



**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**

**FACULTAD DE INGENIERÍA EN ELECTRICIDAD Y COMPUTACIÓN**

**“DISEÑO DE UNA METODOLOGÍA PARA LA MIGRACIÓN DE DATOS  
ENTRE APLICACIONES CRÍTICAS EN LA M.I. MUNICIPALIDAD DE  
GUAYAQUIL”**

**TESIS DE GRADO**

Previa a la obtención del Título de:

**MAGISTER EN SISTEMAS DE INFORMACION GERENCIAL**

Presentado por:

**ING. JORGE RODRIGUEZ ECHEVERRIA**

**GUAYAQUIL – ECUADOR**

**2012**

## AGRADECIMIENTO

A Dios.

A todas personas que de alguna manera ayudaron a la exitosa culminación de esta Tesis.

## DEDICATORIA

A Dios.

A mi Familia.

A mis amigos.

# TRIBUNAL DE GRADO

---

**ING. LENIN FREIRE**  
COORDINADOR MSIG.

---

**ING. ANDRES NAVARRO**  
Miembro Principal

---

**PH.D. JORGE OLAYA**  
Miembro Principal

## DECLARACIÓN EXPRESA

"La responsabilidad del contenido de esta Tesis de Grado, me corresponde exclusivamente; y el patrimonio intelectual de la misma a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL"

(Reglamento de exámenes y títulos profesionales de la ESPOL)

  
ING. JORGE RÓDRIGUEZ E.

## RESUMEN

En la actualidad todas las empresas de una u otra manera intentan estar al día con las Tecnologías de Información, lo que muchas veces representa que se produzcan cambios en el entorno, arquitectura y plataforma sobre las cuales sus aplicaciones se encuentran trabajando. Esta evolución se ve reflejada en los cambios que se deben realizar tanto en las aplicaciones existentes y sus nuevos requerimientos así como en las especificaciones para la creación de nuevas aplicaciones en la empresa.

Como parte de este proceso, cuando se realizan los cambios en las aplicaciones se plantean a la vez mejoras en sus modelos conceptuales generándose muchas veces en el camino cambios en el modelo de base de datos, siendo esta un factor extremadamente importante dentro del funcionamiento de los Sistemas de Información. Debido a los cambios en el modelo de base de datos y/o cambios en plataformas es necesario realizar la migración de la información entre las bases de origen y destino, de manera que las aplicaciones evolucionen a nivel funcional junto la información vinculada a ellas.

La migración de datos a los nuevos sistemas se convierte entonces en una parte fundamental para que los Sistemas de Información puedan satisfacer las necesidades de los usuarios y estos puedan seguir utilizando sus aplicativos sin problema alguno. En base al análisis y relaciones del esquema de datos actual frente al nuevo esquema, se puede generar entonces el plan de migración de datos para que la información pueda ser trasladada de un repositorio a otro; cabe indicar que para este propósito los datos deben de cumplir ciertos requisitos de tal forma que puedan adaptarse durante la ejecución de dicho plan.

En el presente trabajo de tesis se elabora una metodología para poder llevar a cabo la migración de datos minimizando los riesgos existentes para lo cual en su primera parte se define el alcance del proyecto y la metodología a utilizarse, mientras que los demás capítulos detallan cada uno de los pasos que comprende la metodología y su aplicación en la migración de datos de un Sistema de Control de Edificaciones. Por último, se presentan las conclusiones y recomendaciones de la metodología aplicada.



# INDICE GENERAL

RESUMEN .....	VI
INDICE DE FIGURAS .....	XI
INDICE DE TABLAS.....	XIII
INDICE DE ECUACIONES .....	XIV
<b>CAPÍTULO 1 .....</b>	<b>15</b>
1. GENERALIDADES .....	15
1.1. Antecedentes.....	15
1.2. Objetivos y Justificación .....	17
1.3. Alcance del proyecto .....	18
1.4. Metodología de migración .....	18
<b>CAPÍTULO 2.....</b>	<b>21</b>
2. LEVANTAMIENTO DE INFORMACIÓN .....	21
2.1. Modelo Entidad/Relación del Sistema actual.....	21
2.2. Modelo Entidad/Relación del Nuevo Sistema.....	23
2.3. Mapeo general de tablas .....	24
2.4. Volúmenes de datos.....	26
2.5. Estrategia de migración de datos .....	27
<b>CAPÍTULO 3.....</b>	<b>34</b>



