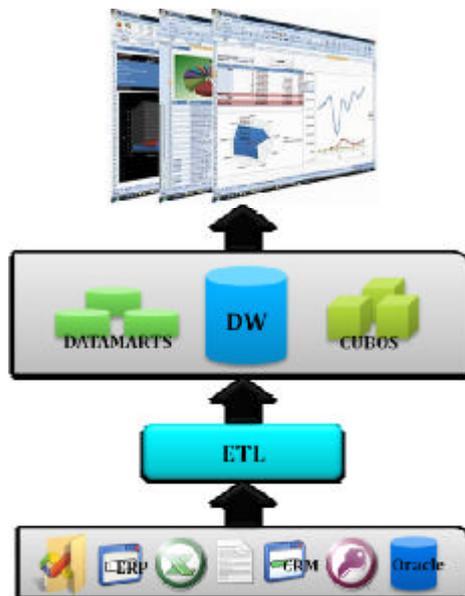
<p>ROLAP</p>	<p>A comienzo de los 90's nacen los ROLAP (Relational OLAP) como nueva alternativa para la creación de bodegas de datos. Establece el ambiente multidimensional sobre una base relacional estructurado en registros y tablas que son consultadas por herramientas de simulación de accesos multidimensionales.</p> <p>Características principales:</p> <ul style="list-style-type: none"> ✓ Altos tiempos de respuesta ✓ Necesidad de uso de herramientas de mejoramiento de performance ✓ Fácil escalabilidad ✓ Amplia flexibilidad ✓ Rápida capacidad de actualización ✓ Manejo de grandes volúmenes de información ✓ Facultad de análisis transaccional
<p>HOLAP</p>	<p>Este tipo de arquitectura HOLAP (Hybrid OLAP) surge de combinar la eficiencia de velocidad de los sistemas MOLAP y la capacidad de análisis y almacenaje de los sistemas ROLAP, para lo cual utiliza cubos pre-calculados como núcleo de la información y cuando faltan datos de consulta busca la respuesta en el modelo relacional.</p>

2.4. Datawarehousing

Bill Inmon fue el primero en utilizar el término datawarehouse a principio de la década de los 90's definiéndolo como *“una colección de datos orientados a elementos específicos de la organización, integrados desde diferentes fuentes, asociados a un período de tiempo, no volátiles e invariables con el tiempo”*. Según Ralph Kimball, pionero en el campo de la inteligencia de negocios y datawarehouse, las bodegas de datos son *“una copia de los datos transaccionales específicamente estructurados para consultas y análisis”*.

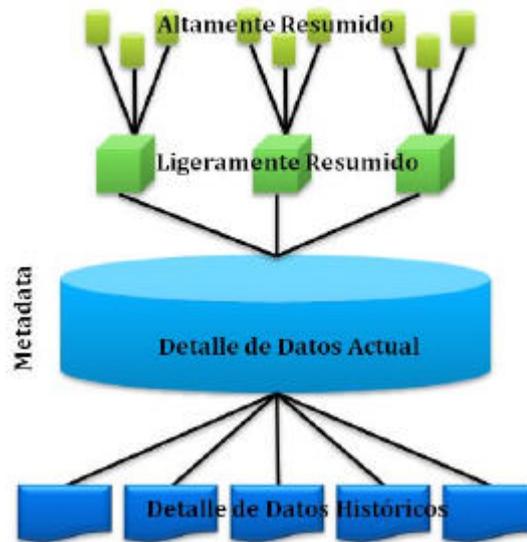
Puede definirse entonces un datawarehouse o bodega de datos como el conjunto de herramientas, metodologías y tecnologías que permiten la manipulación, utilización y construcción de los sistemas de apoyo a la toma de decisiones, que tiene como base de alimentación los sistemas de procesamiento de transacciones.

Figura 13. Esquema del Datawarehouse



Las bodegas de datos están estructuradas en diferentes niveles de esquematización y detalle que le delimitan:

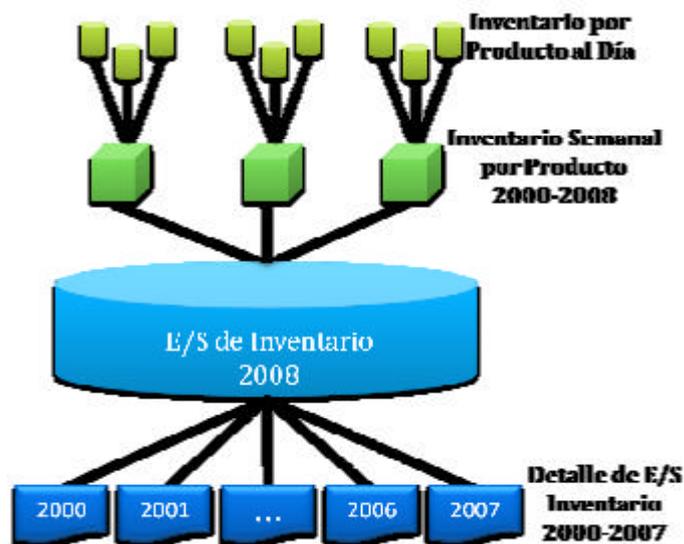
Figura 14. Esquema estructural del Datawarehouse



- **Detalle de datos actual:** Refleja el conjunto de datos más importantes para la organización porque representan las operaciones más recientes de la misma, tienen un alto grado de detalle, volumen moderado y son frecuentemente accedidos.
- **Detalle de datos históricos:** Son el conjunto de datos antiguos que se encuentran almacenados en estructuras de almacenamiento masivo secundario. Representan grandes volúmenes de datos, que no son frecuentemente accedidos por los sistemas operativos transaccionales, cuenta con un nivel de detalle consistente con los datos actuales, constituye la fuente principal de alimentación del datawarehouse.

- **Datos ligeramente resumidos:** Este tipo de datos representa el primer nivel de agregación correspondiente a consultas habituales predeterminadas.
- **Datos altamente resumidos:** Son el conjunto de datos compactos que representan la información que va a ser accedida constantemente y los cálculos más detallados sobre el rendimiento de la organización que establecen los valores totales necesarios para la toma de decisiones.
- **Metadata:** Este tipo de datos no proviene de los datos operacionales transaccionales de la organización y tiene como principal función proveer el marco esquemático que permita la ubicación de contenidos, proporcionar el conjunto de algoritmos que establezca la interacción entre los diferentes niveles de datos estructurales, siendo el motor que administra la transformación de los datos operacionales en datos del datawarehouse.

Figura 15. Ejemplo de niveles de esquematización



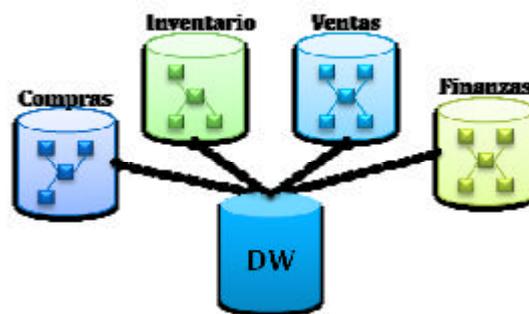
La construcción de un Datawarehouse requiere la definición de estrategias de Extracción-Transformación-Carga (ETL) que permitan la alimentación periódica de los cubos de información con datos provenientes de fuentes heterogéneas. Es importante realizar un análisis completo de las necesidades, pasos y reglas del ETL para obtener un sistema operacional bien diseñado y que genere información confiable y consistente. Además es de suma importancia identificar los niveles de escalabilidad que puedan tener los datos y diseñar los procedimientos correctos para la realización óptima del ETL en cuanto a volúmenes de datos y tiempos de ejecución, esto según las características y metodologías de la organización.

- **Extracción:** Esta primera parte del proceso implica la depuración, integración y análisis de los datos provenientes de los sistemas de origen, los cuáles generalmente están almacenados en sistemas y formatos distintos que pueden ir desde archivos planos hasta aplicaciones a la medida. La tarea de la extracción es establecer los datos que son de utilidad, que cumplen con el formato establecido, garantizan la estructura requerida para luego obtenerlos sin causar impacto en el rendimiento de los sistemas de origen.
- **Transformación:** Este paso es el núcleo principal del proceso y se encarga de disponer y aplicar las reglas del negocio, expresadas en funciones, a los datos obtenidos para que estos cumplan con los estándares necesarios para su almacenaje y posterior utilización.
- **Carga:** Durante esta fase los datos que han sido transformados son cargados en el sistema destino (ROLAP/MOLAP/HOLAP) interactuando directamente con la base de datos destino. Hay 2 tipos de cargas que pueden ser realizadas y la selección de una de ellas dependerá del nivel de detalle con el que se cuente en la bodega de datos:

- **Acumulación Simple:** Este tipo realiza un resumen de todas las transacciones por periodo de tiempo y traslada un único registro al cubo de información, reflejado en un valor calculado que generalmente representa una sumatoria o promedio de valores.
- **Rolling:** Este tipo refleja todas las transacciones realizadas dependiendo del nivel de granularidad y taxonomía con el que se debe contar en el datawarehouse.

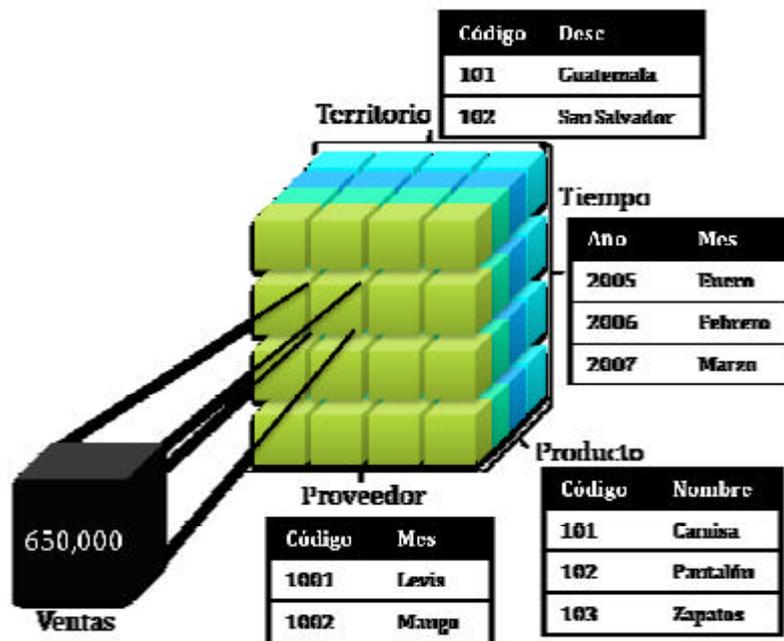
Al igual que las organizaciones se encuentran divididas en departamentos o áreas que cumplen funciones y actividades específicas el datawarehouse de una organización se encuentra segmentado en áreas especializadas de datos que representan un área del negocio donde cada elemento es conocido como Datamart. Ralph Kimball definió el datawarehouse como *“la unión de todos los datamarts de una entidad”*, en resumen un datamart es un subconjunto de datos de un datawarehouse enfocados al soporte de toma de decisiones sobre un área específica de la empresa.

Figura 16. Integración Datamarts -Datawarehouse



La esencia del almacenamiento y navegación OLAP son los cubos de información, los cuales están organizados a través de dimensiones que representan la estructura de los atributos de las áreas temáticas de la organización y jerarquías que describen la taxonomía de la misma. Esto permite el análisis de la información en diferentes niveles de agregación que califica el conjunto de resultados para que puede ser accedida a través de tablas dinámicas o programas especializados.

Figura 17. Cubos de información OLAP



La construcción de un modelo multidimensional requiere la integración de varios componentes básicos:

- **Dimensiones:** Son una representación conceptual o lógica de las áreas de la organización definidas a través de atributos independientes permitiendo el establecimiento de perspectivas múltiples.

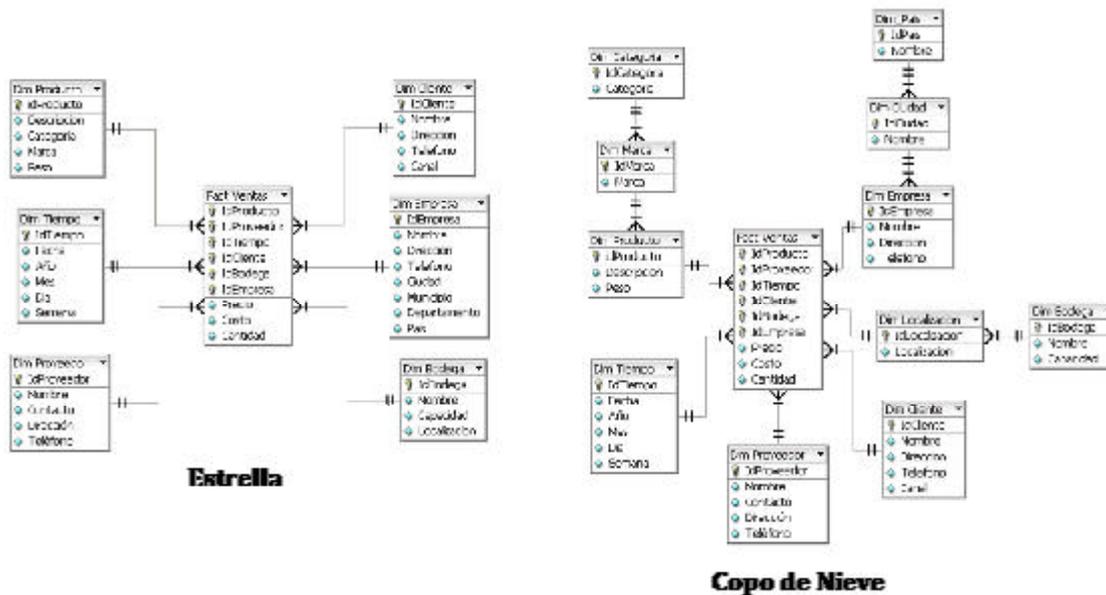
- **Atributos:** Conjunto de elementos de una dimensión agrupados según su categorización o clasificación que representa el nivel de granularidad de la información y su grado segregación.
- **Entidades:** Instancias de los atributos que representan los componentes atómicos del modelo.
- **Relaciones:** Asociaciones lógicas de los atributos a través de la misma dimensión, en cambio las relaciones entre dimensiones se realizan a través de los indicadores o métricas del negocio.
- **Jerarquías:** Establecen el ordenamiento lógico dentro de la dimensión formada por las relaciones existentes entre los atributos de la dimensión.
- **Métricas:** Son el conjunto representaciones que permiten clasificar y medir el rendimiento de la organización con base a los criterios y reglas de la misma.
- **Tabla de hechos:** Se constituye en el elemento principal del modelo siendo la que almacena las métricas e indicadores de la organización, relacionando las diferentes dimensiones establecidas.

Para la construcción del modelo multidimensional existen 2 esquemas físicos:

- **Estrella:** Este modelo tiene como principal característica que parece una estrella, con una tabla de hechos en el centro conteniendo los datos de análisis y rodeada de las tablas de dimensiones. Este esquema es simple y veloz permitiendo la implementación de la funcionalidad multidimensional utilizando como base un modelo relacional.

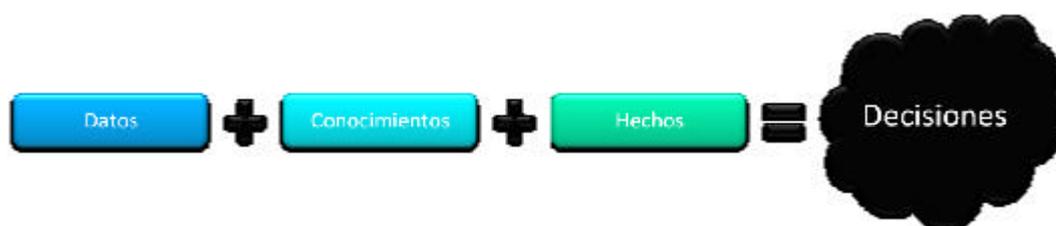
- **Copo de nieve:** Modelo más complejo que el esquema estrella donde las tablas de dimensiones se encuentran implementadas con más de una tabla de datos cuya finalidad es la normalización de las tablas y la reducción de espacio de almacenamiento eliminando redundancia de datos.

Figura 18. Esquemas físicos multidimensionales



Para que el datawarehouse produzca resultados es necesaria la utilización de los datos contenidos en él y transformarlos en información que sea útil en el análisis de escenarios que permitan determinar el estado y el rumbo de la organización. El conjunto de herramientas y estrategias que proporcionan conocimiento a través del análisis de los datos históricos de la organización, con el fin de comprender su funcionamiento actual y establecer con anticipación planes estratégicos de acción, es lo que se conoce como Inteligencia de Negocios (Business Intelligence).

Figura 19. Elementos de BI



Escoger la herramienta correcta para la implementación e implantación de sistemas de inteligencia de negocios es una decisión importante que dependerá de aspectos como estrategia organizativa, compatibilidad con sistemas existentes, cultura organizacional, nivel de complejidad y capacidad adquisitiva de la organización, por lo que, establecer la infraestructura tecnológica apropiada se convierte en un factor fundamental para el éxito o fracaso del proyecto.

2.5. Datamining

Datamining o Minería de Datos, también conocida como Knowledge Data Discovery, es un proceso que integra disciplinas como la inteligencia artificial y la estadística, con el fin de encontrar relaciones complejas entre los datos, que permitan la identificación de patrones de comportamiento. Según C. J. Date la minería de datos puede describirse como *“análisis de datos exploratorio... el propósito es buscar patrones interesantes en los datos, patrones que pueden usarse para especificar la estrategia del negocio o para identificar comportamiento fuera de lo común”*. A su vez Usama Fayyad lo define como *“un proceso no trivial de identificación válida, novedosa, potencialmente útil y entendible de patrones comprensibles que se encuentran ocultos en los datos”*.

Para el análisis de los datos, la minería de datos utiliza algoritmos complejos que son aplicados a los datos con el fin de obtener resultados, a través de técnicas de la inteligencia artificial y la estadística, a continuación se identifica algunas.

Tabla III Técnicas de minería de datos según área de especialización

Área	Técnica
Estadística	<ul style="list-style-type: none"> • Análisis de Varianza: Compara si el conjunto de valores numéricos de una población es representativamente diferente de los de otra. • Regresión: Predice una medida basándose en los parámetros de conocimiento de otra. • Clustering: Agrupa y clasifica un conjunto de individuos según criterios cualitativos o cuantitativos con base a la identificación de semejanzas y diferencias. • Series de Tiempo: Secuencia de puntos de datos medidos en intervalos sucesivos de tiempo que permite la extracción de información representativa.

<p style="text-align: center;">Inteligencia Artificial</p>	<ul style="list-style-type: none"> • Algoritmos Genéticos: Métodos numéricos de algoritmos que tienen la capacidad de mutar, evolucionar y modificar su comportamiento en base a criterios definidos y probabilidades. • Árboles de Decisión: Construcción de diagramas lógicos basados en reglas que pueden representar un conjunto de condiciones. • Redes Neuronales: Metodología que permite la interacción entre variables a través de transformaciones lineales y no lineales hasta obtener salidas.
---	--

La minería de datos puede ser aplicada en cualquier área o sector, desde la identificación de gustos y preferencias de los usuarios según temporadas hasta distinguir estrategias de fraudes a tarjetas de créditos, a través de la interpretación de las relaciones, cambios, dependencias y variaciones de la información generada por los procesos de la institución.

3. DATAWAREHOUSE ELECTORAL

3.1. Introducción

Uno de los pasos más importantes para la efectiva y óptima implementación de un datawarehouse recae en establecer el modelo lógico conceptual del mismo que soporte el almacenamiento y manipulación de la información. La instauración de las relaciones y jerarquías de cada dimensión así como la tabla de hechos son el pilar fundamental para que la interpretación de los datos arroje información relevante y que pueda ser interpretada para la toma de decisiones.

A continuación se detallan los elementos principales necesarios para el análisis, diseño, implementación e implantación del datawarehouse en materia electoral, basado en las características específicas de los procesos y procedimientos involucrados en el sistema electoral guatemalteco. Se establece el marco contextual necesario para que pueda ser instaurado haciendo uso de cualquier tecnología según los parámetros generales establecidos.

3.2. Arquitectura

Dada la característica que la carga de información al sistema, solo se verá afectada cada cuatro años y la eficiencia de los RDBMS para la implantación de sistemas OLAP, se propone el establecimiento de la base de datos a través de ROLAP con la siguiente infraestructura.

Figura 20. Arquitectura del Datawarehouse Electoral



Se define una arquitectura centralizada donde se aglutinan todos los aspectos del sistema relacionadas y almacenadas en una misma plataforma que pueda ser accedida para la consulta de la información y la realización de análisis. Este diseño integrado proporciona eficiencia de accesos, mayor capacidad de procesamiento y menores costos de soporte.

El servidor tiene tres funciones principales:

- Proveer el entorno de base de datos necesario para la implementación del cubo de información, donde serán almacenados los datos.
- Establecer el servidor de herramienta de BI que se encargará de atender las peticiones de los usuarios, procesar la información y devolver los valores de reportes y totales necesarios para la realización de análisis.
- Realizar el proceso de extracción, transformación y carga de los datos, según los parámetros y períodos establecidos.

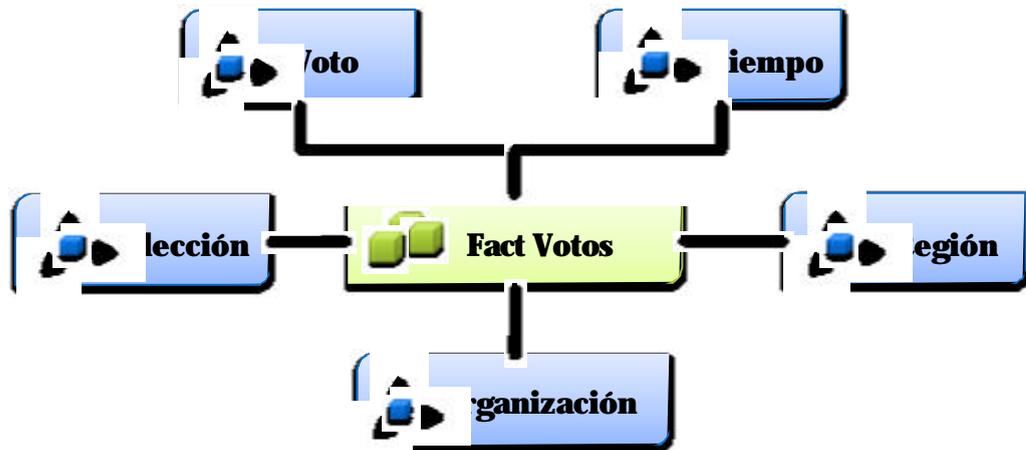
3.3. Modelo de datos

El modelo de datos OLAP está constituido por las dimensiones, jerarquías y tablas de hechos necesarias que permiten la implementación correcta del datawarehouse electoral y que permite la realización de análisis para la construcción de investigaciones socio-políticas.

3.3.1. Descripción general

Se define un datamart que recoge la información electoral correspondiente al desglose de los votos, con una tabla de hechos que recoge la información según características y taxonomías heterogéneas.

Figura 21. Modelo General de Datamart Electoral



A continuación se describe de forma general cada uno de los elementos que componen el datamart a manera de explicar la función de cada uno de estos dentro de la construcción y utilización del modelo estrella.

3.3.2. Dimensiones

Las dimensiones de un modelo OLAP permiten establecer los elementos que van a ser tomados en cuenta para la realización de análisis y no son más que atributos que permiten clasificar y caracterizar las variables, permitiendo el establecimiento de perspectivas seccionadas de análisis. A continuación se detallan los atributos y jerarquías correspondientes a cada una de las dimensiones que constituyen el modelo.

3.3.2.1. Voto

Esta dimensión contiene la información sobre los distintos tipos de votos de los que se lleva información en el datawarehouse, permite clasificar los votos computados durante el conteo de un plebiscito.

Figura 22. Atributos de la dimensión voto



Donde:

- **Voto Sk:** Representa la llave primaria de la dimensión y para su implementación es necesario establecerlo como un tipo de dato autoincremental o una secuencia, dependiendo del gestor de base de datos que se utilice para la implementación del DW.

- **Tipo Voto:** Representa los diferentes votos que pueden contabilizarse durante un conteo.

Tabla IV. Entidades de la dimensión voto

Voto Sk	Tipo Voto
101	Blanco
102	Nulo
103	Válido
104	Emitido

3.3.2.2. Región

Dimensión que establece la estructura geográfica donde se realizan las elecciones, permitiendo la realización de análisis territorial de la emisión de votos identificando de este modo comportamientos regionales.

Figura 23. Atributos de la dimensión región



Donde:

- **Región Sk:** Llave primaria de la dimensión región y que permite llevar control de cada tupla que compone la dimensión.
- **Id Región:** Identificador de cada región en la que se encuentra dividido políticamente el país.
- **Región:** Nombre descriptivo de la región.
- **Id Departamento:** Llave que permite identificar cada departamento del país.
- **Departamento:** Nombre del departamento.
- **Id Municipio:** Código del municipio.
- **Municipio:** Nombre del municipio.

Tabla V. Entidades de la dimensión región

Id Región	Región	Id Depto	Depto	Id Mun	Municipio
101	Región I	101	Guatemala	101	Guatemala
101	Región I	101	Guatemala	102	Villa Nueva
101	Región I	101	Guatemala	103	Mixco

Figura 24. Jerarquía de la dimensión región



A través de esta jerarquía se define la relación existente entre las regiones, departamentos y municipios de Guatemala, donde una Región está compuesta por uno o más Departamentos y a su vez los Departamentos están compuestos por uno o más Municipios.

3.3.2.3. Elección

Esta dimensión permite clasificar los diferentes procesos electorales recogiendo la información referente a los tipos de elecciones que pueden realizarse dentro del sistema electoral.

Figura 25. Atributos de la dimensión elección



Tabla VI. Entidades de la dimensión elección

Id Elección	Elección
101	Presidencial 1ra Vuelta
101	Presidencial 2da Vuelta
102	Diputaciones

3.3.2.4. Organización

Dimensión que mantiene la información sobre las diferentes organizaciones partidarias que forman parte del sistema electoral y han participado en una elección.

Figura 26. Atributos de la dimensión organización



Tabla VII Entidades de la dimensión organización

Id Tipo	Tipo	Id Organización	Siglas
101	Comité Cívico	101	PUA-MEC-FUN
102	Partido Político	101	URNG
102	Partido Político	102	PAN

Figura 27. Jerarquía de la dimensión organización



Esta jerarquía permite establecer la taxonomía de las organizaciones, en base a los tipos de entidades que pueden participar en un proceso electoral.

3.3.2.5. Tiempo

La dimensión tiempo, debido a la periodicidad y temporalidad de los procesos electorales celebrados en el sistema democrático guatemalteco tiene como principal característica llevar control de los años donde se han llevado a cabo procesos electorales (1994, 1999,...., 2007).

Figura 28. Atributos de la dimensión tiempo



Tabla VIII. Entidades de la dimensión tiempo

Id Tiempo	Año
101	1985
102	1990
103	1999

3.1.3. Tabla de hechos

El modelo estrella definido como sistema de datawarehouse electoral, cuenta con una tabla de hechos que permite separar la información según su nivel de granularidad, a continuación se detallan las características de la misma.

3.1.3.1. Fact Votos

Esta tabla de hechos contiene la información sobre los votos obtenidos por las diferentes organizaciones políticas que participan en las elecciones. La métrica número votos es la que lleva el control sobre los datos que se recopilaron y que se encuentran cargados al sistema. Las relaciones existentes entre las dimensiones y la tabla de hechos se representan a través de las llaves Sk que constituyen la llave primaria de cada dimensión, por medio de una llave foránea dentro de la tabla de hechos.

Figura 29. Atributos de la tabla de hechos detalle



3.1.4. Mapeo conceptual

La implantación del datawarehouse en un entorno ROLAP requiere de la creación del esquema físico a través de tablas que representen las dimensiones y tabla de hechos que permita el almacenamiento de los datos necesarios para la realización del análisis multidimensional. A continuación se define el entorno físico de forma general teniendo que modificarse el script de creación para que sea soportado según los tipos de datos y características del DBMS sobre el que se vaya a implementar el cubo electoral.

Figura 30. Modelo estrella

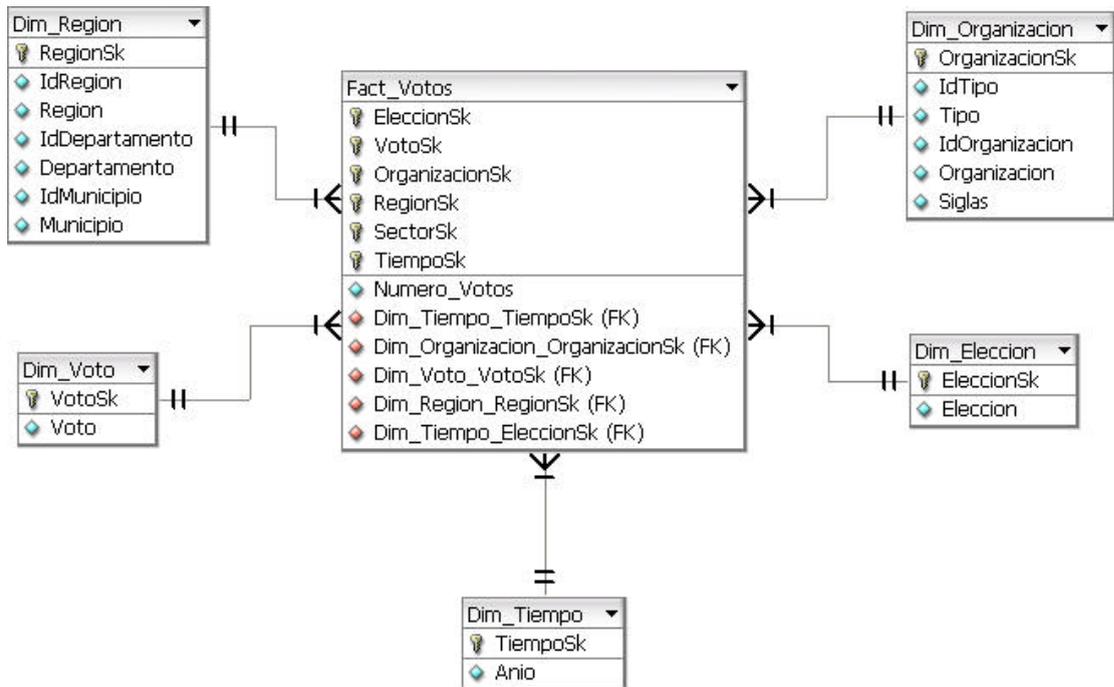


Tabla IX. Script de creación

Tipo	Script
Dimensiones	<pre>CREATE TABLE Dim_Eleccion (EleccionSk NUMERIC NOT NULL, Eleccion VARCHAR NULL, PRIMARY KEY(EleccionSk));</pre>
	<pre>CREATE TABLE Dim_Organizacion (OrganizacionSk NUMERIC NOT NULL, IdTipo NUMERIC, Tipo VARCHAR, IdOrganizacion NUMERIC, Organizacion VARCHAR, Siglas VARCHAR NULL, PRIMARY KEY(OrganizacionSk));</pre>
	<pre>CREATE TABLE Dim_Region (RegionSk NUMERIC NOT NULL, IdRegion NUMERIC, Region VARCHAR, IdDepartamento NUMERIC, Departamento VARCHAR, IdMunicipio NUMERIC, Municipio VARCHAR, PRIMARY KEY(RegionSk));</pre>
	<pre>CREATE TABLE Dim_Tiempo (TiempoSk NUMERIC NOT NULL, Anio VARCHAR, PRIMARY KEY(TiempoSk));</pre>
	<pre>CREATE TABLE Dim_Voto (VotoSk NUMERIC NOT NULL, Voto VARCHAR, PRIMARY KEY(VotoSk));</pre>

Tabla de hechos

```
CREATE TABLE Fact_Votos (  
    EleccionSk NUMERIC NOT NULL,  
    VotoSk NUMERIC NOT NULL,  
    OrganizacionSk NUMERIC NOT NULL,  
    RegionSk NUMERIC NOT NULL,  
    TiempoSk NUMERIC NOT NULL,  
    Numero_Votos NUMERIC NOT NULL,  
    PRIMARY KEY(EleccionSk, VotoSk,  
                OrganizacionSk, RegionSk));
```


4. HERRAMIENTAS DE BI

4.1. Introducción

Actualmente se encuentran disponibles en el mercado gran variedad de herramientas de BI con características técnicas y de instalación particulares, que se constituyen en opciones integrales de acceso, formato y manipulación de información para análisis. La selección de la herramienta de implantación correcta requiere la medición de las características, ventajas y desventajas de las opciones que mejor se adapten a las necesidades de la organización, que provean el entorno ideal de construcción y utilización del modelo subyacente de datawarehouse.

El presente capítulo provee una idea general sobre las herramientas de BI existentes para el año 2008, así como una interpretación de las ventajas competitivas de cada una de ellas y la selección del software de implementación del modelo electoral planteado. El proceso de elección fue llevado a cabo a través de consulta de fuentes y utilización de metodologías aplicadas en proyectos de IT exitosos, tomando como referencia herramientas de alta vigencia tecnológica.

4.2. Características funcionales

Las plataformas de BI tienen como principal función proveer a los usuarios el ambiente para la construcción de aplicaciones que permitan la interpretación de las características de las organizaciones, como elemento de soporte para la toma de decisiones estratégicas de crecimiento y mejoramiento. Para que estas soluciones de software puedan ser consideradas como alternativas viables de construcción del ambiente de inteligencia de negocios del Datawarehouse electoral, deben cumplir al menos con las siguientes características:

Tabla X. Elementos Funcionales de las Plataformas de BI

Característica	Elementos
Integración	<ul style="list-style-type: none"> • Infraestructura de BI: Todas las herramientas deben compartir elementos tales como seguridad, roles, metadatos, formas de administración, modelado de objetos, motores de consultas y mantener un estándar del entorno visual. • Soporte de Metadatos OLAP: Proveer de un entorno sólido de consulta, captura, almacenamiento y publicación de metadatos tales como dimensiones, métricas, jerarquías. • Ambiente de Desarrollo: Proporcionar el conjunto de herramientas de desarrollo que permitan su integración a los procesos del negocio, generación de interfaces de intercambio con otras aplicaciones, construcción de componentes a través de asistentes gráficos, así como el soporte de comunicación a través de web services que permitan el intercambio con entidades externas. • Colaboración y Flujo de Trabajo: Habilitar a los usuarios espacios de intercambio tales como carpetas compartidas y foros de discusión, proveer los elementos para seguimiento de tareas o eventos así como la integración con herramientas de control de flujo de trabajo.

<p>Presentación de la información</p>	<ul style="list-style-type: none"> • Reportes: Contar con un portafolio variado de formatos y estilos de reportes prediseñados e interactivos que abarquen distintas áreas organizacionales con capacidad de intercambio en diversos formatos, así como capacidad de manejo de escalabilidad de los datos. • Consultas Ad-Hoc: Dotar al usuario de los medios para la creación de consultas personalizadas y específicas de los datos, sin necesidad de depender de personal de IT para la creación de reportes facilitando la navegación a través de las fuentes de datos. • Integración con Office: Capacidad de integrarse con herramientas de Office como Microsoft Excel a través de exportaciones a distintos formatos, creación de cuadros pivotes, actualización de las hojas de cálculo, formulación así como el manejo de bloqueo de escritura de celdas.
<p>Análisis</p>	<ul style="list-style-type: none"> • OLAP: Proveen a los usuarios la capacidad de realizar consultas y cálculos de forma rápida abarcando gran variedad de arquitecturas de almacenamiento tales como relacional, multidimensional y en memoria. • Visualización Avanzada: Capacidad de extender los reportes a formas visuales más atractivas y complejas que filas y columnas de información, tales como gráficas interactivas que provean una mejor representación de los datos para que de este modo la información pueda ser comprendida más a fondo y fácilmente.

4.3. Selección de la herramienta

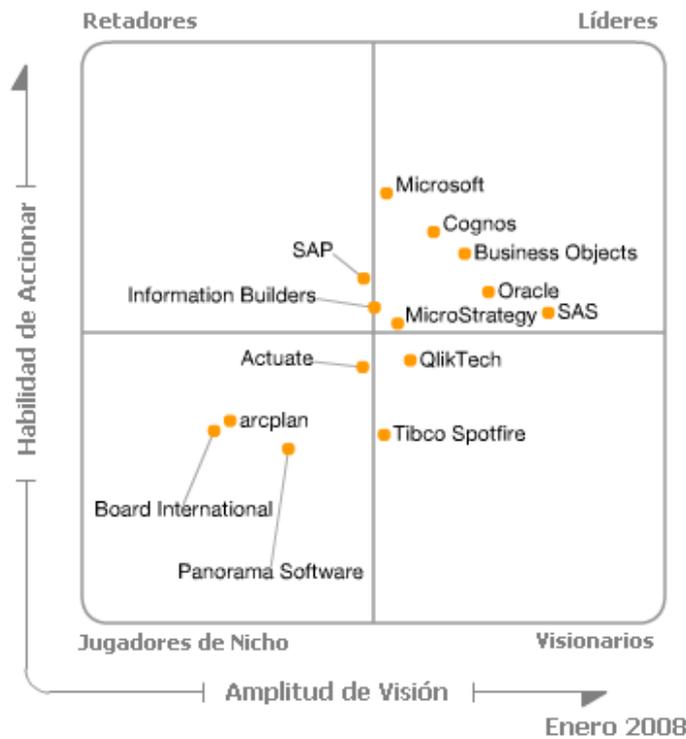
4.3.1. Cuadrante mágico

La empresa Gartner es actualmente un punto de referencia importante utilizado por los gestores de tecnologías de la información, previo a la selección o implementación de proyectos. Proveen estudios de benchmarking de alta calidad para herramientas de IT desde CRM, ERP, BI hasta VoIP que son utilizados para la comparación y evaluación de herramientas.

El cuadrante mágico de plataformas de BI de Gartner, es establecido como una herramienta analítica representativa del estado actual del mercado de proveedores de plataformas de BI, considerados como opciones viables para la implementación de soluciones de inteligencia de negocios y puede servir como guía para la selección de una herramienta específica.

La siguiente gráfica muestra la distribución actual de los diferentes participantes del mercado a nivel mundial:

Figura 31. Cuadrante Mágico de Plataformas de BI



Fuente: Gartner Research (Enero 2008)

Este estudio sólo toma en cuenta aquellas empresas que tienen presencia en el mercado de BI y que tienen ganancias superiores a los \$20 millones anuales. La evaluación de cada proveedor se basa en criterios como: competitividad, continuidad de inversión, evolución y mejoramiento de las herramientas, evolución con las necesidades del cliente, experiencia de los clientes, compromiso a largo plazo, conocimiento de las necesidades del mercado, diversidad de aplicación, estrategias geográficas, entre otros.

4.3.2. Metodología

Para la selección de la herramienta de BI que mejor se ajuste a las necesidades del proyecto de Datawarehouse Electoral, se utilizó la metodología TOC por sus siglas en inglés Total Cost of Ownership (Costo Total de Propiedad). Desarrollada en los años 70's por Hill Kirwin de Gartner Group, plantea que se debe *“conocer su misión y utilizar TCO para priorizar, orientar y evaluar los progresos hacia el cumplimiento de los objetivos de la organización”*. Esta metodología plantea la selección de la herramienta o producto de software que mejor se adapte a las necesidades, basado en la identificación de criterios, asignación de ponderaciones y calificación de las alternativas.

El proceso de evaluación fue llevado a cabo en conjunto con el comité evaluador nombrado por la organización, integrado por el director, el gerente de informática y el jefe del área de investigación. Primero se seleccionaron tres herramientas del mercado basado en el cuadrante mágico de BI de Gartner para enero del 2008, siendo estas: Cognos de IBM, Proclarity de Microsoft e Hyperion de Oracle. Luego se establecieron los criterios y la ponderación de cada uno de estos tomando en cuenta aspectos técnicos, elementos de uso e instalación y la tecnología IT existente en la organización. Como último paso se instalaron y configuraron cada una de las herramientas seleccionadas en ambientes de prueba y fueron calificadas por el comité evaluador para obtener de este modo el resultado final de la herramienta seleccionada.

A continuación se detallan los criterios establecidos para la evaluación de las herramientas, donde la escala de evaluación está definida entre 1 y 10, siendo 10 el valor más satisfactorio:

Tabla XI Criterios de evaluación de primer nivel

Criterio	Porcentaje
Reportes	25%
Análisis	25%
Datawarehousing	25%
Integración	15%
Otros Factores	10%
Total	100%

Tabla XII Criterios de evaluación de segundo nivel

Criterio	Porcentaje
Integración - 25%	
Infraestructura de BI	45%
Soporte de Metadatos OLAP	35%
Ambiente de Desarrollo	10%
Colaboración y Flujo de Trabajo	10%
Presentación de la Información - 25%	
Reportes	50%
Consultas Ad-hoc	20%
Integración con Microsoft Office	30%
Análisis - 25%	
OLAP	20%
Visualización	30%
Análisis Predictivo	30%
Administración	20%

Soporte - 15%	
Garantía	35%
Centro de Servicio	35%
Capacitación	30%
Otros Factores - 10%	
Simplicidad de Instalación-Configuración	30%
Curva de Aprendizaje	30%
Facilidad de ETL	30%
Posición en Cuadrante Mágico	10%

4.3.3. Resultados

Aplicando el modelo de TCO se tiene primeramente los resultados para los criterios de segundo nivel:

Tabla XIII Resultados obtenidos por herramienta en criterios de integración

Integración	Porcentaje	Cognos		Hyperion		Proclarity	
Infraestructura de BI	45%	9	41%	9	41%	9	41%
Soporte de Metadatos OLAP	35%	10	35%	10	35%	10	35%
Ambiente de Desarrollo	10%	10	10%	10	10%	8	8%
Colaboración y Flujo de Trabajo	10%	9	9%	9	9%	6	6%
	100%		95%		95%		90%

Tabla XIV. Resultados de criterios de presentación de la información

Presentación de la Información	Porcentaje	Cognos		Hyperion		Proclarity	
Reportes	50%	9	45%	9	45%	9	45%
Consultas Ad-hoc	30%	10	30%	9	27%	8	24%
Integración con Microsoft Office	20%	9	18%	8	16%	8	16%
	100%		93%		88%		85%

Tabla XV. Resultados para criterios de análisis

Análisis	Porcentaje	Cognos		Hyperion		Proclarity	
OLAP	20%	10	20%	10	20%	10	20%
Visualización	30%	9	27%	9	27%	9	27%
Análisis Predictivo	30%	9	27%	8	24%	7	21%
Administración	20%	10	20%	9	18%	10	20%
	100%		94%		89%		88%

Tabla XVI Resultados para criterios de soporte

Soporte	Porcentaje	Cognos		Hyperion		Proclarity	
Garantía	35%	10	35%	9	32%	9	32%
Centro de Servicio	35%	10	35%	9	32%	9	32%
Capacitación	30%	8	24%	8	24%	9	27%
	100%		94%		87%		90%

Tabla XVII Resultados para criterios de otros factores

Otros Factores	Porcentaje	Cognos		Hyperion		Proclarity	
Simplicidad de Instalación-Configuración	30%	7	21%	7	21%	9	27%
Curva de Aprendizaje	30%	8	24%	9	27%	10	30%
Facilidad de ETL	30%	9	27%	8	24%	9	27%
Posición en Cuadrante Mágico	10%	9	9%	8	8%	10	10%
	100%		81%		80%		94%

Luego de haber evaluado en cada uno de los criterios a las diferentes herramientas, los valores totales sobre 100% obtenidos en cada categoría se trasladan a los criterios de primer nivel, donde fueron obtenidos los siguientes resultados:

Tabla XVIII. Valores obtenidos para criterios de primer nivel

Criterios de Primer Nivel	Porcentaje	Cognos		Hyperion		Proclarity	
Integración	25%	95	24%	95	24%	90	22%
Presentación de la Información	25%	93	23%	88	22%	85	21%
Análisis	25%	94	24%	89	22%	88	22%
Soporte	15%	94	14%	87	13%	90	14%
Otros Factores	10%	81	8%	80	8%	94	9%
	100%		69%		65%		66%

Por último, se valoran con base a los costos promedios de implementación y licenciamiento de cada una de las herramientas, obteniéndose el siguiente resultado:

Tabla XIX. Resultados de TOC

	Costo Promedio*	Resultado	TOC
Cognos	USD \$75,000.00	69	1087.745
Hyperion	USD \$60,000.00	65	918.836
Proclarity	USD \$40,000.00	66	604.686

* Basado en cotizaciones obtenidas de partners en Guatemala

Tomando como base los resultados obtenidos, se tomó la decisión de utilizar la herramienta Proclarity como alternativa de implementación del datawarehouse electoral.

4.4. Entorno

Luego de la selección de la plataforma para la implementación e implantación del sistema de Datawarehouse Electoral es necesario instalar y configurar el conjunto de herramientas Microsoft que permitan establecer el entorno donde será construido y desarrollado el sistema de Datawarehouse electoral. Queda como responsabilidad de la institución la adquisición de la licencia de Proclarity para la puesta en producción del sistema a partir de los parámetros definidos, que le permitan la utilización del mismo. Las herramientas seleccionadas para la implementación del sistema permiten la obtención, almacenamiento, manipulación, transformación y consulta de la información.

Tabla XX. Tabla descriptiva de herramientas BI de la solución

Herramienta	Descripción
Sql Server 2005 	Provee el entorno necesario para el almacenamiento de la información, tanto para el entorno OLAP como OLTP del sistema.
Visual Studio 2005 (BI) 	Define el conjunto de elementos necesarios para la implantación de las dimensiones y el cubo de BI que permitan la consulta de los datos electorales.
Proclarity 	Provee la interfaz de usuarios para la realización de reportes que permitan la realización de análisis sobre los datos electorales.

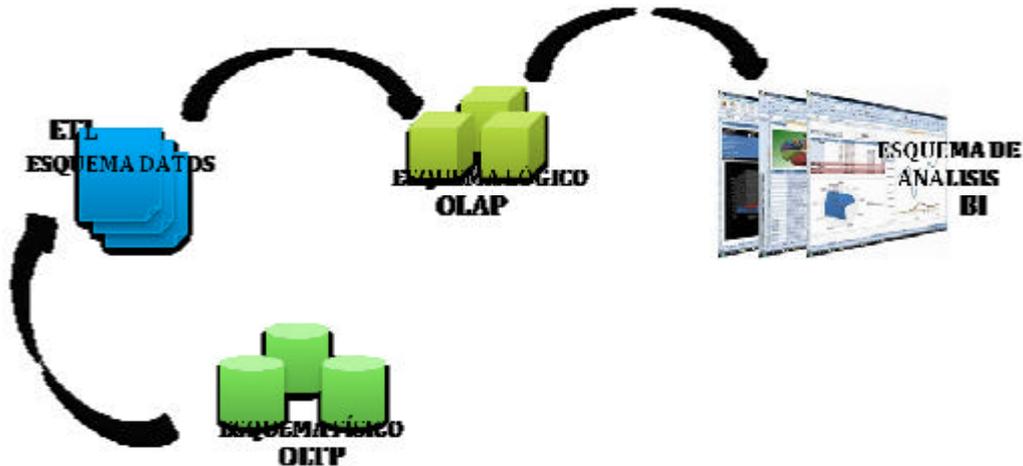
5. IMPLEMENTACIÓN

La correcta implementación del sistema de Datawarehouse Electoral requiere de la consecución de un conjunto de pasos organizados que conlleven a la definición del entorno de BI que permita la generación de información de análisis electoral. El presente capítulo define los parámetros y estándares que deben adoptarse para establecer el ambiente de producción recomendado para la implantación del sistema electoral orientado a la tecnología seleccionada, así como ejemplos de utilización de la misma para la generación de reportes para interpretación.

5.1. Metodología

El procedimiento de implementación de Datawarehouse Electoral requiere de cuatro pasos fundamentales que se entrelazan y se retroalimentan para construir y liberar los elementos necesarios para el correcto funcionamiento del sistema, la siguiente figura muestra el diagrama del proceso y las relaciones entre los elementos.

Figura 32. Proceso de Implementación de Datawarehouse Electoral



Cada área o elemento cuenta con objetivos específicos y tareas establecidas que permiten la correcta integración y comunicación entre ellos.

- **Esquema físico:** Mantener y administrar el entorno de almacenamiento consistente de los datos a través de la creación de la base de datos (script) que contenga el esquema estrella de dimensiones y tabla de hechos definido como base OLTP del datawarehouse.
- **Esquema de datos:** Establecer los parámetros de carga de la información recopilada así como la definición de los estándares que deben tener los datos para que puedan ser utilizados para alimentar el sistema.
- **Esquema lógico:** Organizar el marco lógico OLAP de las estructuras multidimensionales que permitan el establecimiento de las relaciones, jerarquías y dependencias necesarias para que el sistema pueda ser utilizado para la realización de Business Intelligence.
- **Esquema de análisis:** Definir las vistas de información necesarias para los usuarios finales, así como proveer el entorno de manipulación e interpretación de datos a través de gráficos y consolidaciones.

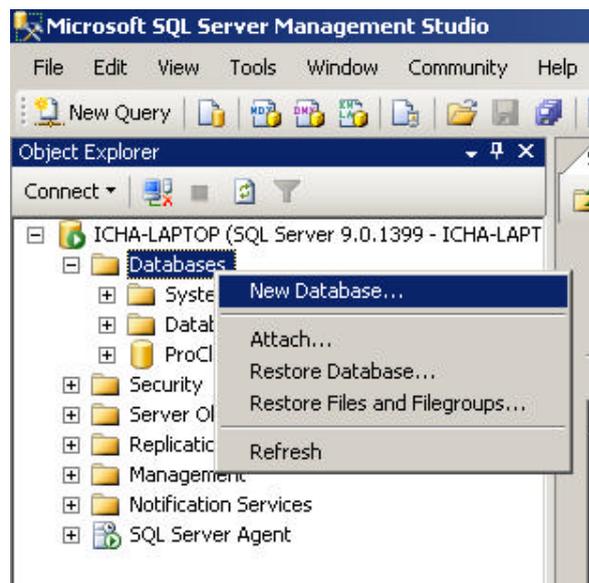
5.1.1. Esquema físico

La creación del esquema de base de datos requiere la utilización de SQL Server 2005 y deberán seguirse los siguientes pasos:

1) Creación de la nueva BDD

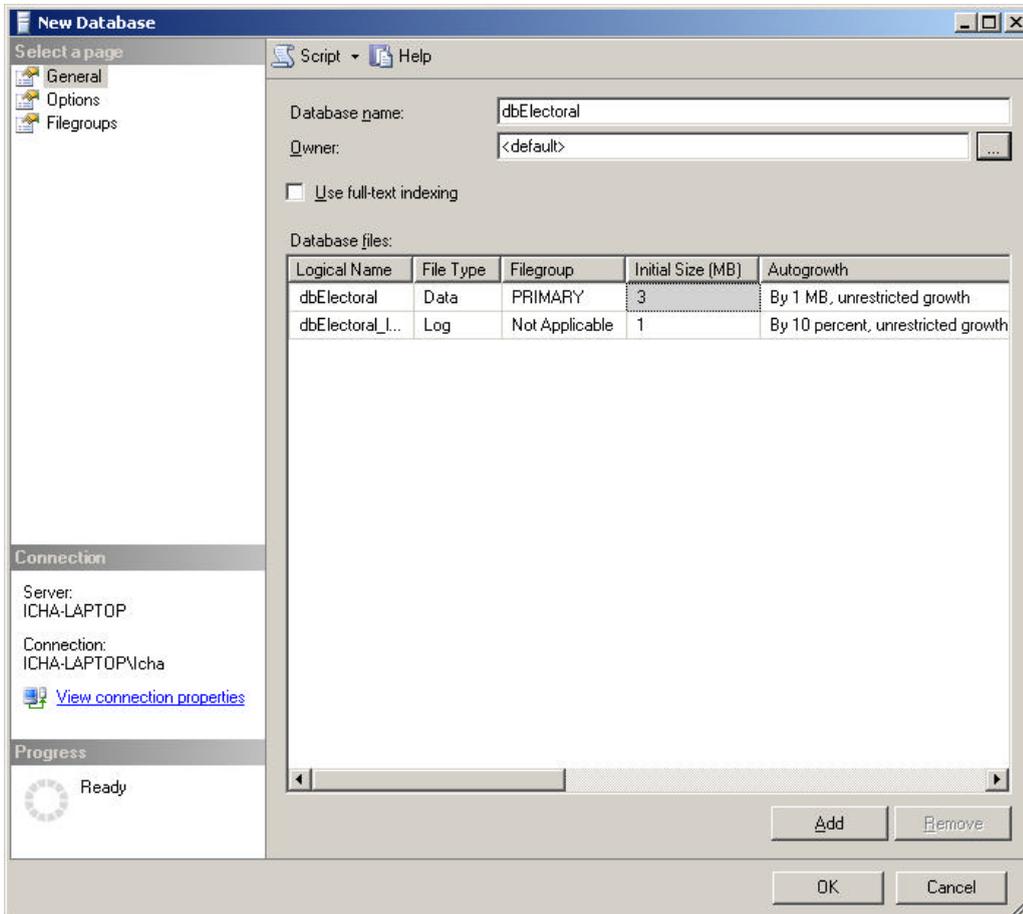
Luego de entrar a SQL Server y seleccionar la pestaña de Databases, se presiona clic derecho sobre ella y se ejecuta la opción New Database, como se muestra en la imagen.

Figura 33. Crear nueva base de datos



Esto mostrará las diferentes opciones de configuración de la BDD, en este caso solo deberá colocarse el nombre que se le quiera dar al entorno (dbElectoral), dejando el resto de parámetros con los parámetros predeterminados.