3.	ANÁLISIS DE CALIDAD .....	34
3.1.	Análisis de integridad.....	34
3.2.	Análisis de consistencia.....	42
3.3.	Análisis de nulidad.....	46
3.4.	Criterios de calidad de datos .....	47
3.5.	Métricas de calidad.....	50
3.6.	Herramienta de migración .....	51
<b>CAPÍTULO 4.....</b>		<b>54</b>
4.	TRANSFORMACIÓN Y LIMPIEZA DE DATOS.....	54
4.1.	Proceso de Extracción de datos .....	54
4.2.	Proceso de Limpieza de datos.....	56
4.3.	Proceso de Transformación de datos .....	59
4.4.	Administración de errores y revisión de información .....	62
4.5.	Proceso de carga de datos.....	64
<b>CAPÍTULO 5.....</b>		<b>67</b>
5.	PRUEBAS .....	67
5.1.	Plan de Pruebas .....	67
5.1.1.	Definición de escenarios .....	68
5.1.2.	Definición de criterios de calidad.....	70
5.2.	Pruebas unitarias de Extracción, Limpieza, Transformación y Carga.....	71
5.3.	Pruebas de volumen de Extracción, Limpieza, Transformación y Carga....	73
5.4.	Análisis y clasificación de errores .....	74

CONCLUSIONES Y RECOMENDACIONES .....	75
APÉNDICES .....	78
ANEXO A.....	79
ANEXO B.....	80
ANEXO C.....	81
ANEXO D.....	82
ANEXO E.....	83
ANEXO F.....	84
ANEXO G.....	94
ANEXO H.....	96
BIBLIOGRAFÍA .....	105

# INDICE DE FIGURAS

FIGURA 1. PASOS DE LA METODOLOGÍA DE MIGRACIÓN UTILIZADA.....	20
FIGURA 2. DIAGRAMA DE BLOQUES DEL ACTUAL MODELO E/R.....	22
FIGURA 3. FORMATO DE DOCUMENTACIÓN DE ENTIDAD DEL SISTEMA ACTUAL. ....	23
FIGURA 4. DIAGRAMA DE BLOQUES DEL NUEVO MODELO E/R.....	24
FIGURA 5. FORMATO DE DOCUMENTACIÓN DE MAPEO DE CAMPOS. ....	25
FIGURA 6. SOLICITUDES CON TIPO DE TRÁMITE NO EXISTENTE EN EL SISTEMA. ...	35
FIGURA 7. DEFINICIÓN Y DATOS DE LA TABLA EDI_TIPO_ESP DEL SISTEMA	
ACTUAL .....	36
FIGURA 8. EJEMPLO DE VALIDACIÓN DE CAMPOS REQUERIDOS. ....	37
FIGURA 9. EJEMPLO DE VALIDACIÓN DE RESTRICCIONES DE CAMPOS. ....	38
FIGURA 10. EJEMPLO DE VALIDACIÓN DE INTEGRIDAD DE ENTIDAD. ....	39
FIGURA 11. EJEMPLO DE VALIDACIÓN DE INTEGRIDAD REFERENCIAL. ....	40
FIGURA 12. EJEMPLO DE VALIDACIÓN DE DATOS DE INTEGRIDAD REFERENCIAL...	41
FIGURA 13. EJEMPLO DE VALIDACIÓN DE INTEGRIDAD DEFINIDA POR EL USUARIO.	42
FIGURA 14. ANÁLISIS DE CONSISTENCIA DE DATOS DE LA TABLA	
EDI_DATOS_SOLIC.....	44
FIGURA 15. HERRAMIENTA SQL PARA IDENTIFICACIÓN DE INCONSISTENCIAS DE	
DATOS.....	45
FIGURA 16. HERRAMIENTA PARA ANÁLISIS DE NULIDAD.....	46
FIGURA 17. RESULTADOS DE EJECUCIÓN DE HERRAMIENTA PARA ANÁLISIS DE	
NULIDAD. ....	47
FIGURA 18. FORMATO DE DOCUMENTACIÓN DE EVALUACIÓN DE MÉTRICAS.....	51