Figura 34. Definición de estructura de la base de datos



2) Ejecución del script de creación

Luego de la creación de la base de datos, se deberá ejecutar el script SQL que permite establecer el esquema OLTP que permitirá la definición de las dimensiones y la tabla de hechos. Este script se ha modificado para soportar los tipos de datos y características de SQL Server 2005:

```

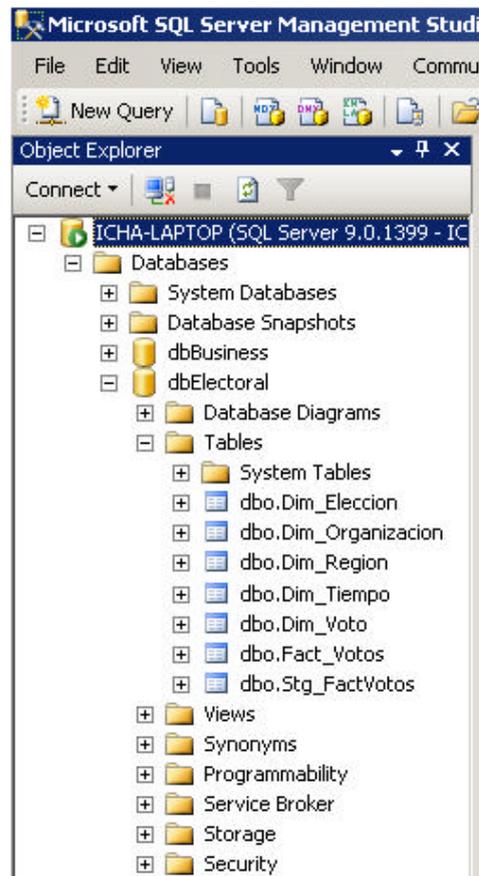
CREATE TABLE Dim_Eleccion (
    EleccionSk NUMERIC IDENTITY(100,1),
    Eleccion VARCHAR(50) NULL,
    PRIMARY KEY(EleccionSk));
CREATE TABLE Dim_Organizacion (
    OrganizacionSk NUMERIC IDENTITY(1,1),
    IdTipo INTEGER NULL,
    Tipo VARCHAR(50) NULL,
    IdOrganizacion INTEGER NULL,
    Organizacion VARCHAR(100) NULL,
    Siglas VARCHAR(50) NULL,
    PRIMARY KEY(OrganizacionSk));
CREATE TABLE Dim_Region (
    RegionSk NUMERIC IDENTITY(1,1),
    IdRegion INTEGER NULL,
    Region VARCHAR(50) NULL,
    IdDepartamento INTEGER NULL,
    Departamento VARCHAR(100) NULL,
    IdMunicipio INTEGER NULL,
    Municipio VARCHAR(100) NULL,
    PRIMARY KEY(RegionSk));
CREATE TABLE Dim_Voto (
    VotoSk NUMERIC IDENTITY(100,1),
    Voto VARCHAR(50) NULL,
    PRIMARY KEY(VotoSk));
CREATE TABLE Dim_Tiempo (
    TiempoSk NUMERIC IDENTITY(100,1),
    Anio VARCHAR(50) NULL,
    PRIMARY KEY(TiempoSk));
GO;

```

```
CREATE TABLE Fact_Votos (  
    EleccionSk NUMERIC NOT NULL,  
    VotoSk NUMERIC NOT NULL,  
    OrganizacionSk NUMERIC NOT NULL,  
    RegionSk NUMERIC NOT NULL,  
    TiempoSk NUMERIC NOT NULL,  
    Numero_Votos NUMERIC NOT NULL,  
    PRIMARY KEY(EleccionSk, VotoSk, OrganizacionSk,  
    RegionSk,TiempoSk));
```

Al finalizar la ejecución del script se pueden observar las tablas creadas en la base de datos “dbElectoral” que soporta el proceso ROLAP del Datawarehouse.

Figura 35. Esquema de tablas de base de datos dbElectoral



5.1.2. Esquema de datos

Esta etapa tiene como principal función definir los parámetros de datos que se cargarán en las dimensiones y tabla de hechos, para lo cual es necesario realizar una carga inicial y establecer los elementos generales para la realización de cargas posteriores. Los datos recopilados de los diferentes procesos electorales se guardaron en archivos de Excel con formato establecido que se utilizan para la carga inicial de los datos del cubo electoral.

- La dimensión tiempo se encuentra reflejada a través del nombre del archivo, el cual indica los diferentes años de procesos electorales realizados en Guatemala.
- La dimensión elección está representada en los nombres de las diferentes hojas que componen el Excel.

Figura 38. Datos de dimensión elección

General	Presidenciales 1ra Vuelta	Presidenciales 2da Vuelta	Diputaciones	Municipales
---------	---------------------------	---------------------------	--------------	-------------

- Los datos de la dimensión organización están reflejados en las columnas de composición de votos donde se identifican los Partidos Políticos y Comités Cívicos por medio de las siglas.

Figura 39. Valores de la dimensión organización

Composición del Voto por Organización Política							
DCG	UCN	PSD	PNR	CAN	MLN-PID	PDCN-PR	PUA-MEC-FUN

- En las columnas de Región, Departamento y Municipio se identifican los valores de la dimensión región donde se ingresan los valores de la regionalización establecida por el TSE para los diferentes procesos electorales.

Figura 40. Parámetros de la dimensión región

Región	Depto.	N°	Municipio
Región I	Guatemala	1	Guatemala
		2	Santa Catarina Pinula
		3	San José Pinula
		4	San José del Golfo
		5	Palencia
		6	Chinautla
		7	San Pedro Ajampuc
		8	Mixco
		9	San Pedro Sacatepéquez
		10	San Juan Sacatepéquez
		11	San Raymundo
		12	Chuarancho
		13	Frailanes
		14	Amatitlan
		15	Villa Nueva
		16	Villa Canales
		17	San Miguel Petapa

- La clasificación de votos del archivo permite identificar los datos de la dimensión votos.

Figura 41. Datos de la dimensión votos

Composición del Voto			
Valido	Nulo	Blanco	Emitido

Estos datos deben ser transformados en valores que puedan ser cargados en las dimensiones que componen el modelo OLAP que permitan la carga de información a la tabla de hechos.

5.1.2.1. ETL de datos

El establecimiento de las cargas tanto inicial como incremental, así como las diferentes transformaciones que son necesarias realizar sobre los datos para que se acoplen de forma correcta al modelo ROLAP y que cumpla con los requerimientos y estándares de procesos electorales, es un paso importante que se definen a continuación:

5.1.2.1.1. Carga inicial

- **Dimensión voto:** Se cargan los valores establecidos para los diferentes tipos de votos que se van a manejar.

Figura 42. Datos cargados en la dimensión voto

Table - dbo.Dim_Voto	
VotoSk	Voto
100	Emitido
101	Blanco
102	Nulo
103	Valido
104	General

Los datos de esta dimensión se establecen con base a los tipos de votos permitidos por el Tribunal Supremo Electoral durante la realización del conteo, en el caso del valor “General” se utiliza para todos aquellos valores que no necesariamente representan sufragio electoral tales como, población empadronada o población analfabeta, así como otros datos estadísticos que puedan ser de interés.

- **Dimensión región:** Los datos de esta dimensión reflejan la jerarquía geográfica de participación electoral, para esto se realizaron inserciones de los datos a través de scripts, basado en información de la división política establecida por el TSE de las diferentes regiones, departamentos y municipios de Guatemala.

```
INSERT INTO
DIM_REGION (IdRegion,Region,IdDepartamento,Departamento,IdMunicipio,Municipio)
VALUES (1,'Región I',1,'Guatemala',1,'Guatemala')
INSERT INTO
DIM_REGION (IdRegion,Region,IdDepartamento,Departamento,IdMunicipio,Municipio)
VALUES (1,'Región I',1,'Guatemala',2,'Santa Catarina Pinula')
INSERT INTO
DIM_REGION (IdRegion,Region,IdDepartamento,Departamento,IdMunicipio,Municipio)
VALUES (1,'Región I',1,'Guatemala',3,'San José Pinula')
INSERT INTO
DIM_REGION (IdRegion,Region,IdDepartamento,Departamento,IdMunicipio,Municipio)
VALUES (1,'Región I',1,'Guatemala',4,'San José del Golfo')
```

Figura 43. Datos cargados en la dimensión región

Table - dbo.Dim_Region							
	RegionSk	IdRegion	Region	IdDepartamento	Departamento	IdMunicipio	Municipio
▶	1	1	Región I	1	Guatemala	1	Guatemala
	2	1	Región I	1	Guatemala	2	Santa Catarina ...
	3	1	Región I	1	Guatemala	3	San José Pinula
	4	1	Región I	1	Guatemala	4	San José del Golfo
	5	1	Región I	1	Guatemala	5	Palencia
	6	1	Región I	1	Guatemala	6	Chinautla
	7	1	Región I	1	Guatemala	7	San Pedro Ayam...
	8	1	Región I	1	Guatemala	8	Mixco
	9	1	Región I	1	Guatemala	9	San Pedro Sacat...
	10	1	Región I	1	Guatemala	10	San Juan Sacate...
	11	1	Región I	1	Guatemala	11	San Raymundo
	12	1	Región I	1	Guatemala	12	Chuarancho
	13	1	Región I	1	Guatemala	13	Fraijanes

- **Dimensión elección:** Se ingresan los diferentes tipos de procesos electorales o entidades de elección popular sobre los que se desea llevar información, estos pueden ser diputaciones, consultas populares, referéndums entre otros, así como el caso de las elecciones presidenciales se identifican la 1ra y la 2da vuelta como elementos diferentes dentro de la dimensión.

Figura 44. Valores cargados a la dimensión elección

Table - dbo.Dim_Eleccion		
	EleccionSk	Eleccion
	100	Presidencial 1ra Vuelta
	101	Presidencial 2da Vuelta
	102	Diputación Nacional
	103	Diputación Distrital
	104	Municipales
	105	Generales

- **Dimensión organización:** Esta dimensión contiene información sobre las distintas organizaciones políticas que han participado en procesos electorales, cada nueva organización política de los posteriores procesos electorales deberá ser registrada. En el caso de los comités cívicos, debido a que varían frecuentemente tanto en número como en nombres durante cada período electoral y dada la característica de que al cierre de cada período electoral dejan de tener vigencia para las autoridades electorales, se utilizarán nombres genéricos como CC1 (Comité Cívico 1), CC2 (Comité Cívico 2). El ingreso de los valores a esta dimensión se hizo a través de un script de inserción que agrega cada Partido Político a la dimensión.

```

INSERT INTO DIM_ORGANIZACION VALUES(1,'Partido Politico',1,'Alianza Democrática','AD')
INSERT INTO DIM_ORGANIZACION VALUES(1,'Partido Politico',2,'Coalición AD-URNG','AD-URNG')
INSERT INTO DIM_ORGANIZACION VALUES(1,'Partido Politico',3,'Alianza Nueva Nación','ANN')
INSERT INTO DIM_ORGANIZACION VALUES(1,'Partido Politico',4,'Alianza Popular 5','AP5')

```

Figura 45. Valores cargados a la dimensión organización

Table - dbo.Dim_Organizacion						
	OrganizacionSk	IdTipo	Tipo	IdOrganizacion	Organizacion	Siglas
	1	1	Partido Politico	1	Alianza Democrá...	AD
	2	1	Partido Politico	2	Coalición AD-URNG	AD-URNG
	3	1	Partido Politico	3	Alianza Nueva N...	ANN
	4	1	Partido Politico	4	Alianza Popular 5	AP5
	5	1	Partido Politico	5	Acción Reconcili...	ARDE
	6	1	Partido Politico	6	Alianza Republic...	ARENA
	7	1	Partido Politico	7	Cambio Histórico...	CAMHINA
	8	1	Partido Politico	8	Central Auténtic...	CAN
	9	1	Partido Politico	9	Democracia Crist...	DCG
	10	1	Partido Politico	10	Coalición DCG-F...	DCG-FDNG
	11	1	Partido Politico	11	Coalición DCG-PR	DCG-PR

- **Dimensión tiempo:** Esta dimensión identifica los años en los que se han realizado procesos electorales, en el caso de Guatemala se verá afectada cada 4 años ingresándose de forma manual el correspondiente valor.

Figura 46. Datos establecidos en la dimensión tiempo

Table - dbo.Dim_Tiempo	
TiempoSk	Anio
100	1985
101	1990
102	1995
103	1999
104	2003

- **Tabla de hechos:** Para la carga inicial de la tabla de hechos se creó una tabla llamada Stg_FactVotos que tiene como función principal almacenar los datos previo a ser cargados a la tabla de hechos basados en las llaves de cada una de las dimensiones, convirtiéndose en el punto intermedio entre la información plana y el cubo de información. Esta tabla se crea a través de la ejecución del siguiente script.

```
CREATE TABLE Stg_FactVotos(
    DV_VOTO varchar(50),
    DE_ELECCION varchar(50),
    DR_REGION varchar(50)
    DR_DEPTO varchar(100),
    DR_MUNI varchar(100),
    DO_TIPO varchar(50),
    DO_SIGLAS varchar(50),
    DT_ANIO numeric(18, 0)
    VALOR numeric(18, 0))
```

Mediante este paso se convierte los datos en información congruente a los datos dimensionales para poder contar de este modo con información consistente y validar los campos de textos con sus correspondientes SK de cada dimensión. Para esta carga, se crearon scripts de inserción de los datos de cada proceso electoral en la tabla intermedia y luego fueron cargados año de año realizando las transformaciones necesarias en la tabla de hechos.

Figura 47. Carga de datos a la tabla Stg_FactVotos

Table - dbo.Stg_FactVotos									
DV_VOTO	DE_ELECCION	DR_REGION	DR_DEPTO	DR_MUNI	DO_TIPO	DO_SIGLAS	ANIO	VALOR	
General	Generales	Región I	Guatemala	Guatemala	Sin Organización	General	1985	282969	
General	Generales	Región I	Guatemala	Santa Catarina Pinula	Sin Organización	General	1985	5296	
General	Generales	Región I	Guatemala	San José Pinula	Sin Organización	General	1985	4792	
General	Generales	Región I	Guatemala	San José del Golfo	Sin Organización	General	1985	1186	
General	Generales	Región I	Guatemala	Palencia	Sin Organización	General	1985	5800	
General	Generales	Región I	Guatemala	Chinautla	Sin Organización	General	1985	4144	
General	Generales	Región I	Guatemala	San Pedro Ayampuc	Sin Organización	General	1985	3558	
General	Generales	Región I	Guatemala	Mixco	Sin Organización	General	1985	33514	
General	Generales	Región I	Guatemala	San Pedro Sacatepéquez	Sin Organización	General	1985	3626	
General	Generales	Región I	Guatemala	San Juan Sacatepéquez	Sin Organización	General	1985	15337	
General	Generales	Región I	Guatemala	San Raymundo	Sin Organización	General	1985	3420	
General	Generales	Región I	Guatemala	Chuarrañcho	Sin Organización	General	1985	2015	
General	Generales	Región I	Guatemala	Fraijanes	Sin Organización	General	1985	2530	
General	Generales	Región I	Guatemala	Amatitlan	Sin Organización	General	1985	9458	
General	Generales	Región I	Guatemala	Villa Nueva	Sin Organización	General	1985	15895	
General	Generales	Región I	Guatemala	Villa Canales	Sin Organización	General	1985	12093	
General	Generales	Región I	Guatemala	San Miguel Petapa	Sin Organización	General	1985	2868	

Luego de haber ingresado a la tabla Stg_FactVotos la información correspondiente a los votos emitidos durante cada proceso electoral, se deberán cargar a la tabla de hechos, para lo cual es necesario ejecutar el siguiente script, para posteriormente borrar los datos de la Stg_FactVotos para que pueda servir en el proceso de cargas incrementales.

```

INSERT INTO FACT_VOTOS (VOTOSK, ELECCIONSK, REGIONSK, ORGANIZACIONSK,
TIEMPOSK, NUMERO_VOTOS)
SELECT DV.VOTOSK, DE.ELECCIONSK, DR.REGIONSK, DO.ORGANIZACIONSK, DT.TIEMPOSK,
VALOR
FROM STG_FACTVOTOS FV, DIM_VOTO DV, DIM_ELECCION DE, DIM_REGION DR,
DIM_ORGANIZACION DO, DIM_TIEMPO DT
WHERE FV.DV_VOTO=DV.VOTO
AND FV.DE_ELECCION=DE.ELECCION
AND FV.DR_REGION=DR.REGION
AND FV.DR_DEPTO=DR.DEPARTAMENTO
AND FV.DR_MUNI=DR.MUNICIPIO
AND FV.DO_TIPO=DO.TIPO
AND FV.DO_SIGLAS=DO.SIGLAS
AND FV.DT_ANIO=DT.ANIO

```

Este script genera la carga de los valores dimensionales en la tabla de hechos generando las relaciones correspondientes a las distintas llaves Sk.

Figura 48. Datos cargados a la tabla de hechos Fact_Votos

ICHA-LAPTOP.d...QLQuery1.sql*		Table - dbo.Fact_Votos*			Summary	
	EleccionSk	VotoSk	OrganizacionSk	RegionSk	TiempoSk	Numero_Votos
	100	100	99	1	100	409722
	100	100	99	1	101	390312
	100	100	99	1	102	308436
	100	100	99	1	103	398806
	100	100	99	1	104	425009
	100	100	99	2	100	6915
	100	100	99	2	101	6933
	100	100	99	2	102	6674
	100	100	99	2	103	9992
	100	100	99	2	104	18047
	100	100	99	3	100	5594
	100	100	99	3	101	5989
	100	100	99	3	102	5530
	100	100	99	3	103	8087
	100	100	99	3	104	12248
	100	100	99	3	100	1150

5.1.2.1.2. Carga incremental

Dada las características de los procesos electorales en Guatemala y basado en la Ley Electoral y de Partidos Políticos, las dimensiones que conforman el Cubo Electoral no sufren cambios drásticos cada 4 años, manteniendo generalmente sus datos invariables, aunque pueden sufrir un crecimiento basado en la generación de nuevos parámetros.

- **Dimensión voto:** Esta dimensión no varía con cada nuevo proceso electoral, en caso de que se quiera llevar control sobre algún otro tipo de voto admitido por el TSE, se deberá hacer una inserción directa de la tupla con el tipo de voto sugerido.

Figura 49. Ejemplo de nuevo valor en la dimensión voto

Table - dbo.Dim_Voto		
	VotoSk	Voto
	100	Emitido
	101	Blanco
	102	Nulo
	103	Valido
	104	General
	NULL	Nuevo Voto
	NULL	NULL

- **Dimensión región:** En caso de que se cree una nueva región, se agregue un departamento o se cree un nuevo municipio, deberá realizarse la inserción del nuevo valor.

Figura 50. Ejemplo de inserción de nuevo dato en dimensión región

325	4	Región IV	22	Jutiapa	325	Zapotitlán	
326	4	Región IV	22	Jutiapa	326	Comapa	
327	4	Región IV	22	Jutiapa	327	Jalpatagua	
328	4	Región IV	22	Jutiapa	328	Conguaco	
329	4	Región IV	22	Jutiapa	329	Moyuta	
330	4	Región IV	22	Jutiapa	330	Pasaco	
331	4	Región IV	22	Jutiapa	331	San José Acate...	
332	4	Región IV	22	Jutiapa	332	Quezada	
	NULL	5	Nueva Region	23	Nuevo Departamento	333	Nuevo Municipio
*	NULL	NULL	NULL	NULL	NULL	NULL	NULL

- **Dimensión elección:** En esta dimensión se deberán agregar nuevos procesos electorales como serían Consultas Populares o Referéndums, en caso de que se desee llevar información sobre ellos.

Figura 51. Ejemplo de nuevo dato en dimensión elección

Table - dbo.Dim_Eleccion	
EleccionSk	Eleccion
100	Presidencial 1ra ...
101	Presidencial 2da ...
102	Diputación Nacio...
103	Diputación Distrital
104	Municipales
105	Generales
	NULL Consulta Popular
*	NULL

- **Dimensión organización:** Según las características de los procesos electorales en Guatemala, esta sería la dimensión que mayor movimiento sufrirá en cada nueva elección y dependerá del número de nuevos partidos políticos que participen en la contienda electoral, los cuales deberán ser insertados en la dimensión.

Figura 52. Ejemplo de ingreso de nuevo dato en dimensión organización

	90	1	Partido Politico	90	Union Nacionalis...	UNO
	91	1	Partido Politico	91	Unidad Revoluci...	URNG
	92	1	Partido Politico	92	Unión Reformist...	URS
	93	2	Comité Cívico	93	Comité Cívico 1	CC1
	94	2	Comité Cívico	94	Comité Cívico 2	CC2
	95	2	Comité Cívico	95	Comité Cívico 3	CC3
	96	2	Comité Cívico	96	Comité Cívico 4	CC4
	97	2	Comité Cívico	97	Comité Cívico 5	CC5
	98	2	Comité Cívico	98	Comité Cívico 6	CC6
	NULL	1	 Partido Politico 	99	 Partido X 	PX
*	NULL	NULL	NULL	NULL	NULL	NULL

- **Dimensión tiempo:** El mantenimiento de esta dimensión requiere el ingreso del nuevo año electoral que va a ser cargado al sistema.

Figura 53. Ejemplo de nuevo dato en dimensión tiempo

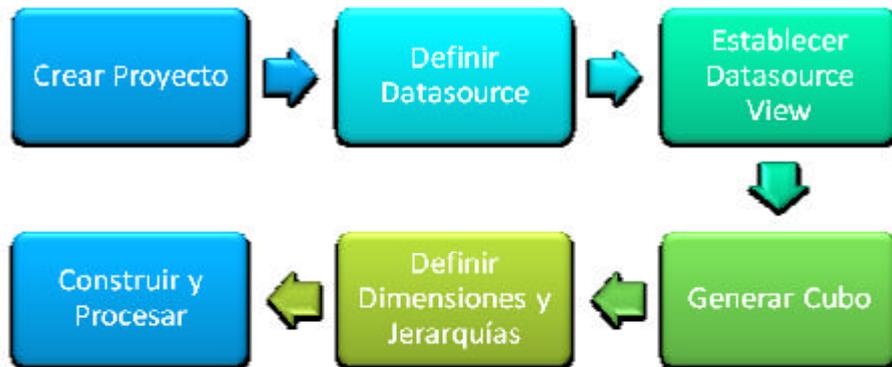
Table - dbo.Dim_Tiempo		
	TiempoSk	Anio
	100	1985
	101	1990
	102	1995
	103	1999
	104	2003
	NULL	2007

- **Tabla de hechos:** Por último, es necesaria la realización de la carga de los datos recopilados del nuevo proceso electoral en la tabla de hechos, esto deberá realizarse utilizando como fase intermedia la tabla stg_factvotos donde serán cargados los datos a través de scripts de inserción, para luego ser transformados en sus correspondientes llaves sk con base a las dimensiones establecidas. En el caso de los procesos electorales, las cargas incrementales no pueden ser automatizadas a través de procesos de Analysis Services producto que dependen de la forma en que los datos sean recopilados y del tipo de información con la que se cuente no teniendo un estándar digital de los mismos. El proceso de carga incremental de la tabla de hechos deberá ser realizado del mismo modo en que se hizo la carga inicial, a través de la ejecución del script de inserción y correlación con las dimensiones.

5.1.2. Esquema lógico

Establecer el esquema lógico de OLAP implica el diseño, creación y manipulación de estructuras multidimensionales que contienen datos detallados y consolidados provenientes de fuentes de datos heterogéneas fusionadas en un modelo único que permite el análisis de grandes cantidades de datos. Este proceso implica la realización de una secuencia organizada de pasos utilizando las herramientas de Analysis Services de Visual Studio .Net:

Figura 54. Proceso de generación de esquema lógico

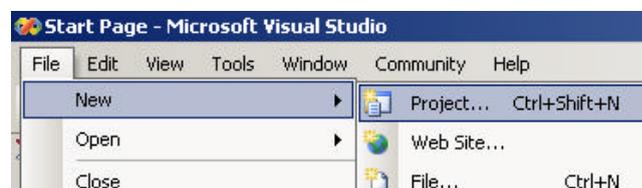


Cada uno de estos pasos genera los elementos necesarios para que el proyecto sea implementado de forma correcta y que pueda establecerse de este modo la interfaz lógica de comunicación OLAP.

5.1.2.1. Crear proyecto

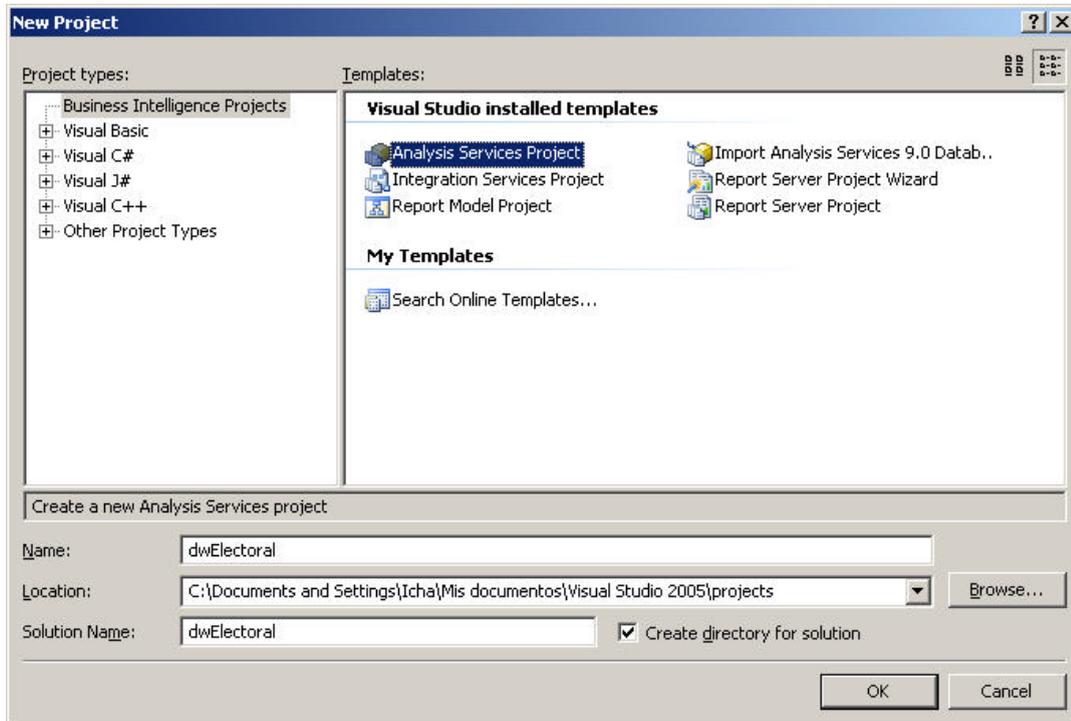
Este paso implica la creación del proyecto de Business Intelligence con Microsoft Visual Studio. Para esto es necesario abrir Visual Studio y utilizar la opción “New\Project”.

Figura 55. Nuevo Proyecto



Se selecciona la opción de Analysis Services Project y se indica como nombre del proyecto “dwElectoral”, en caso de necesitar cambiar la localización donde se almacena el proyecto se debe utilizar el botón Browse e indicar la nueva dirección.

Figura 56. Selección de proyecto de analysis services

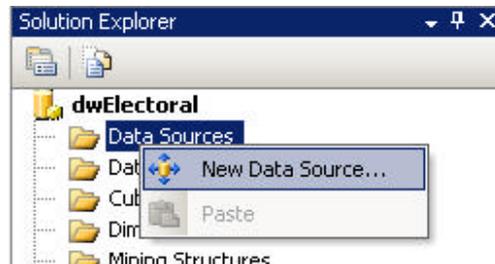


5.1.2.2. Definir Datasource

El siguiente paso corresponde a la creación del datasource, que no es más que un objeto que provee la información necesaria para establecer conexión con las fuentes de datos, permitiendo entonces la comunicación con la base de datos ROLAP donde se encuentran los datos electorales que se desean manipular.

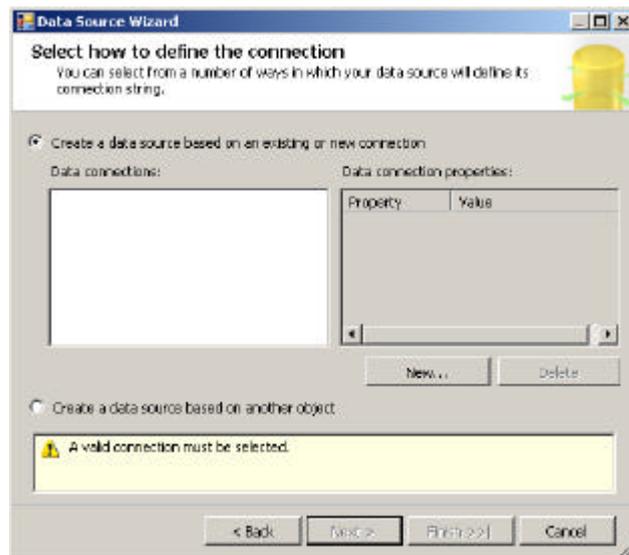
Sobre la carpeta "Data Sources", utilizando el clic derecho del mouse se selecciona la opción de crear un nuevo data source.

Figura 57. Crear nuevo data source



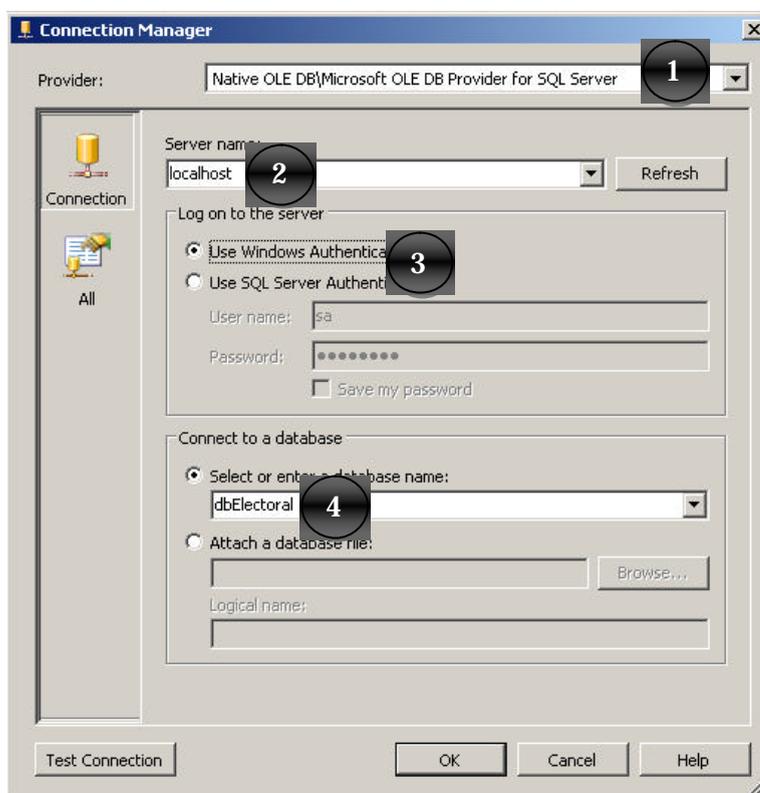
Para la creación del data source es necesario establecer la conexión con la base de datos ROLAP que permita generar el intercambio de información.

Figura 58. Crear conexión a base de datos ROLAP



La siguiente figura muestra los parámetros de configuración establecidos para la conexión que se está creando:

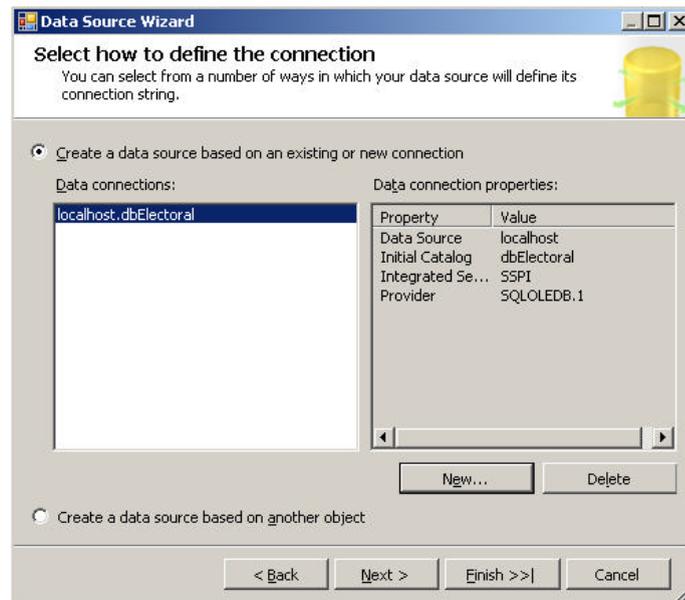
Figura 59. Parámetros de conexión a base de datos ROLAP



- 1) Se define el proveedor OLE DB como SQL Server el cual representa la base de datos ROLAP donde se encuentra creada la base de datos de información.
- 2) El servidor donde se encuentra la base de datos se establece como la máquina local.
- 3) El tipo de autenticación se establece como basado en la autenticación de Windows siendo esta una opción de conexión soportada por SQL Server.
- 4) Se selecciona la base de datos que contiene el esquema ROLAP de la información electoral, el cual sería el creado en la fase de esquema físico.

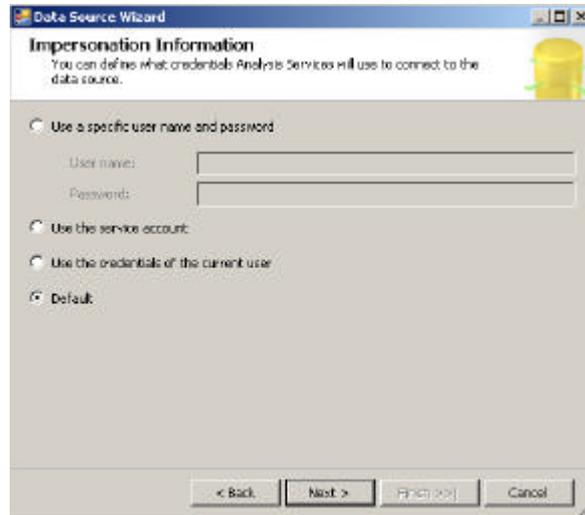
Habiendo creado la conexión con la base de datos se selecciona la misma para ser utilizada por el data source.

Figura 60. Selección de conexión a base de datos



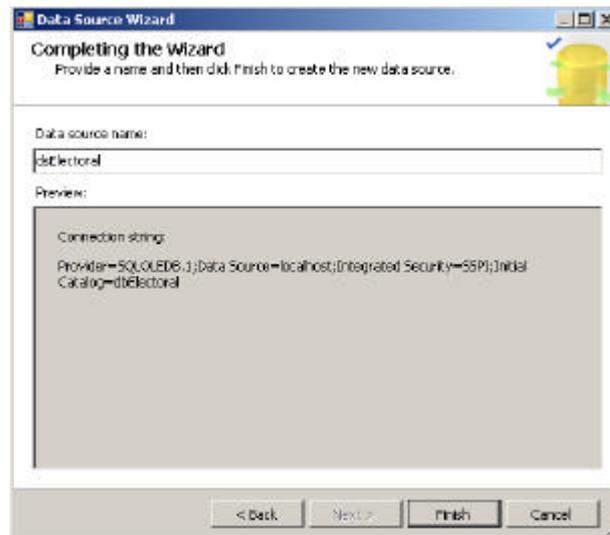
Es importante a continuación definir las credenciales que se utilizarán en la realización de procesos y análisis OLAP, esto a través de la identificación del usuario con los permisos y accesos sobre los objetos de la base de datos, para mayor facilidad y no tener problemas durante el procesamiento del cubo se selecciona la opción Default, la cual selecciona el usuario dueño del esquema.

Figura 61. Parámetro de credenciales



Estos pasos son suficientes para la creación del data sources y se puede observar un resumen de los parámetros establecidos y de la configuración establecida para el data source, debiendo establecerse el nombre del objeto, el cual se llamará "dsElectoral".

Figura 62. Finalización de creación de data source

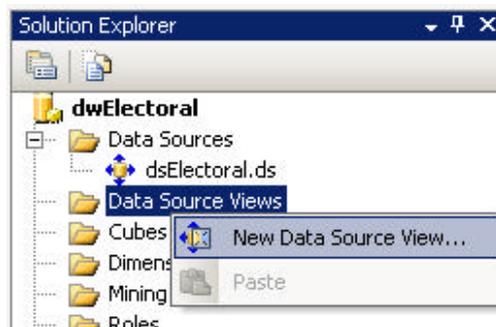


5.1.2.3. Establecer Datasource View

Los proyectos OLAP creados en SQL Server tienen como principal característica que están diseñados a través de un modelo lógico de datos basado en las tablas, vistas, procedimientos y consultas establecidos en uno o más data sources. La creación del data source view establece la metadata de los objetos y sus relaciones generándose entonces una capa espejo de la información ROLAP, a fin de disminuir los accesos a través de la conexión.

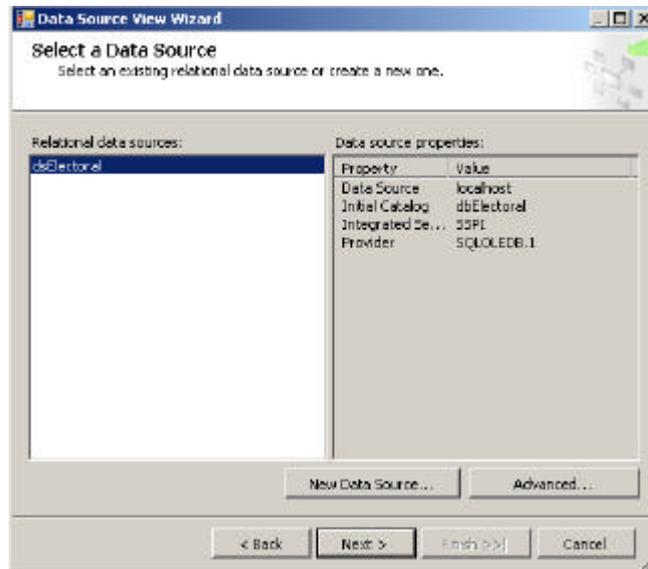
Para crear el data source view se presiona clic derecho sobre la carpeta “Data Source Views” y se selecciona la opción de nuevo data source view, como se muestra en la imagen siguiente.

Figura 63. Creación del data source view



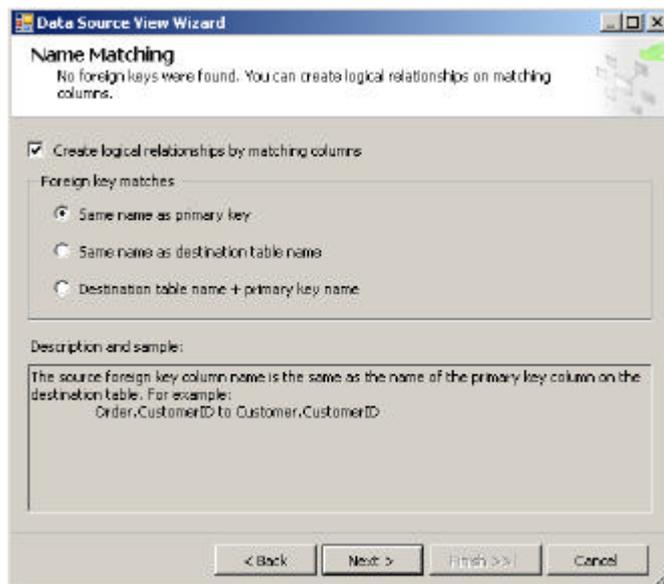
Primeramente aparece la pantalla del asistente para la creación del data source view y a continuación debe indicarse el data source a utilizar siendo en este caso “dsElectoral”.

Figura 64. Selección del data source



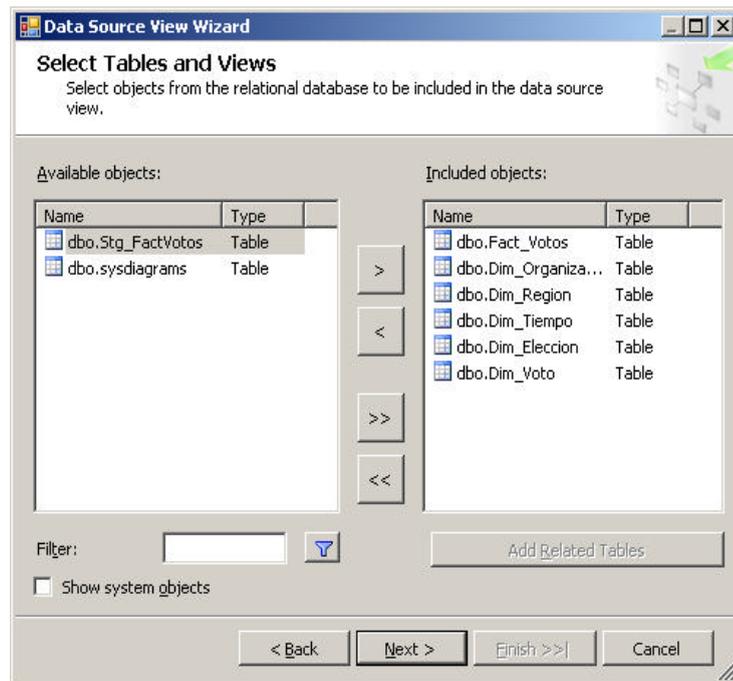
A continuación se establece la relación entre las dimensiones y la tabla de hechos a través de la correspondencia entre los nombres de llaves primarias y foráneas.

Figura 65. Mapeo de relaciones



Luego de haber indicado los parámetros relacionales se identifican las diferentes tablas del data source que definen el cubo de información.

Figura 66. Tablas del data source que forman el cubo



Por último, se identifica el objeto con el nombre “dswElectoral”.

Figura 67. Finalización de la creación de data source view

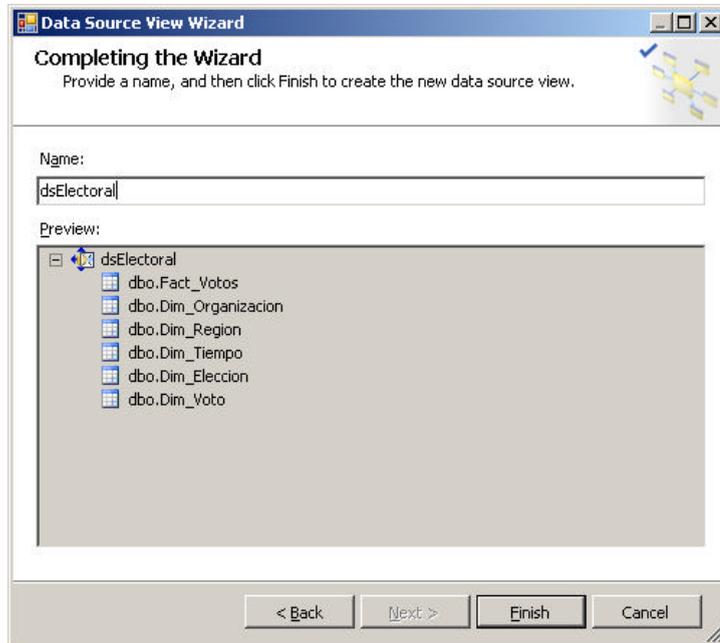
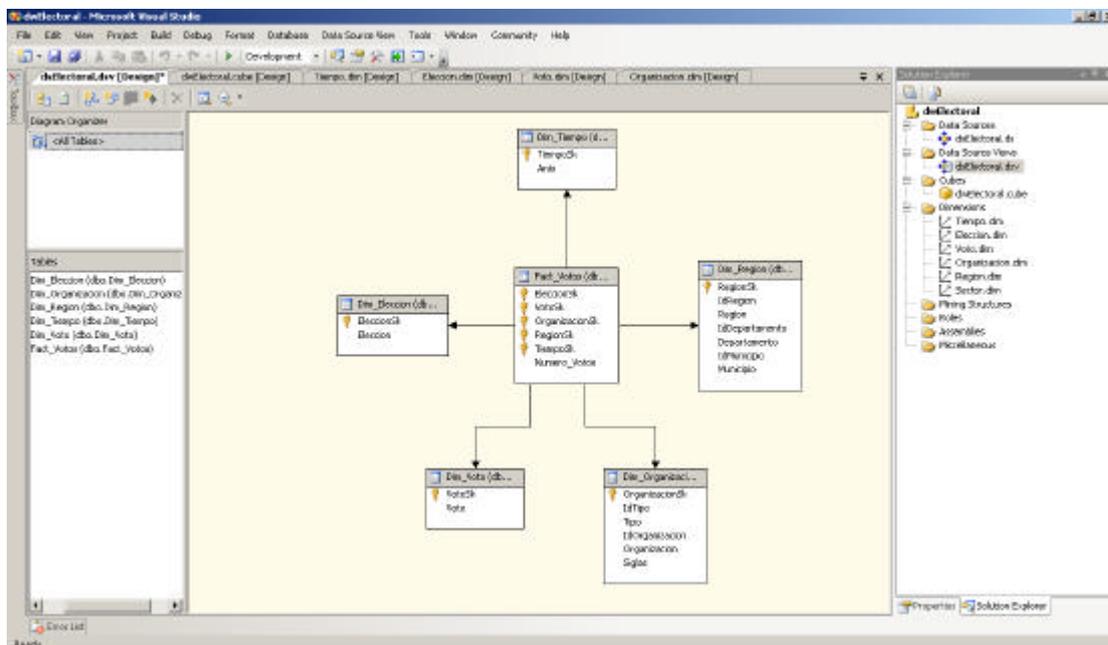


Figura 68. Diseño del data source view



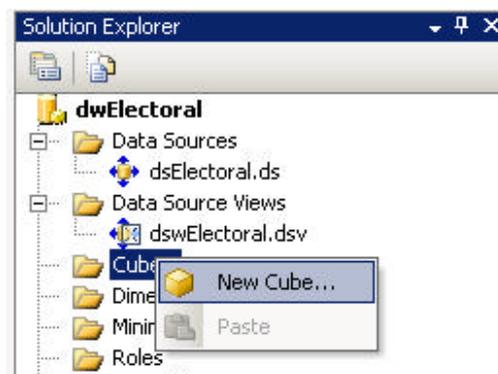
La figura anterior muestra el resultado final de la creación del data source view, pueden verse las dimensiones y sus relaciones con la tabla de hechos a través de las llaves foráneas establecidas durante el proceso de creación del mismo.

5.1.2.4. Generar cubo

Establecer el cubo permite el reconocimiento automático de las dimensiones y la tabla de hechos que componen el Datawarehouse electoral, definiéndose de este modo el eje central de la aplicación.

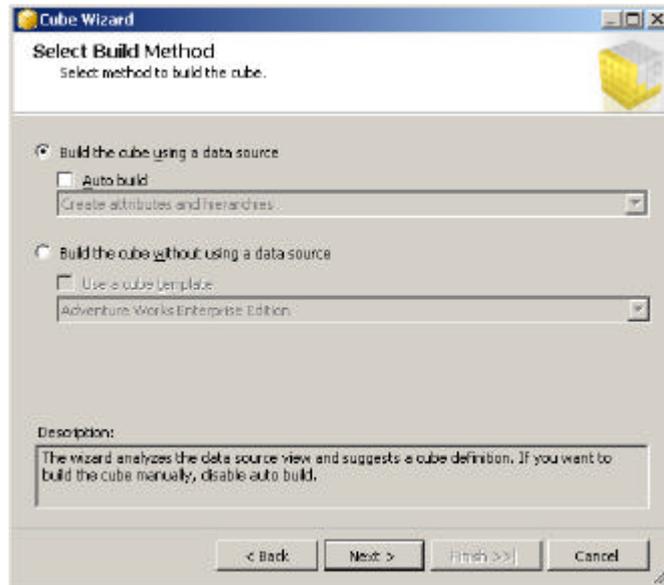
La creación del nuevo cubo debe realizarse a través de la utilización de esta opción sobre la carpeta Cube al dar clic derecho sobre la misma.

Figura 69. Creación de nuevo cubo



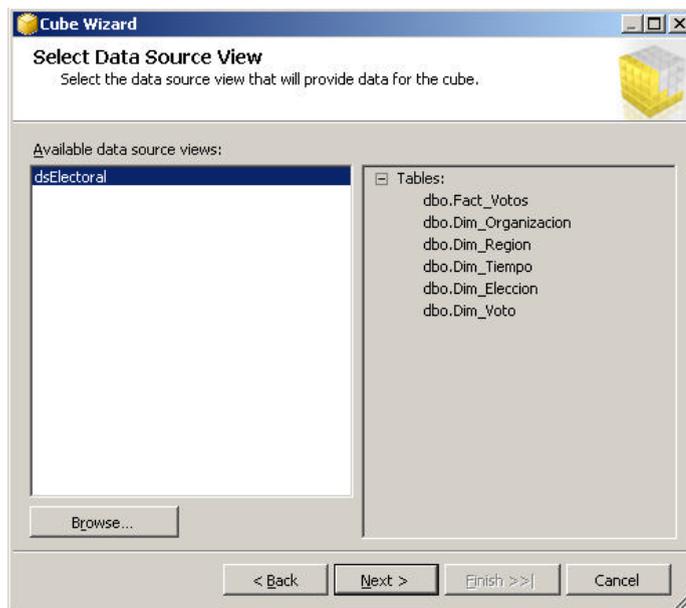
Se establece que el cubo será definido a través de las tablas referenciadas por el data source y las relaciones existente en el mismo tal como se muestra en la figura a continuación.

Figura 70. Selección de fuente de datos



A continuación se indica el data source view que contendrá y proveerá la información del cubo, el cuál fue creado en el paso anterior.

Figura 71. Indicar data source view del cubo



El siguiente paso consiste en la identificación de las dimensiones del modelo, señalización de la tabla de hechos y el establecimiento de la dimensión de tiempo a utilizar, siendo este el paso más importante, ya que permite establecer el modelo del cubo. Este paso también permite ver un diagrama general de las relaciones entre las dimensiones y la tabla de hechos.

Figura 72. Definición de dimensiones, tabla de hechos y tiempo

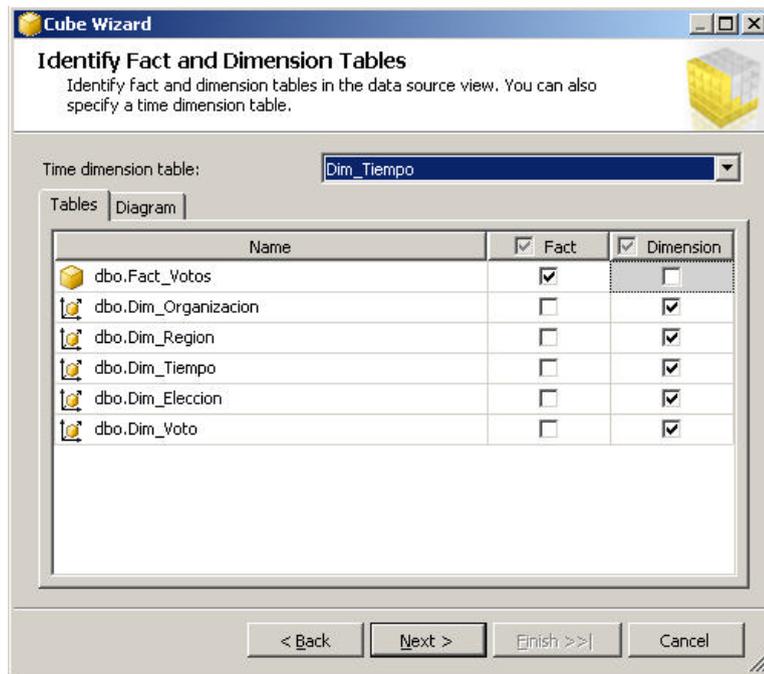
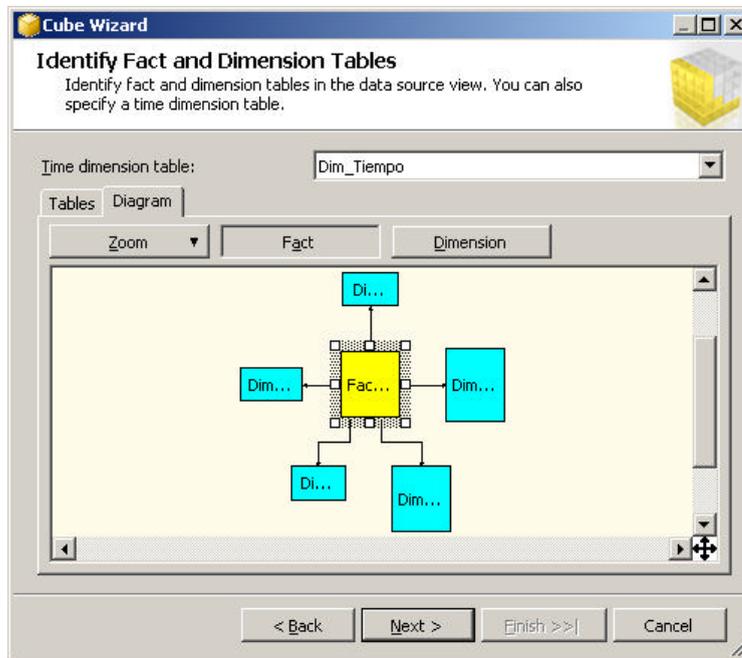
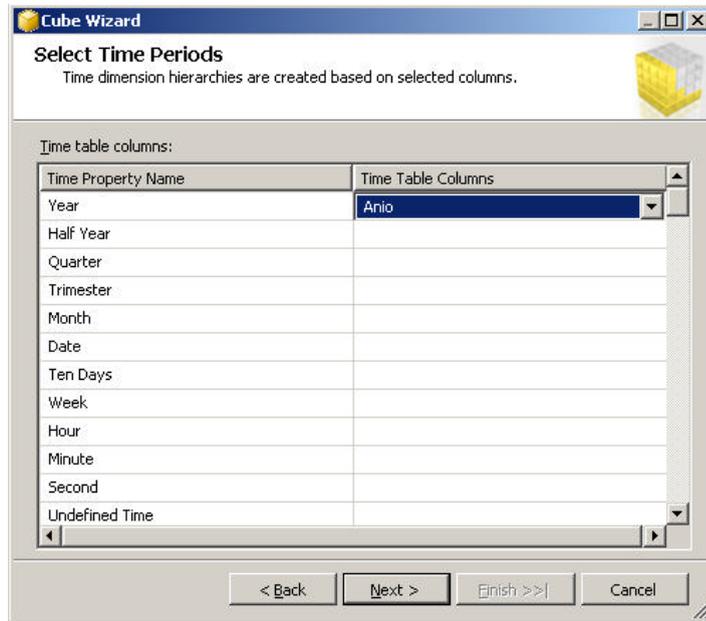


Figura 73. Diseño general de modelo OLAP



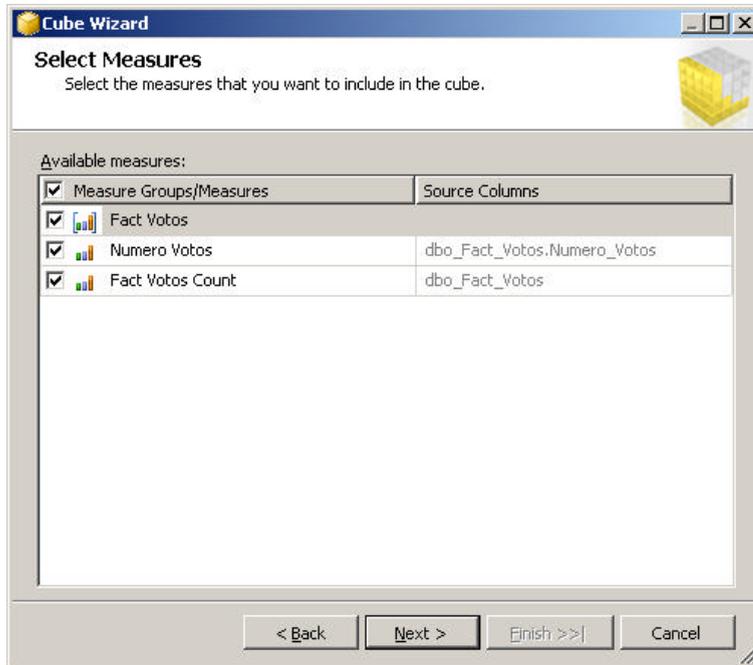
Luego de haber establecido el modelo OLAP a través de las dimensiones y sus relaciones con la tabla de hechos, se definen los periodos de la dimensión tiempo. En el caso de los procesos electorales sólo será el año el valor a tomar como elemento de temporalidad debido a que los procesos electorales se realizan cada 4 años y la fecha exacta de la realización del plebiscito no tiene relevancia dentro del cubo de información.