FIGURA 19. MICROSOFT INTEGRATION SERVICES.....	53
FIGURA 20. PROCESO ETL – EXTRACCIÓN DE DATOS. ....	54
FIGURA 21. PROCESO ETL – CARGA DE DATOS.....	56
FIGURA 22. PROCESO ETL – TRANSFORMACIÓN DE DATOS. ....	60
FIGURA 23. PROCESO ETL – FALLA EN TRANSFORMACIÓN DE DATOS.....	63
FIGURA 24. PROCESO ETL – CARGA DE DATOS.....	65

# INDICE DE TABLAS

TABLA 1. MATRIZ DE RELACIÓN DE DIMENSIONES DE DQ Y FUNCIONALIDADES DE LOS DATOS .....	49
TABLA 2. CRITERIOS DE CALIDAD APLICADOS AL PROCESO DE MIGRACIÓN DE DATOS .....	70

# INDICE DE ECUACIONES

ECUACIÓN 1. FUNCIÓN RATIO PARA MEDICIÓN DE CALIDAD DE DATOS. ....	50
ECUACIÓN 2. FUNCIÓN MIN PARA MEDICIÓN DE CALIDAD DE DATOS. ....	50



# CAPÍTULO 1

## 1. GENERALIDADES

### 1.1. Antecedentes

Los Sistemas de Información utilizados en una empresa, indistintamente si esta es una empresa privada o pública, tienden a sufrir cambios y mejoras durante su vida útil. Muchas veces estos cambios conllevan a la creación de nuevos sistemas que implementen la funcionalidad que actualmente se encuentra disponible y que además se haga un levantamiento de los nuevos requerimientos para satisfacer las nuevas necesidades que los usuarios tienen respecto al actual.

Cuando se ha decidido la implementación de un nuevo sistema que sustituya al existente, es necesario hacer un análisis respecto a los pros y contras de que este sea implementado bajo la misma plataforma y tecnología que el actual. En el caso de aplicaciones que han sido implementadas en viejas plataformas, la posibilidad de que el resultado de este análisis termine en el cambio a una plataforma actual es muy alta, para lo cual siempre se tendrán en cuenta los costos de la misma así como las



mejoras que deben existir tanto en la usabilidad, tiempos de respuesta y rendimiento de las nuevas aplicaciones.

Durante el análisis de los requerimientos para el nuevo sistema, se produce una evolución en el modelo conceptual original de la aplicación, lo cual puede generar automáticamente que se tenga que trabajar con un nuevo modelo de base de datos. Sin embargo, los datos almacenados por la aplicación que está en funcionamiento, se encuentran en un modelo de base de datos distinto y muy probablemente en una plataforma diferente a la que se utilizará. Por tal razón, es necesario realizar la migración entre estas bases de datos ya que la aplicación debe ser migrada junto con la información asociada a la misma.

Una vez que una de las aplicaciones es implementada y su información asociada migrada exitosamente, podemos prever que aquellas aplicaciones existentes en la vieja plataforma de la empresa con el tiempo serán migradas para lo cual es necesario poder contar con metodología de migración de datos.

## 1.2. Objetivos y Justificación

Cuando una organización decide que las aplicaciones de negocio deben evolucionar para que de esta forma pueda incrementar su rendimiento y a la vez mejorar la atención que brinda a sus clientes, esto implícitamente conlleva a que se den cambios en los modelos de datos trayendo consigo una migración de los mismos de los sistemas existentes a los nuevos sistemas. El presente trabajo de Tesis es el "Diseño de una Metodología para la Migración de Datos entre aplicaciones críticas en la M.I. Municipalidad de Guayaquil".

La justificación principal de esta Tesis es diseñar una metodología que pueda ser utilizada para la migración de datos entre las aplicaciones que actualmente se encuentran en un proceso de migración y para aquellas que en un futuro sean migradas.