Figura 74. Establecimiento de dimensión tiempo



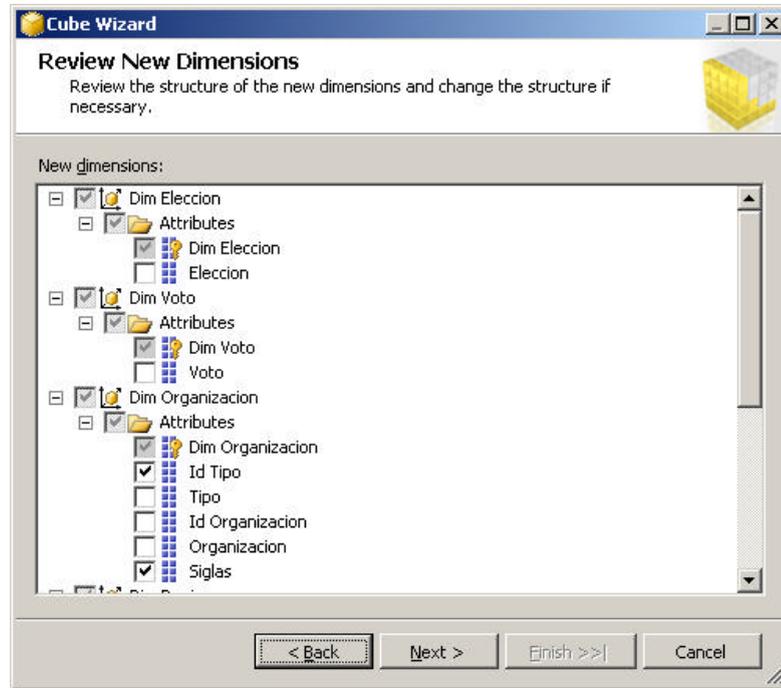
A continuación se definen las diferentes métricas que serán llevadas en la tabla de hechos del cubo OLAP, en este caso sólo serán el número de votos y el conteo de votos.

Figura 75. Definición de métricas



Se establecen los valores que forman las diferentes dimensiones, en este paso es importante tomar en cuenta que deben seleccionarse únicamente aquellos valores que se mostrarán al usuario al momento de realizar reportes.

Figura 76. Selección de valores de dimensiones



Por último, se establece el nombre del cubo y se observa una descripción general de los parámetros configurados previamente.

Figura 77. Finalización de la creación del cubo

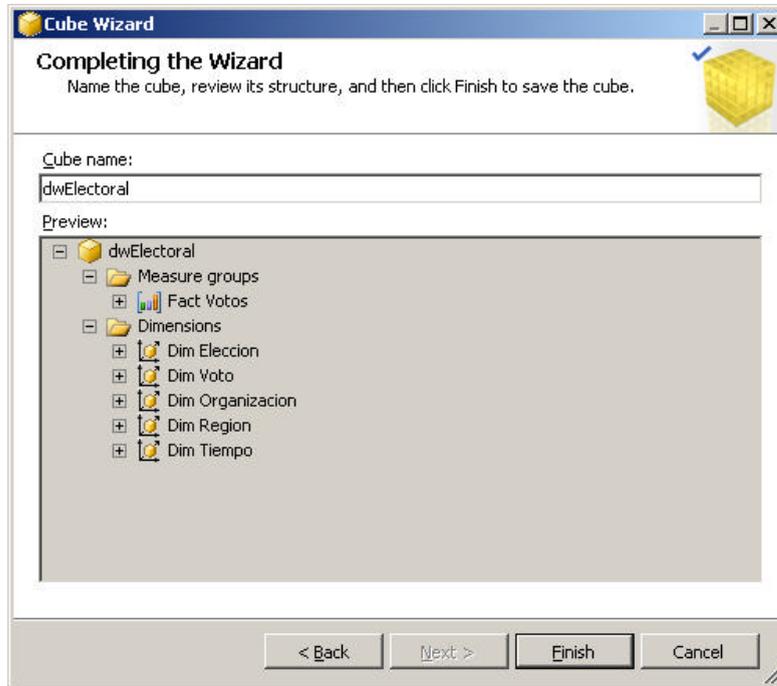
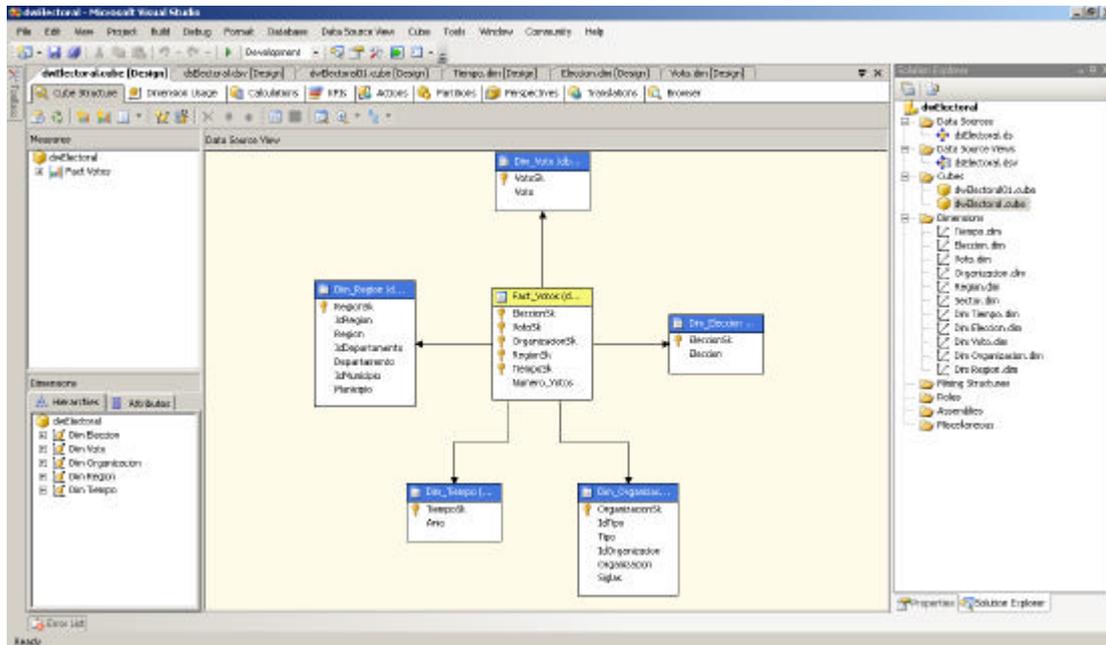


Figura 78. Diseño del cubo dwElectoral



5.1.2.5. Definir dimensiones y jerarquías

Es importante luego de haber creado el cubo de información establecer que las dimensiones proyecten al usuario valores de atributos que sean significativos y que reflejen a su vez las relaciones jerárquicas establecidas durante la fase de análisis y diseño que permitan la navegación intuitiva a través de ello.

Para la realización de este paso es necesario llevar a cabo dos actividades dentro de cada dimensión:

- 1) Establecer los valores descriptivos de los atributos que componen la dimensión, esto implica asignar a cada atributo identificador o id su correspondiente descriptor en la propiedad NameColumn, para lo cual es necesario presionar clic derecho sobre el atributo identificador y seleccionar Propiedades.

Figura 79. Propiedades de los atributos dimensionales

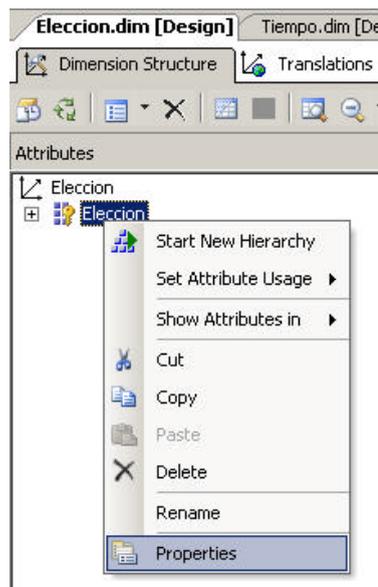
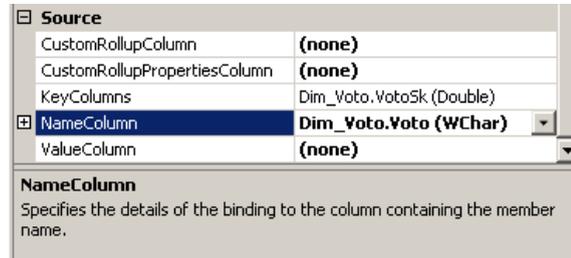
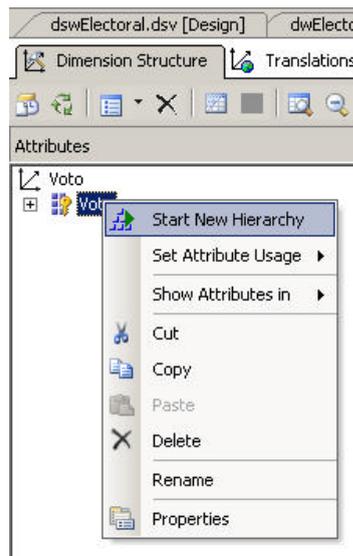


Figura 80. Propiedad NameColumn



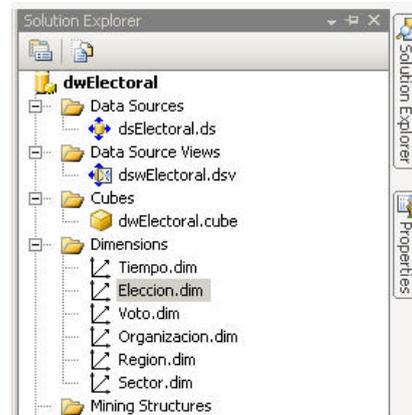
- 2) Definir las relaciones jerárquicas entre los atributos dimensionales, esto se realiza a través de la opción “Start New Hierarchy” del menú contextual.

Figura 81. Creación de nueva jerarquía



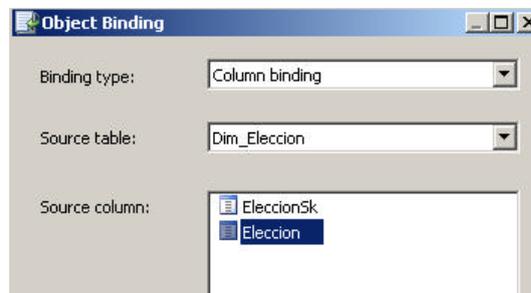
Navegar dentro de cada dimensión puede hacerse al presionar doble clic en la dimensión seleccionada en la ventana del explorador del proyecto y con esto se podrán observar los atributos, jerarquías y fuentes que establecen las dimensiones.

Figura 82. Mantenimiento de dimensiones



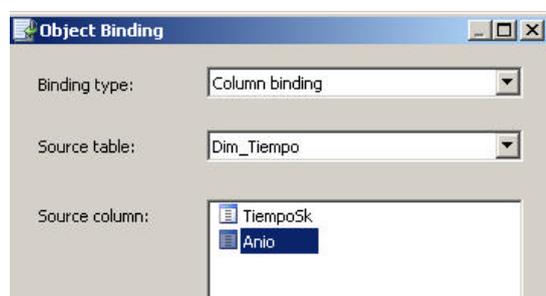
- **Dimensión elección:** La configuración de esta dimensión requiere únicamente establecer el valor descriptivo del atributo ElecciónSk, esto a través de seleccionar el atributo Elección.

Figura 83. Configuración de dimensión elección



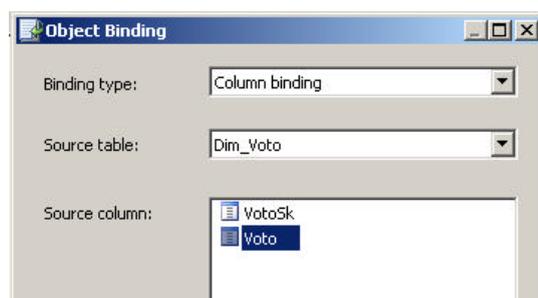
- **Dimensión tiempo:** Esta dimensión requiere definir el atributo año como el valor descriptivo de la dimensión en la propiedad NameColumn.

Figura 84. Configuración de dimensión tiempo



- **Dimensión voto:** Para establecer el valor descriptivo de esta dimensión se identifica con el atributo Voto en la propiedad NameColumn

Figura 85. Parámetros de la dimensión voto



- **Dimensión organización:** Para que la dimensión muestre al usuario los valores descriptivos de cada atributo, se configuró que OrganizacionSk se representa por el atributo Organización y IdTipo por Tipo, quedando el atributo Siglas como un valor descriptivo adicional.

Figura 86. Configuración de dimensión organización (atributo IdTipo)

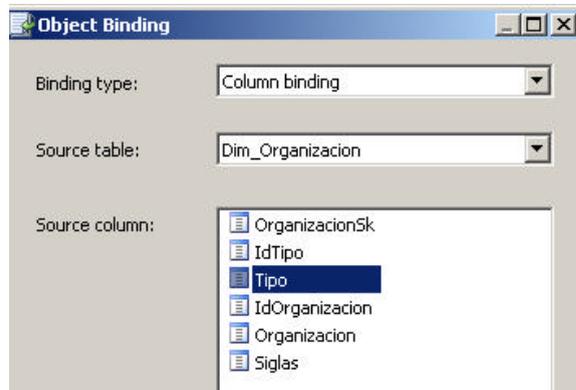
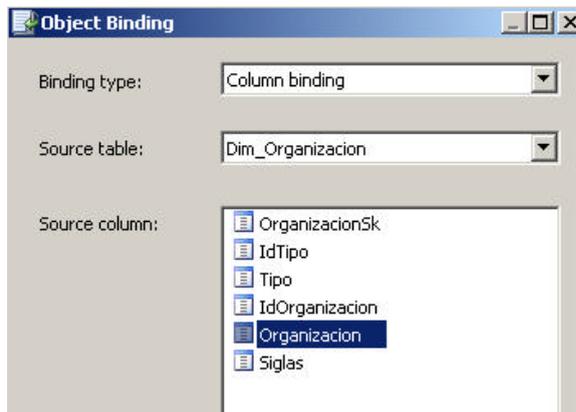
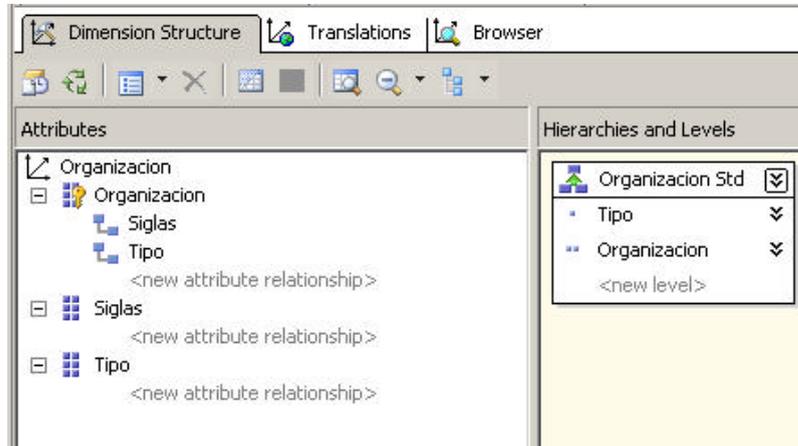


Figura 87. Configuración de dimensión organización (atributo OrganizacionSk)



La jerarquía de esta dimensión está establecida con base al tipo de organización y su nombre o siglas, para esto se crea una nueva jerarquía y se van arrastrando los atributos que componen la jerarquía primero el Tipo y luego Organización.

Figura 88. Jerarquías de la dimensión organización



- **Dimensión región:** Establecer los valores descriptivos de esta dimensión requiere de establecer en la propiedad NameColumn de los atributos RegionSk, DepartamentoId y RegionId los atributos Municipio, Departamento y Región respectivamente.

Figura 89. Configuración de dimensión región (atributo RegionSk)

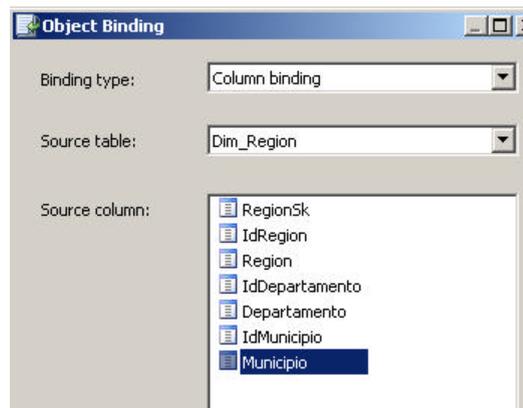
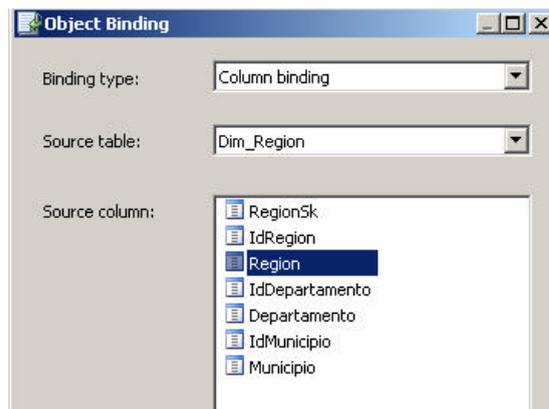


Figura 90. Parámetros de dimensión región (atributo IdDepartamento)

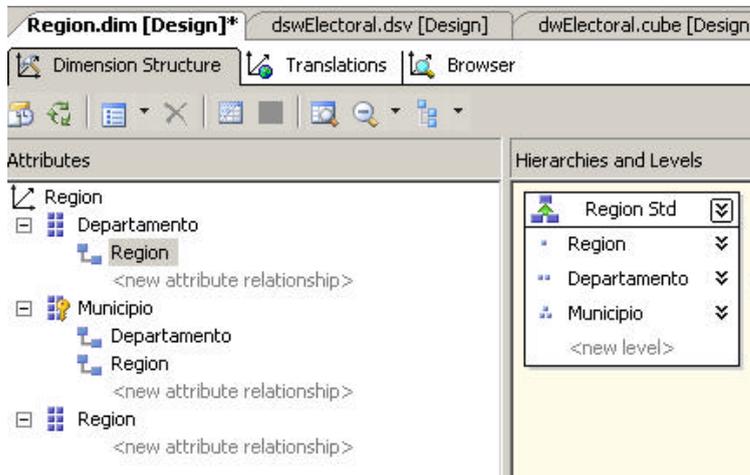


Figura 91. Configuración de dimensión región (atributo IdRegion)



Luego de haber definido la configuración de la dimensión se debe establecer la jerarquía generada por la relación entre Región, Departamento y Municipio, para lo que debe crearse la nueva jerarquía, arrastrando los atributos relacionados.

Figura 92. Jerarquía de dimensión región



Para que la jerarquía creada tenga relaciones naturales, es necesario ir poniendo los atributos relacionados jerárquicamente debajo del hijo inmediato, como se muestra en la figura, el atributo Departamento tiene la relación establecida con Región y a su vez el atributo Municipio con Departamento y Región.

5.1.2.6. Construir y procesar

Luego de haber establecido la configuración lógica del modelo de dimensiones, jerarquías y del cubo de datos corresponde la creación del esquema físico que soporte el manejo de la información del Datawarehouse.

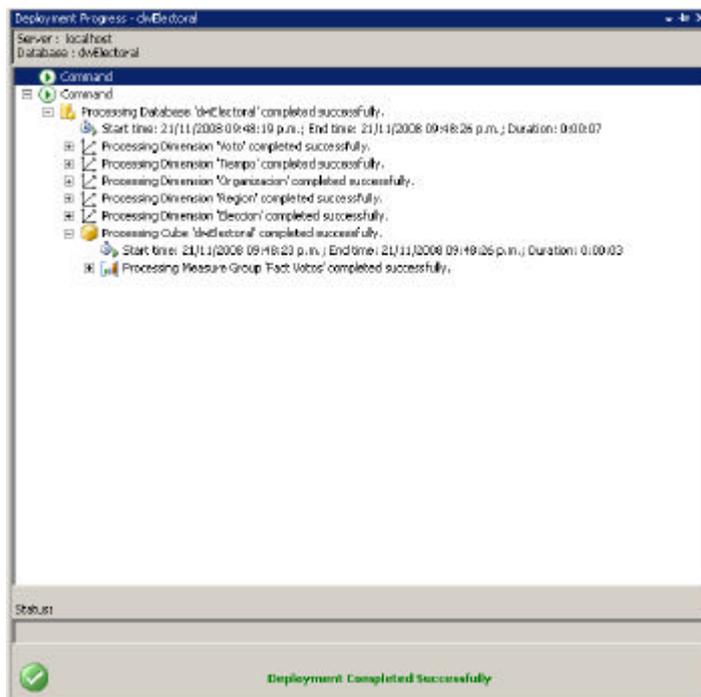
La opción de “Build\Deploy” permite crear el esquema físico en la base de datos SQL Server del modelo lógico del Datawarehouse, generando de este modo el modelo OLAP de Analysis Services.

Figura 93. Creación del esquema físico del modelo OLAP



En caso de que las dimensiones, relaciones, jerarquías y demás componentes del cubo estén creadas de forma correcta la opción de construcción mostrará el mensaje de que el cubo ha sido generado de forma satisfactoria.

Figura 94. Resultados de la construcción del cubo

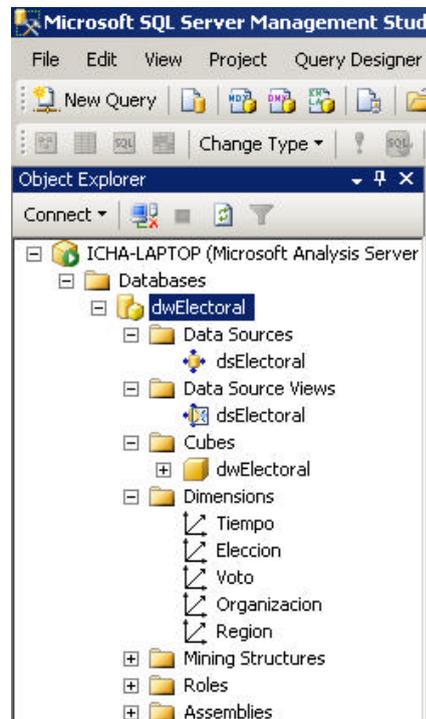


Para observar que toda la estructura ha sido creada de forma satisfactoria puede conectarse al SQL Server en el esquema de Analysis Services y ver el modelo generado.

Figura 95. Ingreso a SQL Server Analysis Services

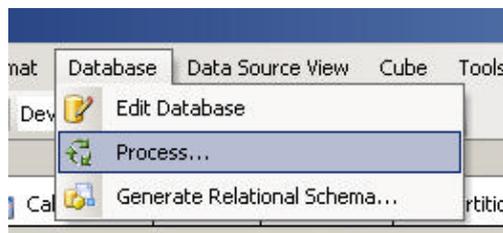


Figura 96. Estructura física del cubo dwElectoral



A continuación debe usarse la opción “Database\Process” que tiene como principal función la extracción de los datos de la base de datos ROLAP y la carga de estos en el modelo lógico y físico OLAP para que puedan ser consultados posteriormente.

Figura 97. Procesamiento del cubo OLAP



La siguiente ventana muestra los valores de configuración para el procesamiento de los datos y deberá seleccionarse el botón Run para que comience a procesarse el cubo de información.