El objetivo principal es establecer los lineamientos generales que permitan realizar una migración de datos de los sistemas existentes a los nuevos sistemas minimizando los problemas que se pudieran suscitar durante este proceso.

### 1.3. Alcance del proyecto

En el presente proyecto de Tesis se establecerán los lineamientos básicos de una metodología para realizar la migración de datos desde cualquier motor de base de datos hacia otro motor o una versión más reciente del que la empresa se encuentra utilizando.

Adicionalmente, la metodología descrita será aplicada durante el proceso de migración de datos del Sistema de Control de Edificaciones (SCE) de la M.I. Municipalidad de Guayaquil, el cual se encuentra implementado actualmente en Oracle 7.3.1 y se lo desea implementar en Microsoft SQL Server 2008.

### 1.4. Metodología de migración

La metodología propuesta para realizar la migración de datos consta de cuatro fases, las cuales se detallan a continuación:

#### **Fase de levantamiento de información**

En esta fase se deberá establecer el Modelo Entidad-Relación del nuevo Sistema, así como un mapeo de tablas entre dicho modelo y el modelo

existente. Por último se establecerá cual es la estrategia de migración a ser utilizada.

### **Fase de mapeo de campos**

Durante esta fase se identificarán las diferencias y características de los campos existentes en las tablas del actual Modelo E-R versus los campos en las tablas del nuevo Modelo E-R, para lo cual se hará un análisis de nulidad, integridad y consistencia de datos de tal manera que se pueda determinar la calidad de los datos y aplicar las métricas correspondientes. Como parte de esta fase se debe realizar una validación con el levantamiento de información previamente realizado.

### **Fase de transformación y limpieza de datos**

En esta fase se deberán establecer los procesos que serán utilizados para realizar la extracción, transformación, limpieza y carga de la información que se encuentran almacenada bajo la estructura del modelo actual y poder ser migrados al nuevo modelo.

## Fase de Pruebas

Como última fase de la metodología de migración se encuentra las pruebas de los procesos establecidos para la migración, para cual se deberá establecer un plan de pruebas en el cual se describan los escenarios y criterios de calidad que deben ser aplicados a los procesos de Extracción, Transformación y Limpieza de datos (ETL).

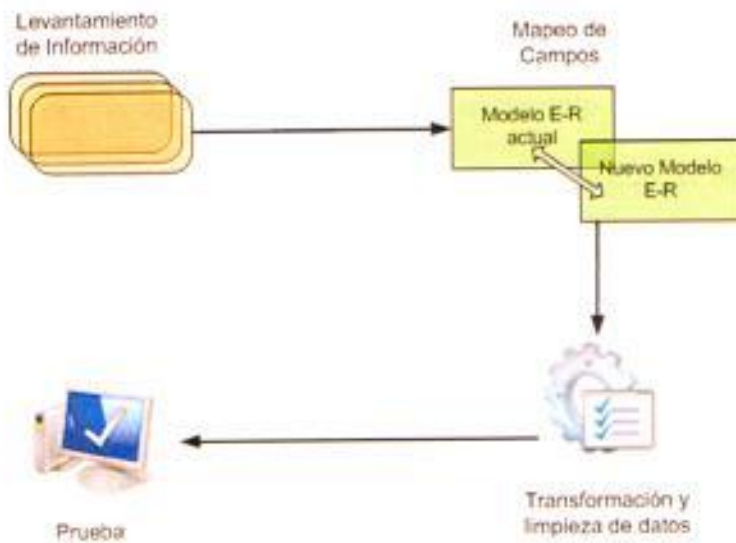


Figura 1. Pasos de la metodología de migración utilizada



## CAPÍTULO 2

### 2. LEVANTAMIENTO DE INFORMACIÓN

#### 2.1. Modelo Entidad/Relación del Sistema actual

La primera tarea a llevar a cabo es obtener el modelo Entidad/Relación de la base de datos del Sistema actual, es decir el modelo que se encuentra en el ambiente de producción de la empresa; el cual deberá ser proporcionado por el administrador de bases de datos.

A partir del momento en que se cuenta con el diagrama se procederá con el análisis todas las entidades y sus relaciones, para de esta manera determinar cuál es la función que tiene cada una y a su vez establecer el uso que se les está dando. Como resultado de este análisis se podrán:

- Identificar las tablas que contienen información repetida.
- Identificar tablas que contienen información que pertenece a otros sistemas.
- Identificar tablas que no están siendo utilizadas.
- Identificar que transacciones afectan a las tablas.

- Identificar las características que poseen los campos en cada una de las entidades para de esta manera más adelante realizar un análisis de integridad, consistencia y nulidad de los mismos.
- Identificar las entidades que almacenan la información crítica para el negocio.
- Identificar entidades que contiene información de referencia.

A continuación, se presenta un esquema de bloques de la clasificación que se realizó sobre las entidades del actual Sistema de Edificaciones (EDI). El Diagrama Entidad Relación del sistema actual se detalla en el Anexo A.

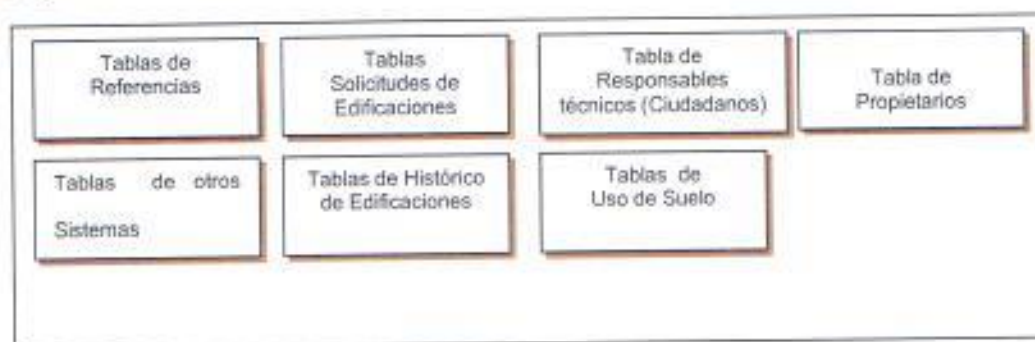


Figura 2. Diagrama de bloques del actual Modelo E/R.

Luego de que se han clasificado las entidades, se deberá identificar la función que cumplen los campos dentro de las mismas, es decir que se deberá construir el Diccionario de Dato sino se lo tiene disponible, de tal forma que esta información pueda ser utilizada durante el mapeo entre las tablas del sistema actual y del nuevo Sistema manteniéndose así una



concordancia de la información. El formato que se utilizara para documentar cada tabla del Diccionario de Datos es el siguiente:

<b>Tabla:</b>	<Nombre de tabla a documentar>		
<b>Descripción:</b>	<Descripción acerca de los datos que guarda la tabla>		
<b>Columna</b>	<b>Tipo de dato</b>	<b>Nulo</b>	<b>Descripción</b>
col1	Tipo de dato	S/N	<Descripción de lo que almacena la columna>
col2	tipo de dato	S/N	<Descripción de lo que almacena la columna>
...	tipo de dato	S/N	<Descripción de lo que almacena la columna>
colN	tipo de dato	S/N	<Descripción de lo que almacena la columna>

Figura 3. Formato de documentación de entidad del sistema actual.

En el Anexo B se encuentra documentada una de las entidades del modelo en base a la plantilla antes indicada.

## 2.2. Modelo Entidad/Relación del Nuevo Sistema

Una vez analizado el modelo E-R del sistema actual es necesario proceder con el análisis del modelo E-R del nuevo sistema. En caso de que el nuevo sistema haya sido desarrollado por terceros, se realizará la construcción de dicho modelo a partir del modelo físico proporcionado por el proveedor.

Cuando se cuente con el modelo E/R se deberán identificar cual son las entidades que se utilizan para almacenar la información bajo este nuevo

esquema, estableciendo que entidades se han mantenido, cuáles han sido creadas y cuáles han desaparecido, estas últimas ya sea porque contenían información redundante o por pertenecen a otros sistemas de la empresa. Adicionalmente, se deberá realizar un análisis de las relaciones existentes en dicho modelo y los cambios que han existido con relación al modelo del sistema actual.