Figura 98. Inicio del procesamiento del cubo OLAP

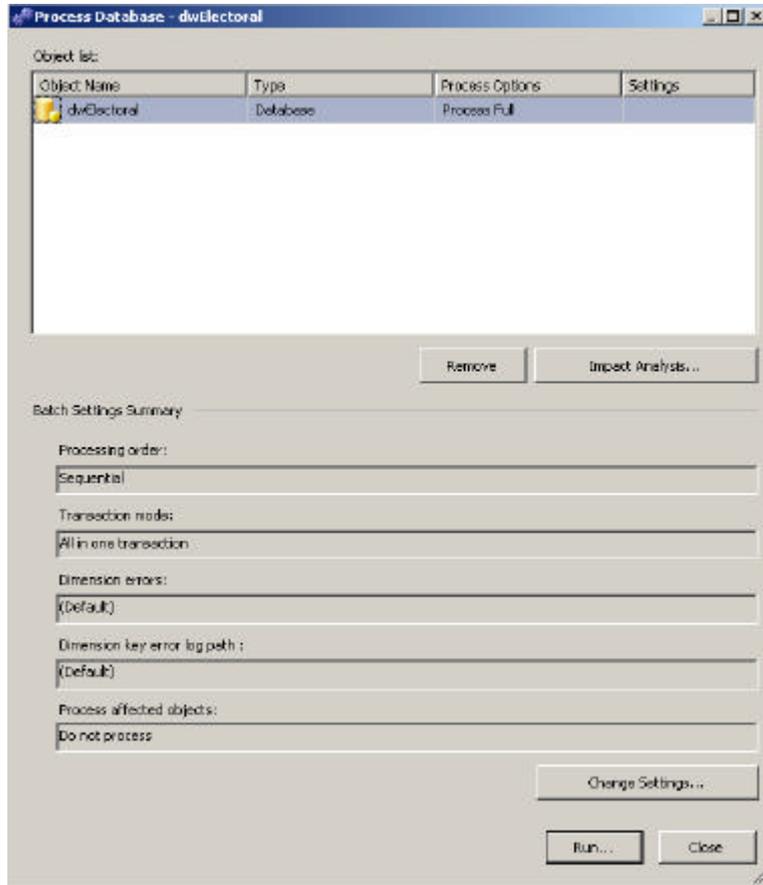
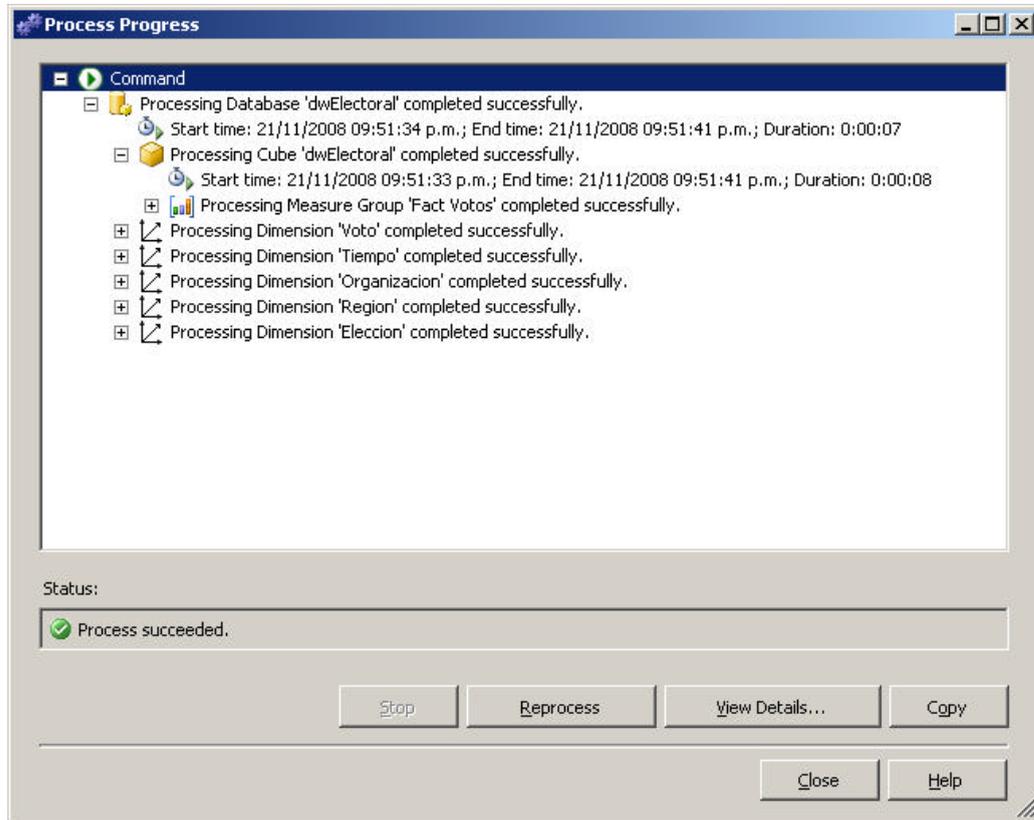


Figura 99. Procesamiento del cubo OLAP



Luego de haber realizado el procesamiento de los datos y el vaciado de los mismos en el modelo OLAP, Visual Studio provee de la herramienta llamada Browse que permite realizar consultas de los datos contenidos en el cubo a fin de verificar que la información cargada sea la correcta.

Figura 100. Herramienta browse de Analysis Services

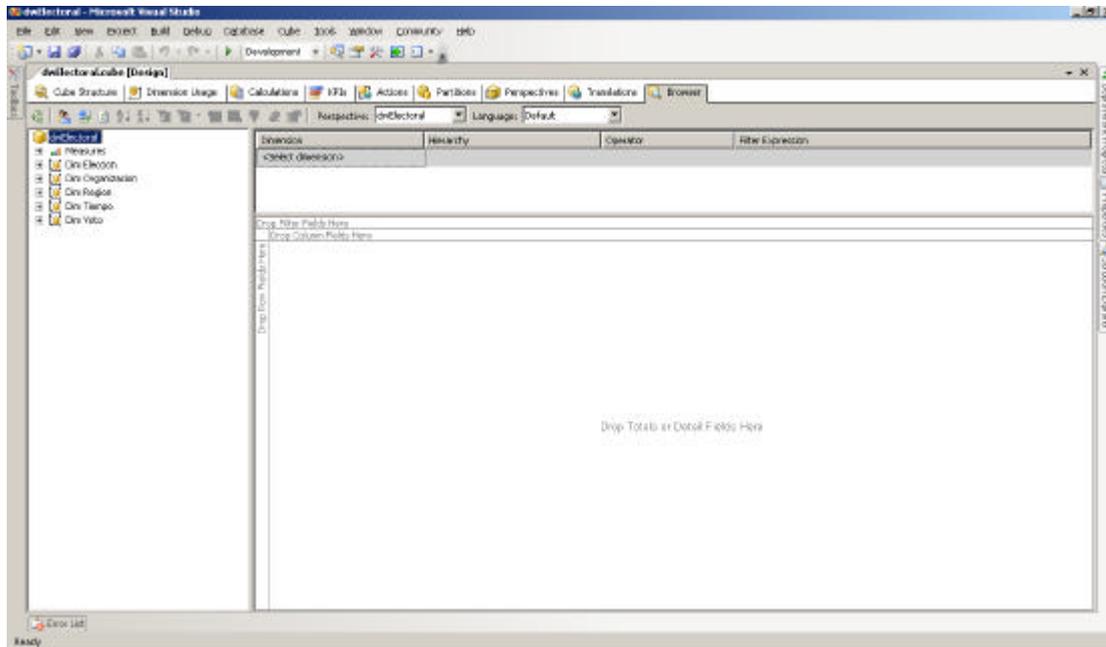
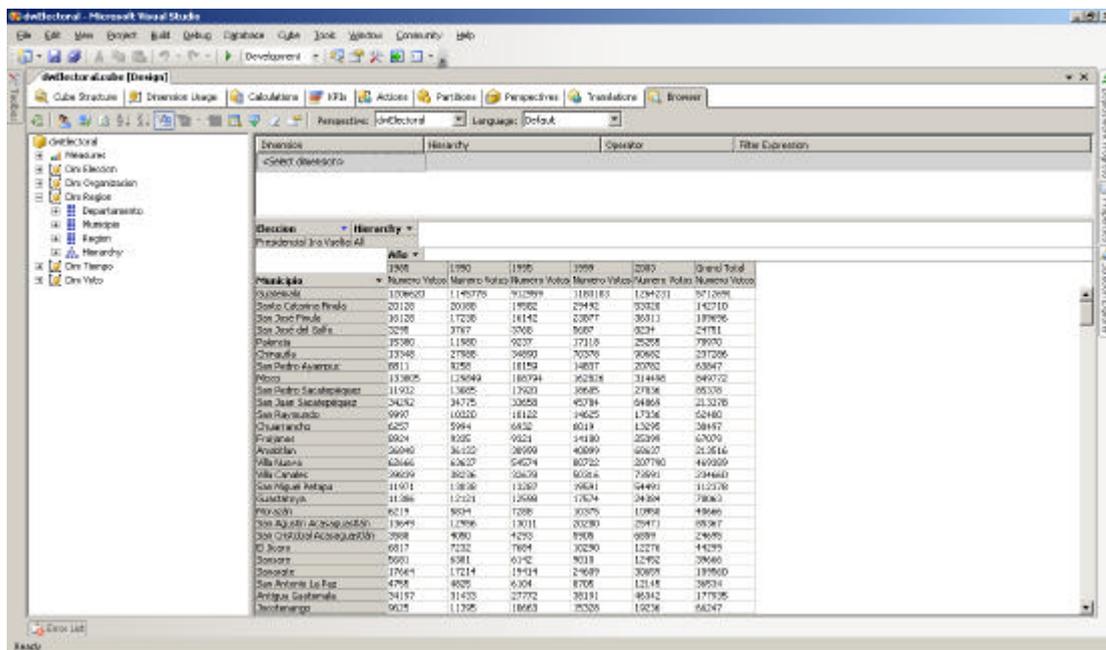


Figura 101. Resultados de consulta

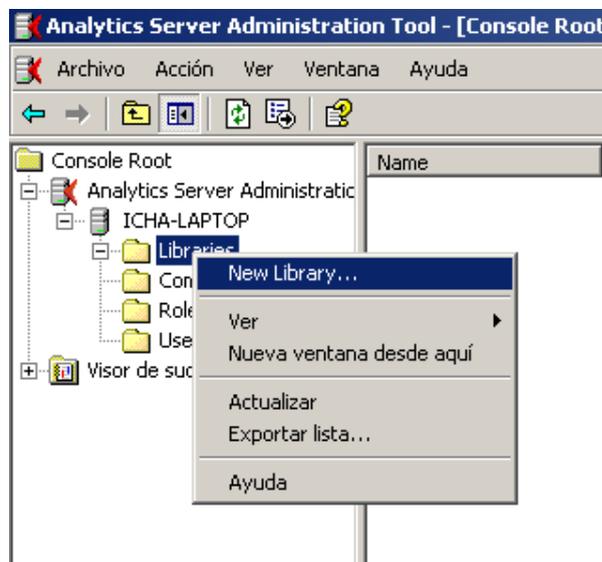


5.1.3. Esquema de análisis

Crear el esquema de análisis requiere el establecimiento de la conexión entre el modelo lógico creado a través de Visual Studio y una vista de Proclarity que permita la realización de consultas personalizadas sobre los datos del Datawarehouse.

El primer paso corresponde a la creación de una nueva librería, para esto debe ingresarse a la herramienta de administración del servidor de Proclarity y en la carpeta de librerías, usando el botón derecho del mouse seleccionar la creación de una nueva.

Figura 102. Creación de nueva librería de Proclarity



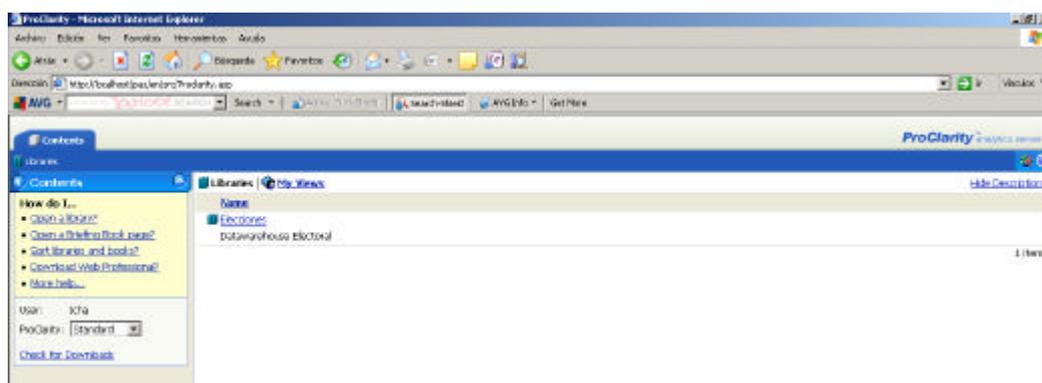
Debe indicarse el nombre de la nueva librería así como una descripción, en este caso será como se muestra en la siguiente figura.

Figura 103. Asignación de nombre y descripción a la librería



Con este paso ya es suficiente para conectarse a Proclarity usando la dirección <http://localhost/pas>, en caso de estar implementado en un servidor los clientes deberán conectarse usando el nombre del servidor en vez de localhost.

Figura 104. Librería creada en Proclarity



A continuación deberá darse clic sobre la nueva librería Elecciones, para establecer el cubo de información al que hace referencia indicando el servidor y la base OLAP.

Figura 105. Escoger el servidor OLAP

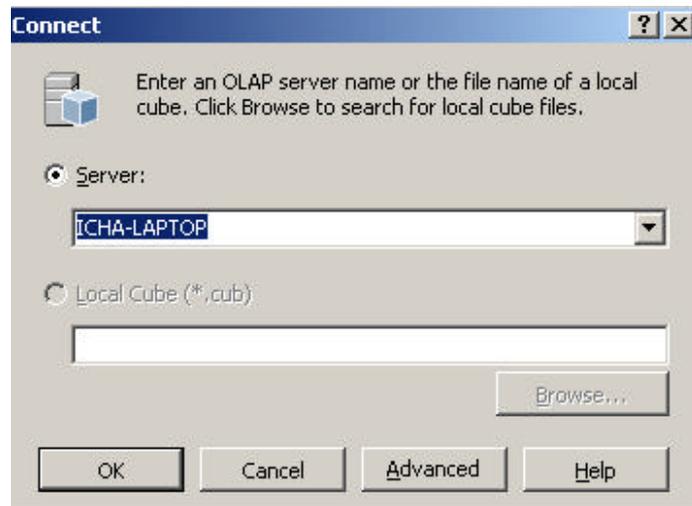
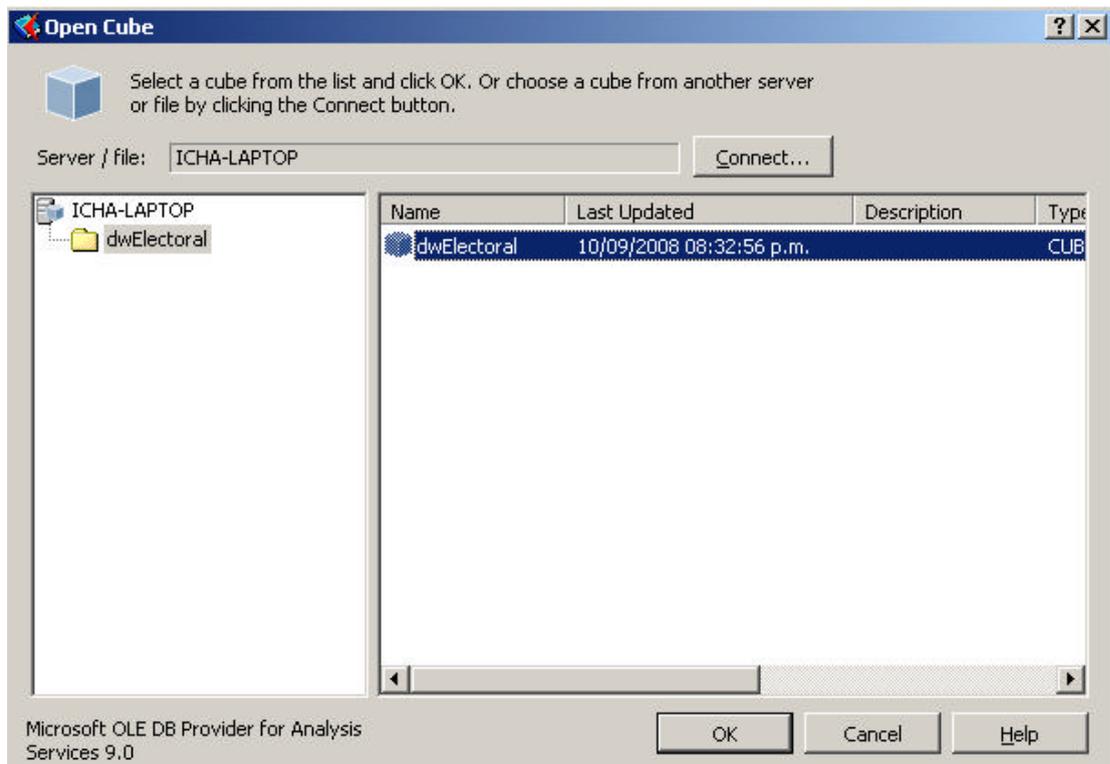


Figura 106. Selección de cubo dwElectoral



5.2. Reportes

Luego de haber realizado todos los pasos anteriores, se cuenta con la estructura física y lógica necesaria para la realización de reportes personalizados que permitan la realización de análisis de datos electorales, para esto se hace uso de la herramienta Proclarity Professional, a continuación se muestran ejemplos de algunos reportes personalizados.

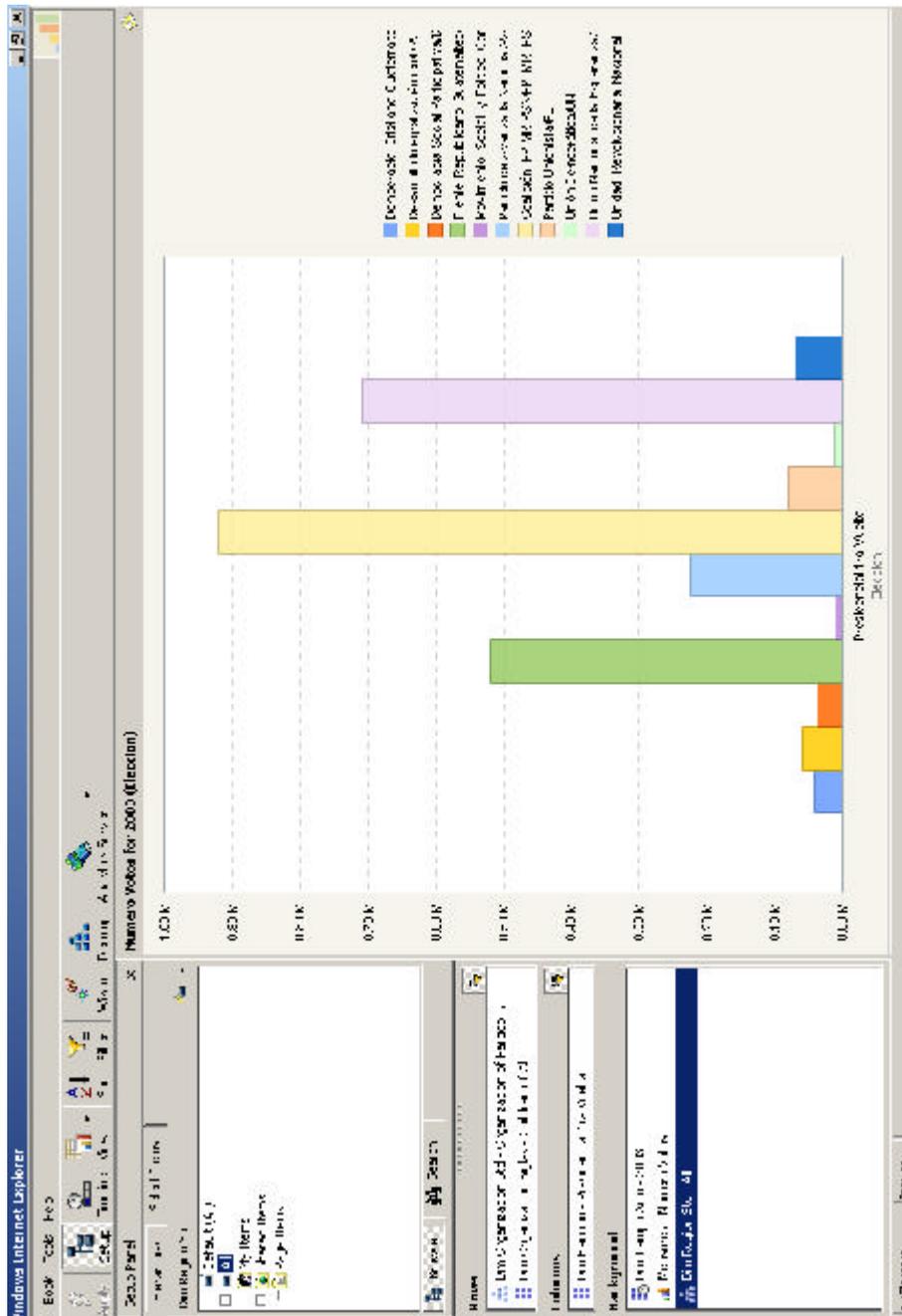
5.2.1. Votos por Organización Política en la 1ra vuelta presidencial

El siguiente reporte muestra los votos obtenidos por organización política en la 1ra vuelta de elecciones presidenciales en las elecciones realizadas en el año 2003 a nivel nacional. Estos resultados pueden verse en el archivo de Excel correspondiente al año 2003, estableciéndose la relación de resultados.

Figura 107. Resultados en excel de votos por organización en 1ra vuelta presidencial del año 2003

Composición del Voto por Organización Política											
PP-MR-PSN	UN	FRG	DIA	URNG	PU	DSP	UNE	MSPCN	DCG	PAN	Emitidos
1,707	7	2,783	10	11	19	9	253	8	41	83	4,931
3,058	9	2,085	47	56	356	11	1,315	13	234	1,254	8,438
680	12	452	12	3	63	1	1,482	3	315	46	3,069
295	19	1,060	16	8	10	1	293	12	21	2,197	3,932
0	0	0	0	0	0	0	0	0	0	0	0
921,316	11,980	518,464	59,776	69,301	80,987	37,506	707,635	10,830	42,205	224,179	2,684,179

Figura 109. Gráfica de resultados electorales del año 2003, por organización política en 1ra vuelta presidencial.



5.2.2. Votos emitidos en los distintos procesos electorales en municipios del departamento de Guatemala

Puede observarse el comportamiento del voto en los distintos años electorales para los municipios del departamento de Guatemala.

Figura 110. Emisión de voto en municipios de Guatemala

Numero Votos for Municipales (Municipio by Año)					
	1985	1990	1995	1999	2003
Guatemala	382,625	374,341	910,439	1,183,654	1,263,327
Santa Catarina Pinula	6,153	6,389	19,587	29,635	53,256
San José Pinula	4,947	5,424	16,243	23,895	35,743
San José del Golfo	964	1,216	3,822	5,728	8,271
Palencia	3,001	0	0	17,081	25,216
Chimutla	4,065	0	0	70,162	90,543
San Pedro Ayampuc	2,664	2,259	10,243	14,915	20,674
Mixco	40,904	40,946	108,372	162,294	312,902
San Pedro Sacatepéquez	3,687	4,140	14,027	18,719	27,661
San Juan Sacatepéquez	10,256	10,563	33,691	45,733	64,372
San Raymundo	3,004	3,247	10,170	14,634	17,367
Chuamrancho	1,900	1,870	7,299	6,028	13,300
Fraijanes	2,769	2,944	9,434	13,932	25,295
Amatitlan	11,200	11,238	30,922	40,856	68,391
Villa Nueva	19,157	19,801	54,136	80,285	205,795
Villa Canales	11,976	11,625	32,554	50,233	73,516
San Miguel Petapa	3,595	4,134	13,271	19,515	54,323
Grand Total	514,067	500,837	1,274,220	1,797,309	2,362,212

5.2.3. Votos de elecciones de diputación nacional en el año 1995, por departamento según el tipo de voto.

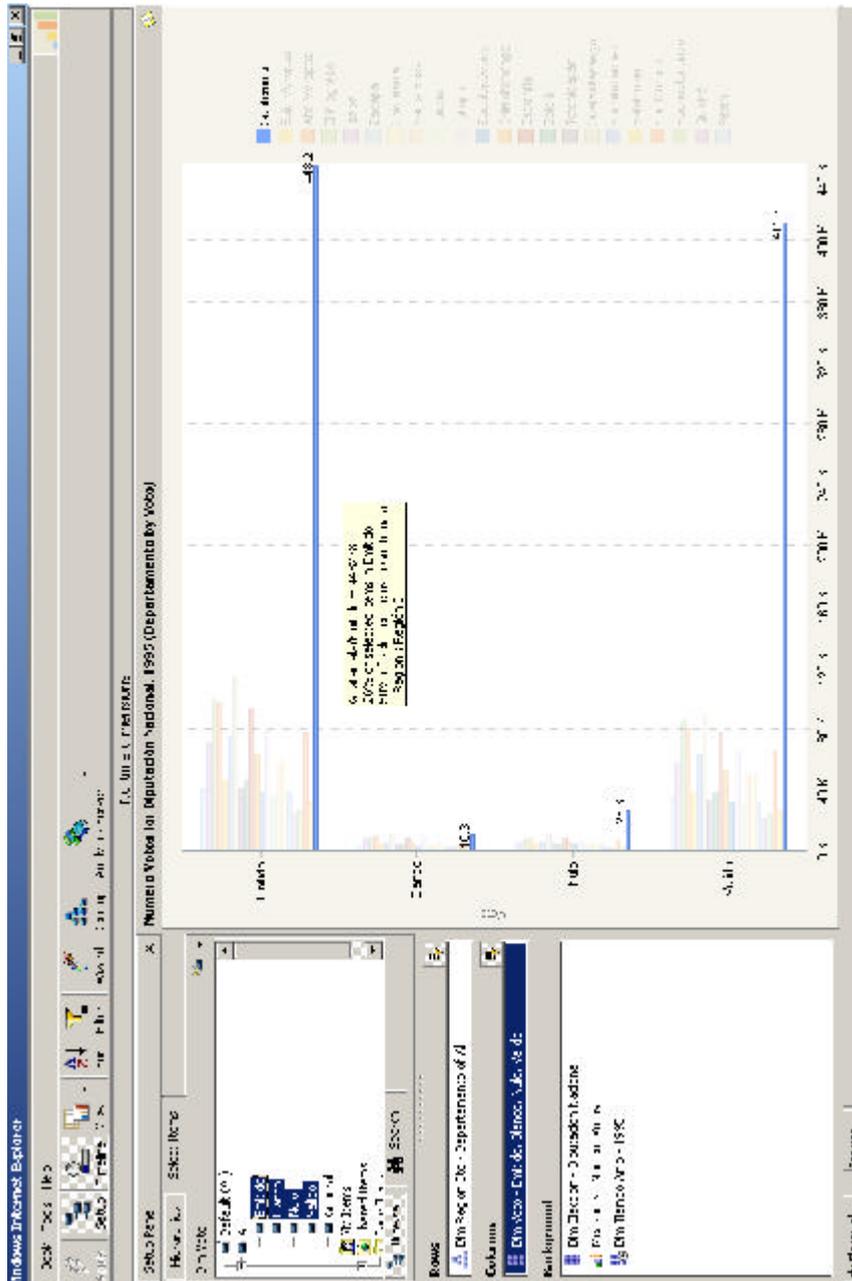
El siguiente reporte muestra el número de votos emitidos durante las elecciones de 1995, para la elección de diputados nacionales resumidos por departamentos.