A continuación, se presenta un esquema de bloques de la clasificación que se realizó sobre las entidades del nuevo Sistema de Edificaciones (EDI). El Diagrama Entidad Relación del nuevo sistema se detalla en el Anexo C.



Figura 4. Diagrama de bloques del nuevo Modelo E/R.

### 2.3. Mapeo general de tablas

Luego de haber sido analizados los Modelos E/R de cada uno de los sistemas, es necesario establecer cuál es la correspondencia entre las entidades y los campos de la Base de Datos del sistema actual frente a la Base de Datos del nuevo sistema, para de esta manera determinar:

- Campos de origen para llenar los campos en las nuevas tablas.

- Campos que cambian de tipo de dato.
- Campos numéricos que cambian de precisión.
- Valores que deben ser convertidos, como por ejemplo campos de moneda (sucres a dólares).
- Campos de secuencias que deben ser generadas.
- Campos que permitan el enlace de la información con otros sistemas.
- Campos para llenar los rastros de auditoría implementados.
- Campos que son requeridos.
- Campos que soportan valores nulos.
- Valores por defecto de los campos

Para documentar el mapeo existente entre las entidades y sus campos se hará uso de una plantilla, la cual registra toda la información antes mencionada. A continuación se presenta este modelo:

Mapeo de Tabla				
Tabla	Tabla_destino1			
Tablas Fuentes	Tabla_fuente1	Sistema origen		
	Tabla_fuente2	Sistema origen		
Campo	Origen	Campo	Valor por defecto	Unico
Campo1	--	--	Secuencia	--
Campo2	Tabla_fuente1	Campo1	--	--
Campo3	Tabla_fuente2	Campo1	--	--
...	--	--	Null	--
CampoN	--	--	Fecha sistema	--
Observaciones				

Figura 5. Formato de documentación de mapeo de campos

En el Anexo D se encuentra documentado el mapeo de campos para una de las entidades del nuevo modelo en base a la plantilla antes indicada.

## **2.4. Volúmenes de datos**

Una vez establecidas las correspondencias entre campos y diferencias entre los modelos, se deberá conocer cuál es el volumen de los datos que van a ser migrados, lo cual permitirá establecer la capacidad de recursos a nivel de hardware (espacio en discos, memoria RAM y procesador) que son requeridos por el proceso.

A su vez esto permitirá seleccionar que tipos de herramientas de software y procedimientos son los más adecuados para ser utilizados en el proceso de migración de tal forma que podemos determinar si los datos serán migrados ya sea exportando e importándolos en caso de que no requieran conversiones, haciendo uso de herramientas de ETL, a través de la ejecución de sentencias SQL entre bases o a través de procesos de copias masivas de datos.

En el Sistema de Edificaciones el volumen de datos existentes es de alrededor de un millón de registros. En el Anexo E se encuentra el detalle del volumen de datos por cada tabla del sistema actual.



## 2.5. Estrategia de migración de datos

Actualmente las organizaciones realizan la migración de datos como un proceso que es parte de la actualización de sistemas, medios de almacenamiento ó como parte del control de contingencia de datos. Por esta razón es muy importante que el impacto de la migración de datos en los procesos de negocio de la empresa sea el menor posible; es decir que la estrategia de migración utilizada debe considerar minimizar el tiempo que los sistemas están fuera de línea, problemas en integridad de datos, costos implicados en la operación, etc.

Una buena estrategia no solo puede considerar que la migración de datos sea realizada durante el tiempo que los sistemas no están siendo utilizados ya que esto incurre en implicaciones como por ejemplo el incremento de gastos por pago de horas extras, interrupción de los horarios en los cuales se realizan operaciones de respaldo de datos ó puede darse el caso de que la ventana de tiempo para realizar esta operación sea muy pequeña debido al tiempo en que los sistemas deben estar nuevamente en línea ya sea por la demanda de procesos internos o externos; en base a lo antes mencionado, la estrategia que se aplicará considera un balance entre integridad de datos, velocidad de migración, reducción del tiempo de baja del sistema y la minimización de costos.

Lo primero que se realizará es la clasificación de los datos para lo cual debemos conocer cómo y dónde estos se encuentran almacenados, así como entender las estructuras de datos en las cuales se encuentran almacenados; cabe indicar que parte de esta tarea ya se ha realizado con el análisis del modelo E-R del sistema actual. Por otro lado esta clasificación implica también que se deberá conocer las condiciones de acceso a los datos, controles y seguridades implementadas, de manera que se tengan todas las claves de acceso necesarias y también se conozca el tipo de encriptación de manejan los datos, en caso de existir.

Además, se debe conocer la forma como se encuentran distribuidos los servidores de datos en la red, el ancho de banda y las seguridades existentes en la red para tener la seguridad de que los servidores se podrán ver entre sí; esto en caso de que estemos realizando una migración entre diferentes equipos, como es el caso de la M.I. Municipalidad de Guayaquil.

Una vez que se conoce que va a ser migrado, que tiempo es el aceptado para que el sistema esté fuera de línea y cuál es el presupuesto disponible, se deben considerar los estándares relacionados a los datos que maneja la organización en caso de poseerlos. En el caso de la M.I. Municipalidad de Guayaquil se deberán tomar en consideración los servicios de auditoría

implementados en todas las bases de datos y la replicación de datos con el centro de contingencia.

Luego se realizarán ciertas consideraciones a nivel tecnológico, consideraciones como: que protocolos se necesitan para realizar el control de la migración de manera local o remota, se pueden migrar los datos directamente desde la plataforma actual en que se encuentran o es necesario realizar un paso por una plataforma más reciente y compatible con la nueva, la herramienta que se utilizará tiene soporte o no para realizar la extracción de datos desde la plataforma actual, las estructuras son idénticas o no entre la fuente el destino lo cual implicaría una transformación de datos, se deben comprobar o no los datos durante la migración, se cuenta con los respaldos de datos en la base de datos origen en caso de que exista una falla, entre otras.

La forma en que los datos serán migrados será realizada en fases:

1. Fase de migración de pre-requisitos ó tablas referenciales

Esta es la primera fase de la migración y en ella se procederán a crear y llenar con información válida todas aquellas tablas o estructuras referenciales y de pre-requisitos que son comunes a todos los sistemas



que se encuentran en la nueva plataforma hacia la cual se está realizando la migración.

La M.I. Municipalidad de Guayaquil maneja un Módulo General de Tablas comunes, en el cual se encuentran creadas tablas que contienen información común a toda la organización. A continuación, se listan las tablas del Sistema de Control de Edificaciones que serán migradas al modulo:

- Tipos de Funcionarios
- Profesiones

## 2. Fase de creación de nuevas tablas

Durante esta fase se procederá la creación de las entidades del nuevo modelo E-R del sistema. En el caso del Sistema de Control de Edificaciones, este contempla tablas definidas en el nuevo modelo que no se encuentran en el modelo actual, dichas tablas deberán ser cargadas con su correspondiente información durante la ejecución del plan de Migración.

En este paso se clasifican las tablas de la siguiente manera:

a) Tablas que nacen sin información

Estas tablas corresponden a aquellas que no existen actualmente ya que son resultado de los nuevos requerimientos del sistema.

En esta categoría se encuentra la tabla RetiroResponsable, la cual debe almacenar un tramite que actualmente se realiza manualmente y no se registra en el sistema.

b) Tablas que nacen con información

Crear todas aquellas tablas que contendrán la información referencial del nuevo sistema ya sean estas referencias entre tablas del mismo sistema o entre tablas de sistema con tablas de otros sistemas; cabe indicar que estas luego de su creación deberán ser llenadas con su correspondiente información la cual ha sido previamente validada por los usuarios. A continuación, se definen todas las tablas:

- CambioResponsable
- Funcionario
- InspeccionFinal
- Liquidacion
- Material
- ModificacionPlano
- ObraMenor
- Predio
- RegistroConstruccion
- ResponsableTecnico
- RetiroResponsable
- SceMaterial
- SrcRetiro
- Solicitante
- Solicitud
- Funcionario

- SrcAcera
- SrcEdificación
- TipoInmueble
- SubTramite
- Tramite

Las tablas que durante esta fase son creadas sin información serán pobladas una vez que la información con la cual se deben llenar haya sido validada y aprobada por los usuarios correspondientes. Esta información será transformada en sentencias SQL para la generación de los scripts que deberán ser ejecutados en el proceso de migración.