Figura 112. Total de votos emitidos según el tipo en elección de 1995

	Emitido	Blanco	Nulo	Valido
Guatemala	448,238	10,278	26,298	411,662
Baja Verapaz	31,330	2,358	2,811	26,161
Alta Verapaz	76,994	4,755	6,799	65,440
El Progreso	26,597	1,947	1,622	23,028
Izabal	24,314	1,636	1,846	20,832
Zacapa	37,994	2,588	3,545	31,861
Chiquimula	59,490	4,802	4,437	50,251
Santa Rosa	57,746	4,859	4,551	48,336
Jalapa	36,349	3,035	2,578	30,736
Jutiapa	79,211	7,331	5,516	66,364
Sacatepéquez	38,218	3,005	3,412	31,801
Chimaltenango	63,099	5,739	5,010	52,350
Escuintla	93,357	7,636	8,274	77,447
Sololá	45,956	3,587	4,111	38,258
Totonicapán	40,793	3,594	4,125	33,074
Quetzaltenango	114,043	11,207	11,418	91,418
Suchitepéquez	74,792	6,003	5,729	63,060
Retalhuleu	45,685	3,949	3,586	38,150
San Marcos	97,258	9,203	8,219	79,836
Huehuetenango	100,438	7,405	7,794	85,239
Quiché	70,919	7,347	5,409	58,163
Petén	40,795	3,087	3,833	33,875
Grand Total	1,703,616	115,351	130,923	1,457,342

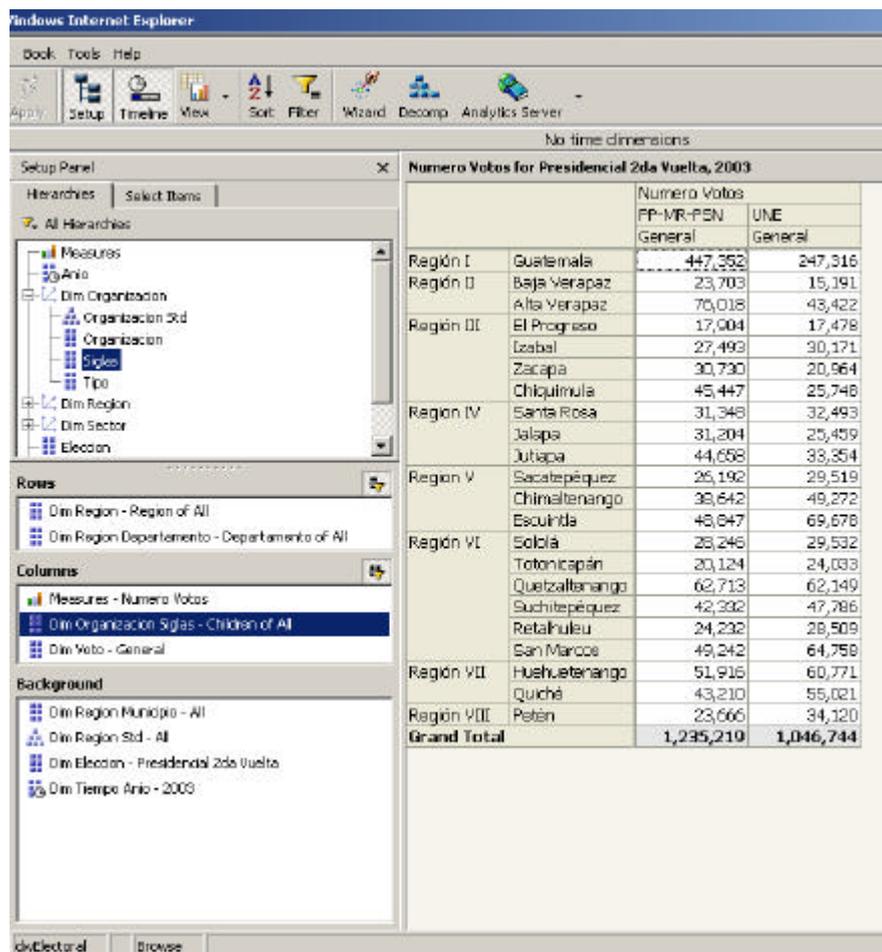
Además puede crearse el gráfico que permita ver la selección para un departamento específico.

Figura 113. Gráfico de selección por un departamento específico



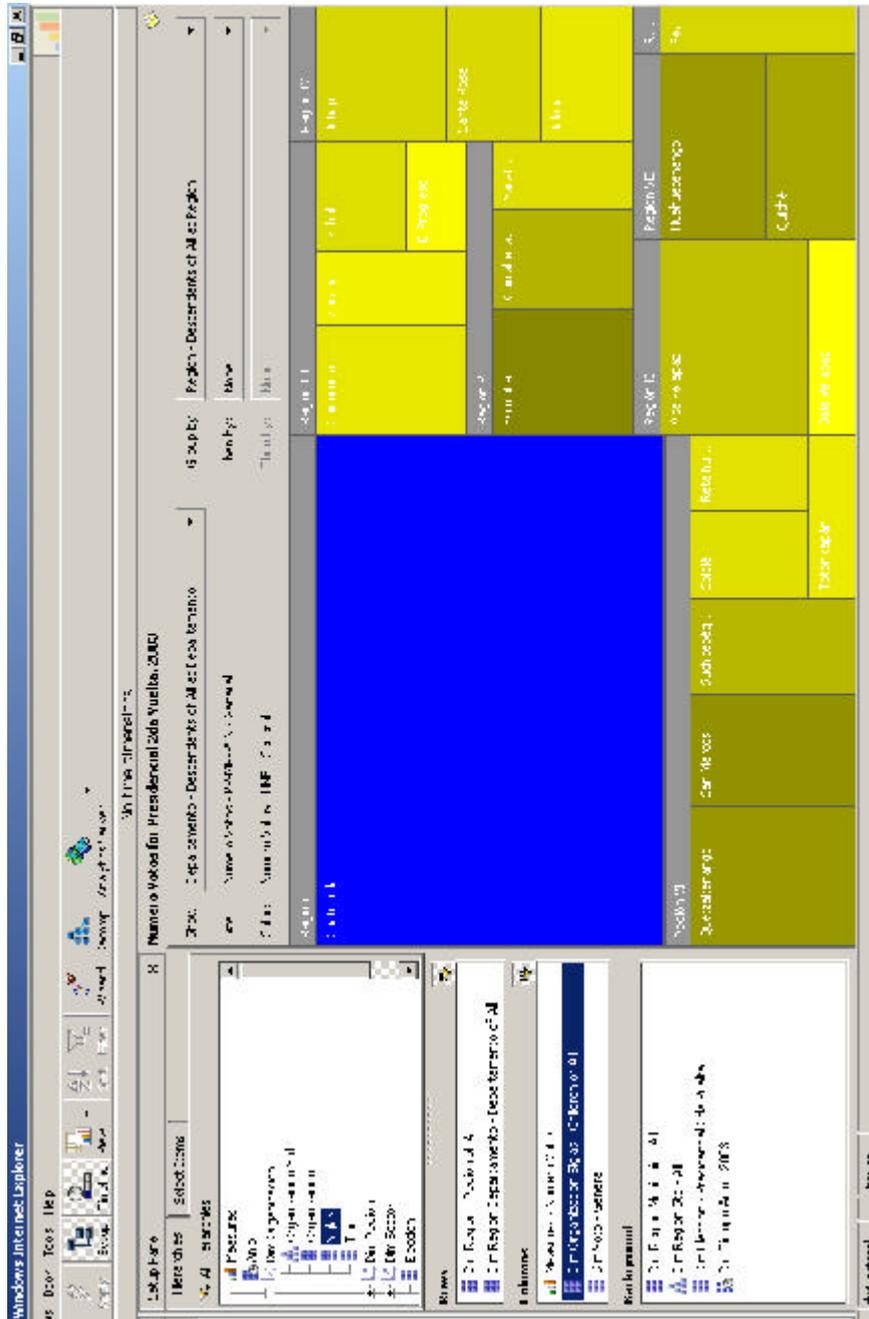
5.2.4. Votos emitidos por organización política en la 2da vuelta presidencial del año 2003.

Figura 114 Votos por Partido Político en la 2da vuelta presidencial 2003, por región y departamento



Puede realizarse un mapa de comportamiento con el fin de comparar cuánto abarca del departamento cada partido político o comité cívico.

Figura 115. Mapa de comportamiento por organización política, departamento y región



CONCLUSIONES

- 1. El apoyo de la investigación a través de herramientas de inteligencia de negocios, permite que, el análisis de comportamiento e identificación de patrones, soportado por las plataformas de BI, sea utilizado para la realización de crítica constructiva a las entidades electorales, con el fin de fortalecer la democracia guatemalteca.**
- 2. El empleo de plataformas de BI para la interpretación de datos se constituye en un elemento que proporciona los fundamentos para el planteamiento de alternativas de organización y representación democrática.**
- 3. Se hace necesario fomentar la utilización de otras tecnologías de IT para la automatización de las instituciones electorales y los procesos electorales venciendo las barreras de resistencia al cambio y sobre todo eliminar de este modo la burocracia instaurada en la mayoría de instituciones del Estado.**
- 4. Las plataformas de BI se constituyen en alternativas viables para la predicción de cualquier área del conocimiento y abarcando todo tipo de institución sin quedar limitado a soluciones comerciales de mercado.**
- 5. Es de vital importancia la generación de un modelo estructural lógico de datos que soporte la carga de los datos electorales y que permita la generación del cubo OLAP de información.**

6. **A pesar que existen muchas plataformas de BI, actualmente en el mercado, la selección debe basarse en aspectos como integración, facilidad de uso, escalabilidad, arquitectura web, sencillez de administración y que cumpla con los requisitos específicos relacionas con la bodega de datos que se construyó.**

7. **El uso de Proclarity como herramienta para creación de reportes e interpretación de los datos resultó en una experiencia simple y práctica para los usuarios finales del sistema de BI.**

RECOMENDACIONES

- 1. Utilizar alguna herramienta de datamining para la identificación de tendencias, construcción de modelos analíticos más complejos que permitan establecer las causas de comportamientos ciudadanos participativos, así como predecir el desarrollo de los subsiguientes procesos electorales.**
- 2. Establecer una vía de comunicación con el Tribunal Supremo Electoral para contar con la información electoral generada por cada nuevo plebiscito.**
- 3. Agregar a la bodega de datos electoral en una nueva fact las métricas correspondientes a número de afiliados por partido político, para hacer otros análisis enfocados a la dispersión del voto y diversidad electoral.**
- 4. Crear un nuevo cubo de información de información complementaria, que permita la correlación de datos electorales de votos y organizaciones políticas contra indicadores poblacionales.**

REFERENCIAS BIBLIOGRÁFICAS

1. Luis Fernando Mack, Paola Loaiza Ortiz, María Alejandra Erazo y otros. **Cuadernos de Información Política (Volúmenes del 1 al 15)**. Guatemala: Facultad Latinoamericana de Ciencias Sociales, 2006.
2. James A. Senn. **Análisis y diseño de sistemas de información**. Estados Unidos: Mc Graw Hill, 2000. 962pp.
3. Richard A. Johnson, Fremon E. Kast, James E. Rosenzweig. **Teoría, Integración y Administración de Sistemas**. Estados Unidos: Mc Graw Hill, 1996. 399pp.
4. Ralph M. Stair, George W. Reynolds. **Principios de sistemas de información: Enfoque Administrativo**. México: Cengage Learning, 2000. 692pp.
5. Carmen de Pablos, Jose Joaquín López, Sonia Medina. **Informática y comunicaciones en la empresa**. España: Universidad Rey Juan Carlos, 2004. 316pp.
6. John M. Ivancevich, Peter Lorenzi, Steven J. Skinner, Philip B. Crosby. **Gestión, calidad y competitividad**. Estados Unidos: Mc Graw Hill, 1997. 800pp.
7. Henry Mintzberg, Jose Nicolau Medina. **Mintzberg y la dirección**. España: Ediciones Díaz de Santos, 1991. 494pp.
8. Henry Mintzberg, James Brian Quinn, John Voyer. **El proceso estratégico: Conceptos, contextos y casos**. España: Pearson Education, 1997. 688pp.
9. Vicenç Fernández Alarcón. **Desarrollo de sistemas de información: Una metodología basada en el modelado**. Catalunya: Ediciones UPC, 2006. 219pp.

10. Larry Downes, Chunka Mui, Adriana Oklander. **Estrategias digitales para dominar el mercado: Una guía para convertir a las aplicaciones asesinas en un recurso de avanzada.** Argentina: Ediciones Granica SA, 1999. 303pp.
11. Joseph H. Boyett, Jimmie T. Boyett. **Hablan los gurús: Las mejores ideas de los máximos pensadores de la administración** España: Editorial Norma, 1999. 381pp.
12. Fernando Giner de la Fuente, Maria de los Angeles Gil Estallo. **Los sistemas de información en la sociedad del conocimiento.** España: Editorial ESIC, 2004. 215pp.
13. C. J. Date, Sergio Luis María Ruiz Faudón, Felipe López Gamino. **Introducción a los sistemas de bases de datos.** España: Pearson Education, 2001. 936pp.
14. Peter Rob Carlos Coronel. **Sistemas de bases de datos: Diseño, implementación y administración.** México: Cengage Learning Editores, 2004. 838pp.
15. Jesus M^a Minguet, Tim Read. **Informática Fundamental.** España: Editorial Ramos Areces. 398pp.
16. Francisco de la Torre. **Individuo y sociedad.** México: Editorial El Progreso.
17. Octavio Uña Juárez, Alfredo Hernández Sánchez, José Manuel Prado Antúnez. **Diccionario de Sociología** España: ESIC Editorial, 2004. 1675pp.
18. Leandro Valdes. **Sistemas Electorales y de Partidos.** México: Instituto Federal Electoral, 2001.
19. J. Valles y A. Bosch. **Sistemas electorales y gobiernos representativos.** España: Editorial Ariel, 1997. 210pp

20. Centro de Asesoría y Promoción Electoral (CAPEL) IIDH. **Diccionario Electoral Tomos I y II**. 2000
21. Philip Schlesinger. **Los intelectuales en la sociedad de la información**. España: Editorial Anthropos, 1987. 299pp
22. José Sanmartín Esplugues, Stephen H. Cutcliffe, Steven L. Goldman, Manuel Medina. **Estudios sobre sociedad y tecnología**. España: Editorial Anthropos, 1992. 334pp
23. Antonio Camou. **Los desafíos de la gobernabilidad**. México: Facultad Latinoamericana de Ciencias Sociales, 2001. 338pp
24. Ricardo Monge González, Cindy Alfaro Azofeifa, José Alfaro Chamberlain. **TICs en las PYMES de Centroamérica**. Canadá: Editorial IDRC, 2005. 268pp.
25. Carlos Roberto Brys. **Plan Estratégico para el Gobierno Electrónico de la Provincia de Misiones**. Argentina: Editorial Universitaria de Misiones.
26. Josep Villalta, Eduard Pallejá. **Universidades y desarrollo territorial en la sociedad del conocimiento**. Catalunya: Editorial Universitaria Politècnica de Catalunya, 2002. 297pp
27. Manuel Peralta. **Sistemas de Información**. <http://www.monografias.com/trabajos7/sisinf/sisinf.shtml>. (Marzo 2008).
28. Norma Ortiz. **Sistemas de Información de la Empresa**. <http://www.slideshare.net/normyser/sistemas-de-informacin-en-la-empresa-cap-2>. (Marzo 2008)
29. Janhil Aurora Trejo Martínez. **Bases de Datos**. Universidad Autónoma de Nuevo León. <http://www.monografias.com/trabajos11/basda/basda.shtml>. (Marzo 2008).

30. Kioskea. **Niveles de Datos.** <http://es.kioskea.net/bdd/bddansi.php3>. (Abril 2008)
31. Sinnexus. **Business Intelligence.** <http://www.sinnexus.com/business%5Fintelligence>. (Abril 2008)
32. iWorld. **La empresa multidimensional: OLAP.** <http://www.idg.es/iWorld/impart.asp?id=143456>. (Abril 2008)
33. Dataprix. **Olap, Molap y Rolap.** <http://www.dataprix.com/olap-rolap-molap#Introducción>. (Abril 2008)
34. Gartner. **Magic Quadrant for Business Intelligence Platforms.** [http://www.telefonica.net/web2/todobi/Pixfolder/magic quadrant for business intelligence int Q106.pdf](http://www.telefonica.net/web2/todobi/Pixfolder/magic%20quadrant%20for%20business%20intelligence%20int%20Q106.pdf). (Abril 2008)
35. Usama Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth. **From Data Mining to Knowledge Discovery in Databases.** <http://www.daedalus.es/fileadmin/daedalus/doc/MineriaDeDatos/fayyad96.pdf>. (Mayo 2008)
36. William Frawley, Gregory Piatetsky-Shapiro, Christopher Matheus. **Knowledge Discovery in Databases: An Overview.** <http://www.daedalus.es/fileadmin/daedalus/doc/MineriaDeDatos/frawley92.pdf>. (Mayo 2008)
37. Sofia Vallejos. **Minería de Datos.** Facultad de Ciencias Exactas y Agrimensura. http://exa.unne.edu.ar/depar/areas/informatica/SistemasOperativos/Mineria_Datos_Vallejos.pdf. (Mayo 2008)
38. Walter Fritz. **El concepto de una sociedad.** <http://www.intelligent-systems.com.ar/intsys/concsocSp.htm>. (Junio 2008)
39. Talcott Parsons. **Estructuralismo Funcional.** <http://es.geocities.com/diazbandres/hwct/resumen2.html>. (Abril 2008)

40. **La filosofía de Aristóteles.** http://www.webdianoia.com/aristoteles/aristoteles_polis.html. (Abril 2008)
41. **Formas de gobierno y de Estado.** <http://pdf.rincondelvago.com/formas-de-gobierno-y-de-estado.html>. (Junio 2008)
42. **Características de la democracia.** <http://www.america.gov/st/democracy-spanish/2008/September/20080915170033pii0.5193445.html>. (Junio 2008)
43. **Hegel y su teoría del Estado.** <http://hegel.pais-global.com.ar/index.php/3105>. (Abril 2008)
44. **Cumbre mundial sobre la sociedad de la información.** <http://www.yachay.com.pe/cumbre/doc/doc8.pdf> (Julio 2008)
45. Octavio Islas, Fernando Gutiérrez. **La sociedad de la información como utopía.** <http://www.cem.itesm.mx/dacs/publicaciones/logos/espejo/2004/mayo3.html>. (Julio 2008)
46. **La sociedad de la información como estrategia, económica, política y social: una nueva era.** [http://www.euroresidentes.com/internet/sociedad de la informacion/sociedad informacion.htm](http://www.euroresidentes.com/internet/sociedad%20de%20la%20informacion/sociedad%20informacion.htm) (Julio 2008)
47. **Sociedad de la Información/Sociedad del conocimiento.** [http://www.insumisos.com/lecturasinsumisas/Sociedad de la informacion y del conocimiento.pdf](http://www.insumisos.com/lecturasinsumisas/Sociedad%20de%20la%20informacion%20y%20del%20conocimiento.pdf) (Agosto 2008)
48. Adrián de la Rosa Nolasco. **Las nuevas tecnologías de la información y comunicación.** <http://www.correodelmaestro.com/anteriores/2006/mayo/1anteaula120.htm>. (Agosto 2008)

49. **Concepto de la Sociedad de la Información.** http://www.telefonica.es/sociedaddeinformacion/pdf/informes/espana_2001/parte1_1.pdf. (Agosto 2008)
50. Luis Jesús Padrón Arredondo. **Las nuevas tecnologías de la información y las comunicaciones (NTIC) en la formación del hombre nuevo.** <http://www.monografias.com/trabajos23/nuevas-tecnologias/nuevas-tecnologias.shtml>. (Agosto 2008)
51. José Huidobro. **Tecnologías de información y comunicación.** <http://www.monografias.com/trabajos37/tecnologias-comunicacion/tecnologias-comunicacion.shtml>. (Agosto 2008)
52. **La sociedad de la información en España 2007.** http://e-libros.fundacion.telefonica.com/sie07/aplicacion_sie/ParteA/pdf/SIE_2007.pdf. (Agosto 2008)

BIBLIOGRAFÍA

1. Ludwig von Bertalanffy. **Teoría General de Sistemas**. México: Fondo de Cultura Económica, 1976. 311pp
2. Dr. Ravi Kalakota, Marcia Robinson. **Del e-Commerce al e-Business, el siguiente paso**. México: Pearson Education, 2001. 400pp
3. Gabriela Aguilera Peralta, Jorge Solares y otros. **Los problemas de la democracia**. Guatemala: Facultad Latinoamericana de Ciencias Sociales, 1992.
4. **Ley Electoral y de Partidos Políticos y su Reglamento**. Decreto 1-85 de Asamblea Nacional Constituyente y sus Reformas.
5. **Acuerdos de Paz** 3ra Ed. Guatemala: Universidad Rafael Landívar, 1998.
6. Manuel Camacho Solís. **Cambio sin ruptura** México: Alianza Editorial, 1994.
7. Mario Roberto Reyes Marroquín, Pablo Augussto Rosales Tejada. Trabajo de Graduación de Ingeniería en Ciencias y Sistemas: Desarrollo de un Datamart de Información Académica de Estudiantes de la escuela de Ciencias y Sistemas de la Facultad de Ingeniería de la USAC. Guatemala, Universidad de San Carlos de Guatemala, agosto 2007. 127pp
8. Oscar René Alvarado Palencia. Trabajo de Graduación de Ingeniería en Ciencias y Sistemas: Fundamentos para el diseño de “Datawarehouses” en el ambiente de Microsoft SQL Server 7.0 y su integración a la inteligencia empresarial de Cognos. Guatemala, Universidad de San Carlos de Guatemala, septiembre 2000. 132pp.

9. Elder Omidio Sandoval Molina. Trabajo de Graduación de Ingeniería en Ciencias y Sistemas: Diseño del Datawarehouse. Guatemala, Universidad de San Carlos de Guatemala, agosto 2000. 98pp.
10. Usama Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth. **From Data Mining to Knowledge Discovery in Databases**. AI Magazine, 1996. 54pp.
11. William J. Frawley, Gregory Piatetsky-Shapiro y Christopher J. Matheus. **Knowledge Discovery in Databases: An Overview**. AI Magazine, 1992. 70pp.
12. Jan Zytkow, Mohamed Quafafou. **Principles of Data Mining and Knowledge Discovery: Second European Symposium, PKDD '98**. Springer, 1998. 482pp.
13. Joy Mundy, Warren Thornthwaite, Ralph Kimball. **The Microsoft Data Warehouse Toolkit: With SQL Server 2005 and the Microsoft Business Intelligence Toolset**. John Wiley and Sons, 2006. 746pp.
14. Sholom Weiss, Nitin Indurkha. **Data Mining: A Practical Guide**. Morgan Kaufmann, 1998. 228pp.
15. Jiawei Han, Micheline Kamber. **Data Mining: Concept and Techniques**. Morgan Kaufmann, 2006. 770pp.
16. D. J. Hand, Heikki Mannila, Padhraic Smyth. **Principles of Data Mining**. MIT Press, 2001. 546pp.
17. María Elena Simón Rodríguez. **Democracia vital: Mujeres y hombres hacia la plena ciudadanía**. España: Editorial Narcea, 1999. 222pp.
18. Douglas Lummis. **Democracia radical**. España: Editorial Siglo XXI, 2002. 224pp.

19. Niklas Luhmann. **El derecho de la sociedad** México: Universidad Iberoamericana, 2002. 673pp
20. Isabel Antúnez Pérez. **Las TICs en la investigación universitaria de la comunicación no verbal.** Universidad de Sevilla, 2006. 10pp
21. Michael D. Soo, Christian S. Jensen, Richard T. Snodgrass. **An Architectural Framework.** <http://www.cs.aau.dk/~csj/Thesis/pdf/chapter20.pdf>. (Abril 2008)
22. Data and Knowledge Engineering Group. **Module 1: Database Systems.** <http://www.itee.uq.edu.au/~infs1200/Lectures/Module1.pdf>. (Abril 2008)
23. Estrategia Magazine. **Sistemas de Información y Tecnología de la Información.** <http://www.estrategiamagazine.com/tecnologia/sistema-de-informacion-y-tecnologia-de-la-informacion-ciencia-tecnica-it-is-ti-ntic-que-es-un-sistema-de-informacion-que-es-tecnologia-definicion-que-es-informatica-definicion/>. (Abril 2008)
24. IBM. **Managing Total Cost of Ownership for Business Intelligence.** http://www.sun.com/third-party/srsc/resources/iway/IBI_TCO_white_paper.pdf. (Julio 2008)
25. ActionAid. **Diálogos sobre pobreza y desigualdad.** http://www.actionaid.org/assets/pdf%5Cguatemala_esp.pdf. (Abril 2008)
26. José Luis Simon. **Introducción a la Sociología.** <http://www.une.edu.py/maestriacs/sociologiaprg.pdf>. (Abril 2008)
27. Ignacio Lago, José Ramón Montero. **“Todavía no sé quiénes, pero ganaremos”: manipulación del sistema electoral español.** Universidad Autónoma de Madrid, 2005. http://portal.uam.es/portal/page/portal/UAM_ORGANIZATIVO/Departamentos/CienciaPoliticaRelacionesInternacionales/publicaciones_en_red/working_papers/archivos/45_2005.pdf. (Junio 2008)

28. Guillermo Molinelli, Mark Jones, Sebastián Saiegh. **Sistemas electorales hacia una mejora en la representación política.** Julio 1998. <http://www.udesa.edu.ar/Faculty/Tommasi/cedi/dts/dt9.pdf>. (Julio 2008)
29. Fernando Sanchez. **Sistema Electoral y Partidos Políticos: Incentivos hacia el bipartidismo en Costa Rica.** <http://cariari.ucr.ac.cr/~anuario/fsanchez.pdf>. (Julio 2008)
30. **Sistema Electoral y Sistema Político en Centroamérica, Panamá y República Dominicana.** <http://www.jce.do/web/Portals/0/publicaciones/ensayos/Sistema Electoral y Sistema Político en Centro América, Panamá y República Dominicana.pdf>. (Junio 2008)
31. Oracle. **Lista de Precios oficiales.** http://www.oracle.com/corporate/pricing/GSA_pricelist.pdf. (Julio 2008)
32. Dafne Sabanes Plou. **Debate por una sociedad de la información inclusiva.** <http://alainet.org/active/8557&lang=es>. (Agosto 2008)

APÉNDICE

A. 1. INSTALACIÓN Y CONFIGURACIÓN DEL ENTORNO

A.1.1. SQL SERVER

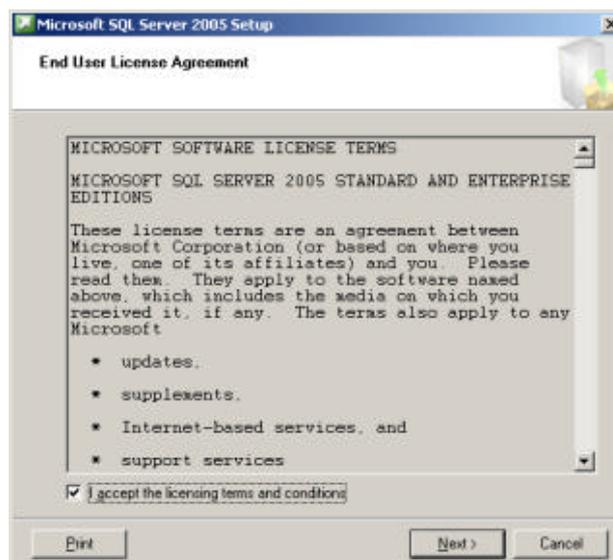
Microsoft SQL Server es la herramienta desarrollada por Microsoft Corporation para la gestión de bases de datos, basada en el lenguaje Transact SQL. El primer paso de la instalación de SQL Server 2005 es la selección de los componentes que se desean instalar, en nuestro caso será el servidor.

Figura 116. Pantalla inicial de instalación



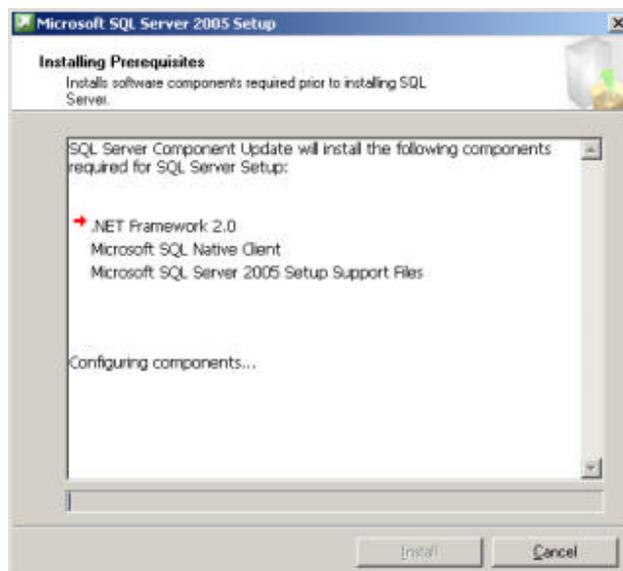
A continuación debemos aceptar el acuerdo de licenciamiento.

Figura 117. Aceptación de licencia



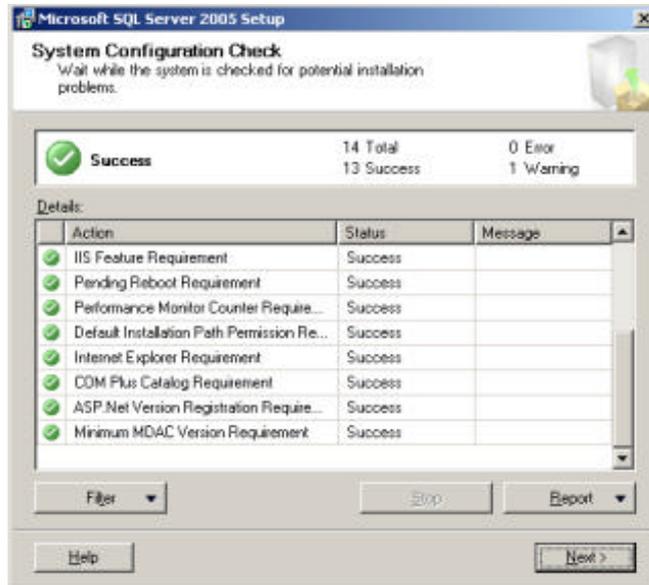
Paso seguido, el instalador se encarga de la instalación de los prerequisites, en caso de que todos se instalen correctamente debe proseguirse con la instalación.

Figura 118. Instalación de prerequisites



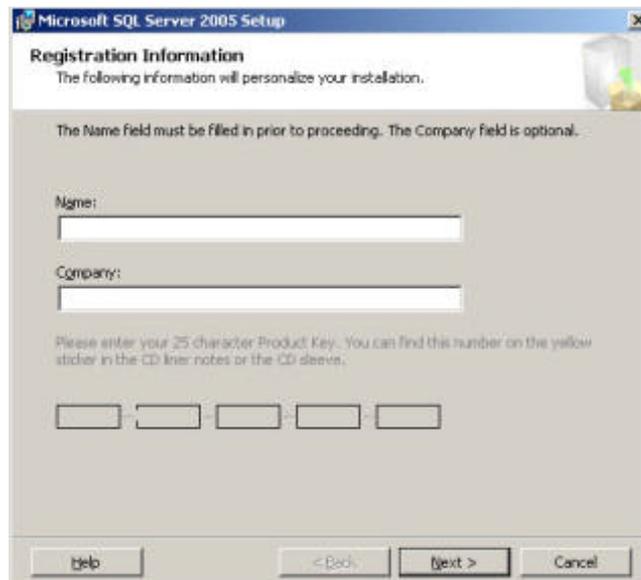
Seguidamente, se verifican todos los elementos de sistema operativo que deben estar configurados para que la instalación se realice de forma exitosa, mostrándose un resumen de los resultados obtenidos.

Figura 119 Validación de instalaciones previas



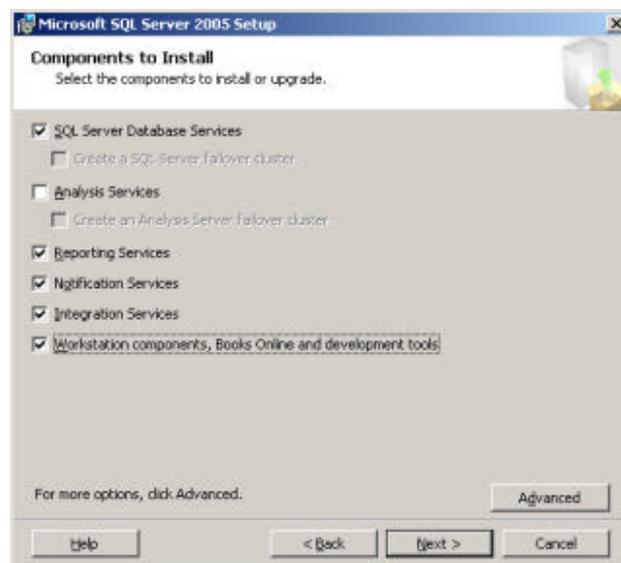
Luego deberán indicarse los valores de licenciamiento.

Figura 120 Ingreso de valores de licencia



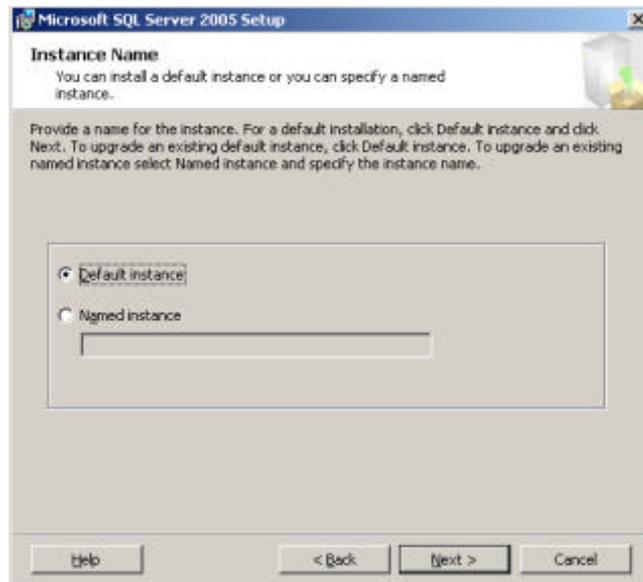
En el siguiente paso, deberán seleccionarse los elementos que se desean instalar, en nuestro caso, es importante seleccionar análisis services, integration services, reporting services y sql server, como productos mínimos de instalación.

Figura 121. Componentes de instalación



Debe indicarse posteriormente la instancia de base de datos a utilizar.

Figura 122. Instancia de base de datos



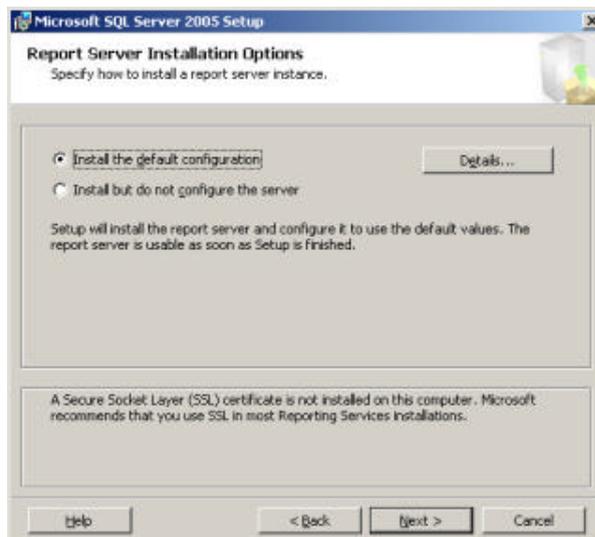
El siguiente paso corresponde al establecimiento del método de autenticación del usuario de base de datos de la instancia definida previamente. Es importante seleccionar el modo mixto, debido a que la integración con Visual Estudio requiere la utilización de usuario y clave para el establecimiento de la conexión.

Figura 123. Método de autenticación



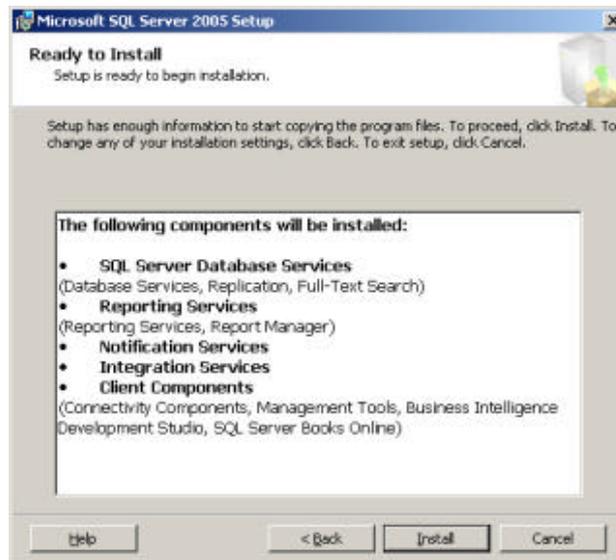
Deberá indicarse el tipo de instalación para el reporting services, en nuestro caso dejamos el valor predeterminado.

Figura 124. Opciones de Report Server



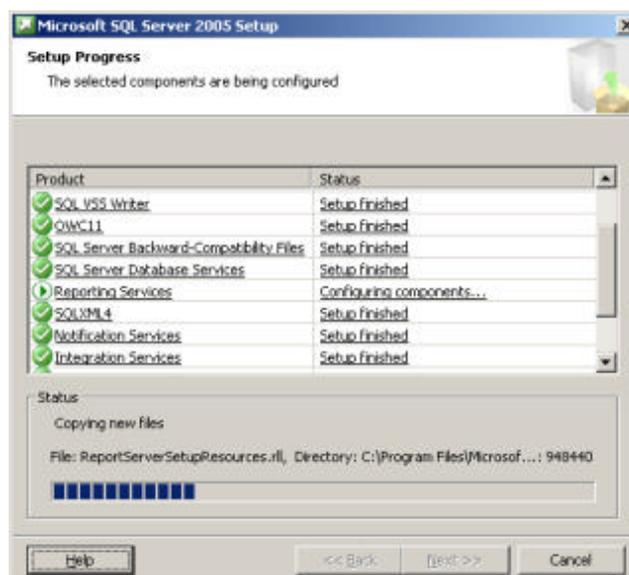
A continuación se muestra un resumen de los parámetros establecidos durante los pasos previos.

Figura 125. Parámetros de instalación



El programa de instalación se encarga de la instalación y configuración de los diferentes elementos seleccionados y todas sus dependencias, mostrando una pantalla de avances.

Figura 126. Pantalla de progreso



A.1.2. VISUAL ESTUDIO

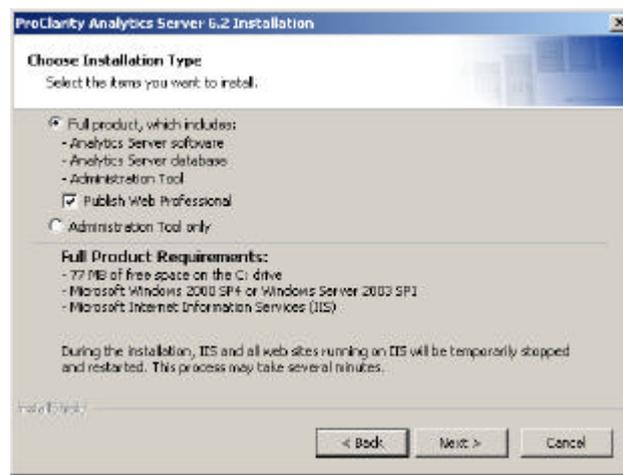
Microsoft Visual Studio es una plataforma de desarrollo para sistemas Windows, desarrollada por Microsoft Corporation. Debido a que las herramientas de Analysis Services a utilizar durante el proyecto de datawarehouse electoral son instaladas por el instalador de Sql Server 2005, no es necesaria la realización de ninguna instalación o configuración adicional para el Visual Studio.

A.1.3. PROCLARITY

Proclarity es la plataforma analítica de datos multidimensionales de Microsoft Corporation. Para la instalación de Proclarity, primeramente es necesario utilizar la opción de Install Analytics Server, que es la correspondiente a la configuración de servidores, tal como necesitamos en nuestro caso.

Seguidamente se define el tipo de instalación a utilizar, en nuestro caso se trata de instalación completa, tanto de los elementos administrativos como los correspondientes software y base de datos del servidor.

Figura 129 Definición del tipo de instalación



Luego, debe indicarse el nombre del directorio virtual donde se encontrará el servidor de Proclarity (en nuestro caso se utilizó PAS), así como la dirección física del mismo dentro de la computadora, siendo parámetros necesario para la comunicación con el servidor de internet de Microsoft (IIS).

Figura 130. Definición del directorio virtual de trabajo

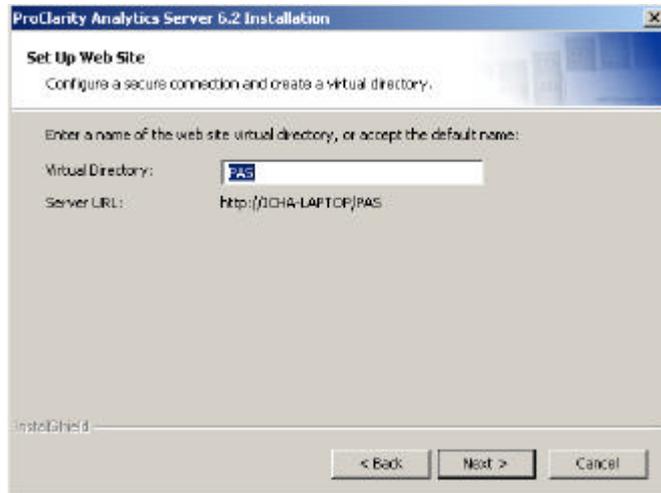
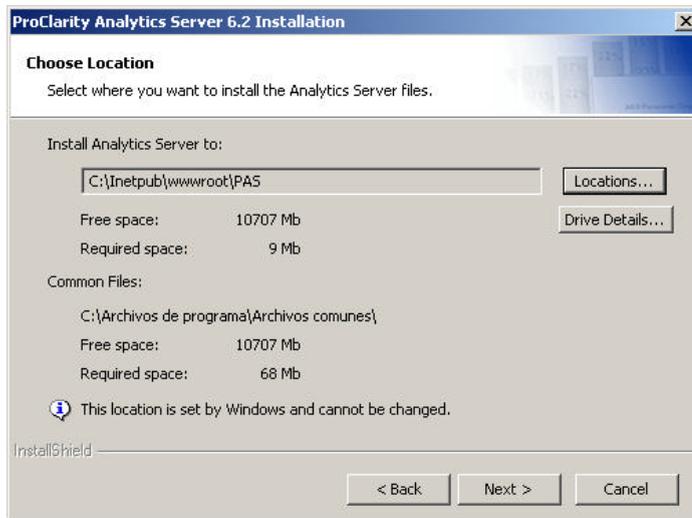
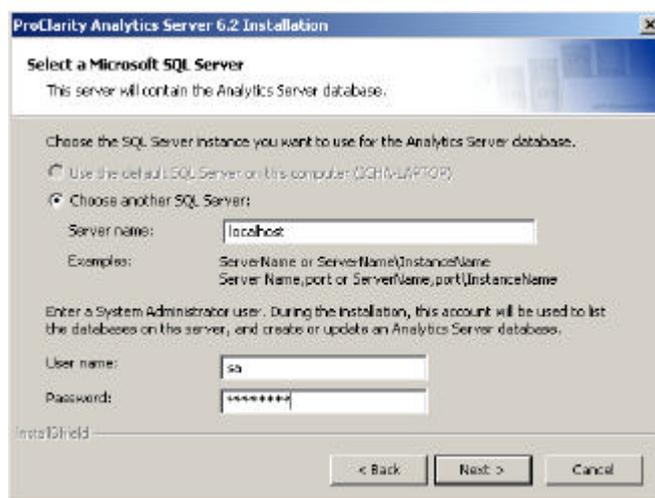


Figura 131. Localización física del servidor



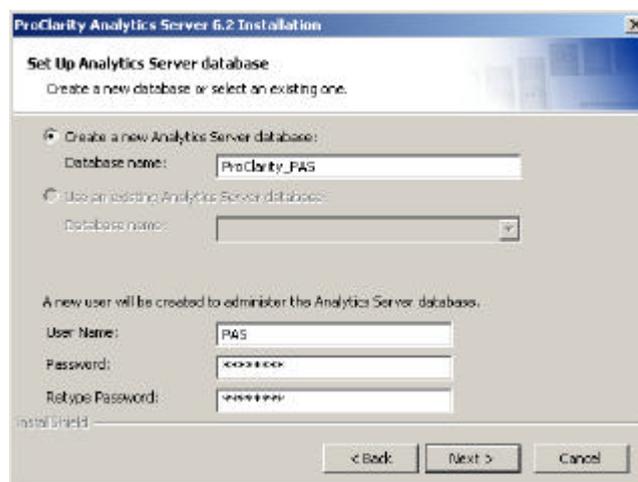
Posteriormente deben indicarse los parámetros de conexión al servidor de SQL, (nombre, ubicación, usuario y clave), que servirán para la creación del repositorio de datos del cubo lógico de Proclarity.

Figura 132. Indicar servidor de SQL



A continuación deberá configurarse el nombre del usuario administrador y la clave a utilizar para la gestión del servidor analítico de Proclarity.

Figura 133. Proporcionar usuario y clave del administrador de Proclarity



A continuación el instalador muestra un resumen de los parámetros establecidos durante el proceso y en caso de que todo esté correcto y según las opciones seleccionadas, deberá seguirse con la instalación.

Figura 134. Resumen

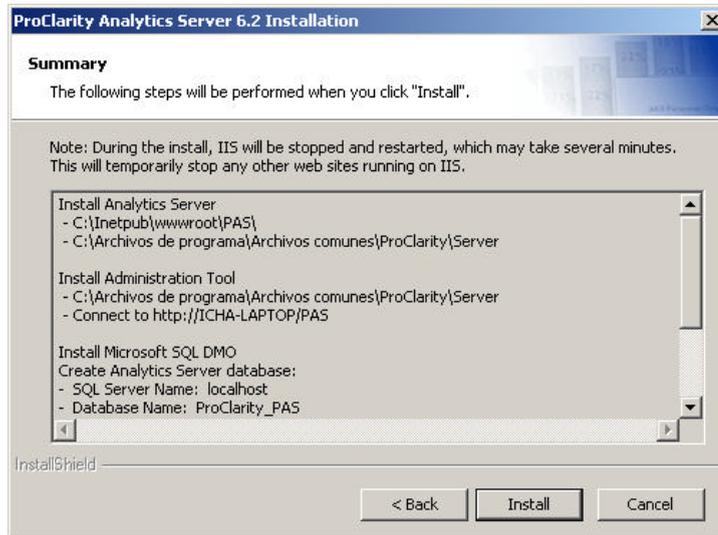
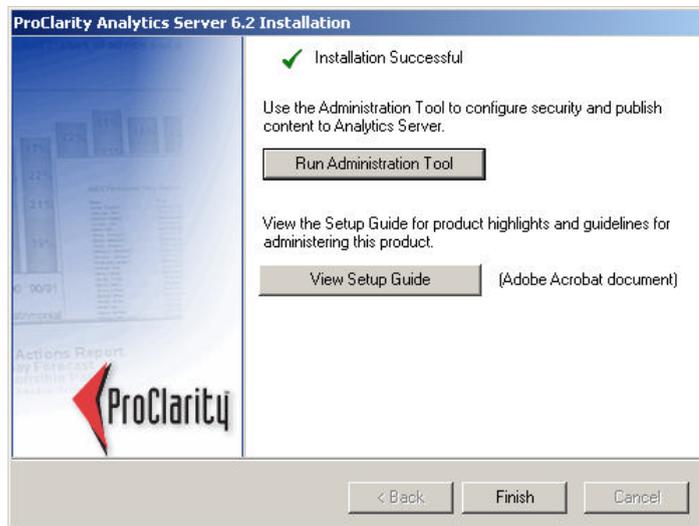


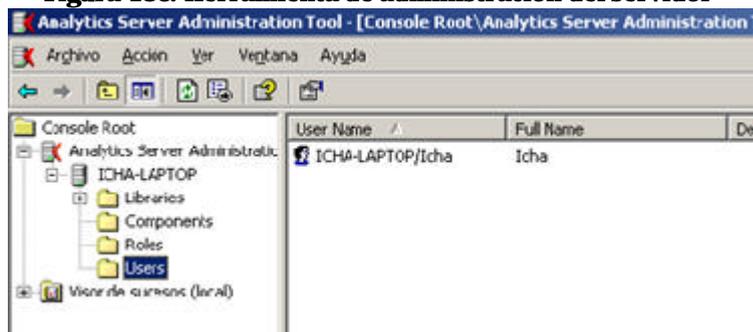
Figura 135. Finalización de la instalación



Para que los usuarios puedan consultar la información y realizar reportes personalizados, a través del explorador de Microsoft, podrán acceder al cubo de información a través de la url: <http://servername/pas>. Es importante señalar que para que el servidor de Proclarity reconozca a los usuarios, estos deben haber

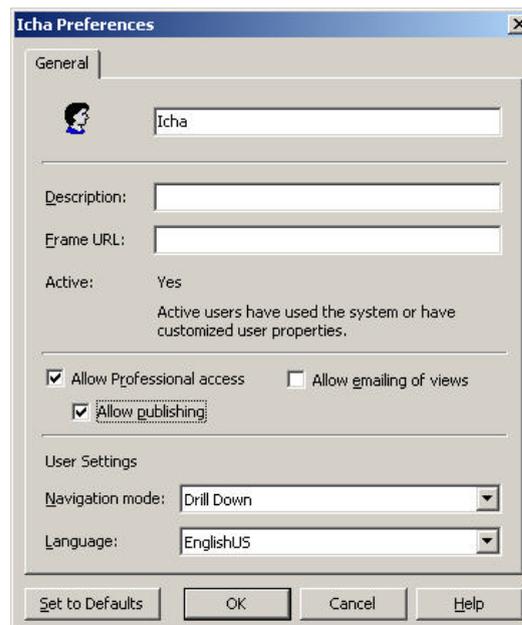
ingresado al menos una vez a la herramienta, usando el navegador. Los permisos y niveles de acceso que tendrá cada usuario podrá gestionarse a través de la herramienta de administración del servidor.

Figura 136. Herramienta de administración del servidor



Esta opción de permite establecer si los usuarios pueden utilizar la versión profesional y la publicación de vistas y reportes personalizados.

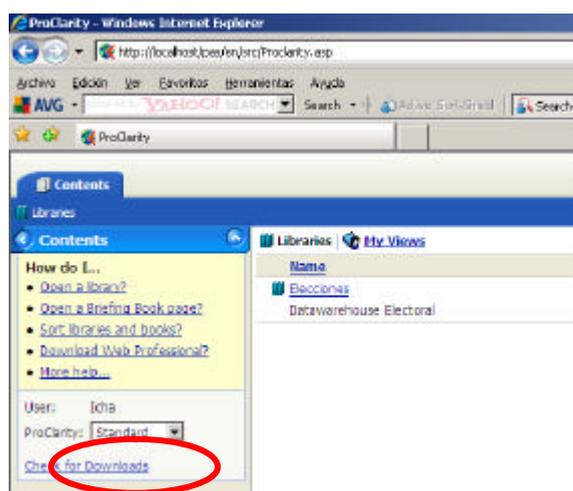
Figura 137. Asignación de permisos de usuario



En caso de que un usuario tenga permisos para la instalación del Proclarity profesional, se deberán seguir los siguientes pasos en la máquina cliente.

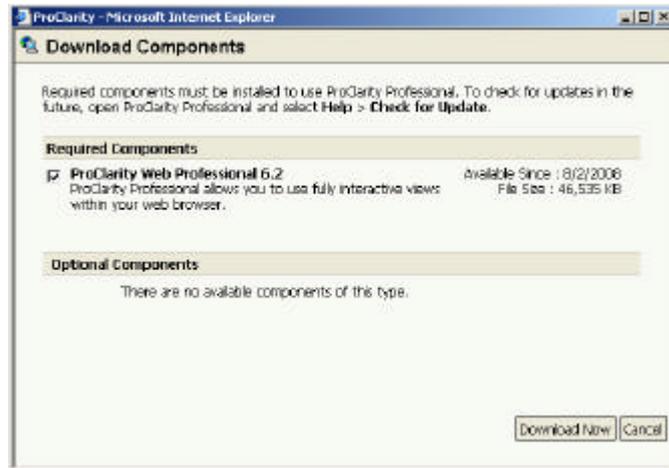
Primeramente deberá utilizarse la opción del panel izquierdo “*Check for Downloads*” la cual permite la búsqueda de actualizaciones. Esta opción también es importante en caso de que se quieran publicar parches de corrección de errores o de actualizaciones de software en el servidor y que los usuarios deberán instalar.

Figura 138 Búsqueda de actualizaciones



El paso anterior provoca que la ventana de componentes aparezca y esta es donde los usuarios deberán activar la opción de Proclarity Professional y luego “*Download Now*”.

Figura 139. Ventana de actualización de componentes



Estos pasos son suficientes para contar con la versión profesional de Proclarity, la cual podrá ser accedida a través del ícono que se encuentra en la esquina superior derecha del navegador de Proclarity Standard.

Figura 140. Icono para Proclarity Professional

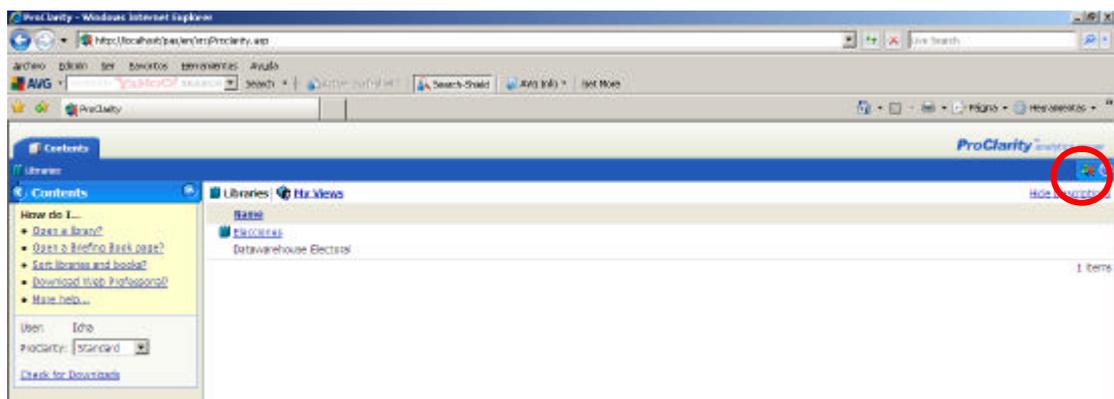


Figura 141. Entorno de Proclarity Professional