### 3. Migración de tablas existentes en la actual plataforma

En esta fase se realizará la migración de las tablas que existen en ambos modelos en base al mapeo de tablas ya realizado. Se deberán seguir los siguientes pasos:

- a) Crear las estructuras del nuevo modelo.
- b) Crear los procesos para realizar la extracción, limpieza, transformación y carga de los datos de las tablas actuales a las nuevas.
- c) Analizar los registros rechazados durante el proceso de transformación a fin de depurar la información que presente fallas u omisiones.

- d) Verificar que el mapeo sea correcto y que las características de los campos se mantienen, características como por ejemplo la precisión de campos numéricos.
- e) Verificar que el número de registros existentes en el Sistema anterior sea el mismo que el transformado y migrado al nuevo, esto con el objetivo de garantizar que se trasladó toda la información.

#### 4. Migración de históricos

En esta fase se realizará la migración de las tablas que existen en el histórico al nuevo modelo de base de datos. En el caso para el cual estamos aplicando esta metodología no existe actualmente un sistema de histórico razón por la cual se creará la estructura pero sin datos.

Al final de esta fase debemos de contar con un documento de estrategia de migración en la que se detalle el proceso antes descrito para un sistema en particular.



# CAPÍTULO 3

## 3. ANÁLISIS DE CALIDAD

### 3.1. Análisis de integridad

Una vez que se ha concluido con el levantamiento de información del sistema y se ha establecido la estrategia de migración de datos a ser utilizada, se debe proceder con la revisión de los datos que se encuentran almacenados en la base de datos del sistema actual, procediendo a realizar un análisis de la integridad el cual consiste en revisar la corrección y completitud de los datos que están almacenados.

A lo largo del ciclo de vida del sistema actual han sido ejecutadas una serie de sentencias SQL sobre los datos, ya sea para realizar una inserción (INSERT), eliminación de datos (DELETE) ó actualización de los mismos (UPDATE) las cuales han podido ocasionar que se pierda la integridad que estos presentan de modo que pueden generarse valores incorrectos, como por ejemplo solicitudes con un tipo de trámite que no existe en el sistema de edificaciones, Figura 6, ó como la ejecución de un bloque de sentencia en la cuales no se manejó transaccionalidad de datos y falló pero no fue

detectado el error en su momento habiéndose aplicado los cambios parcialmente.

```
--      edi_solicitudes
-----
-- Análisis de integridad:
-- * No tiene integridad referencial
-- * Si presenta integridad de entidad
-- * Si presenta integridad de dominio

-- Análisis de consistencia
-- * 2296 solicitudes con un trámite que no existe ('OPE') en la tabla de tramites del sistema.
select *
from   edi_solicitudes a
where  not exists (
        select *
        from edi_tramite b
        where a.tramite = b.codigo
      )
and   tramite <> 'XXX';
```

	TRAMITE	ANIO	SOLICITUD	FECHA_SYS	STATUS	USUARIO	TIPO
1	OPE	2000	1	31/08/2000 9:30:30 *		EDIFICACION	--
2	OPE	2000	2	31/08/2000 9:30:30 *		EDIFICACION	--
3	OPE	2000	3	31/08/2000 9:30:30 *		EDIFICACION	--
4	OPE	2000	4	31/08/2000 9:30:30 *		EDIFICACION	--
5	OPE	2000	5	31/08/2000 9:30:30 *		EDIFICACION	--
6	OPE	2000	6	31/08/2000 9:30:30 *		EDIFICACION	--
7	OPE	2000	7	31/08/2000 9:30:30 *		EDIFICACION	--
8	OPE	2000	8	31/08/2000 9:30:30 *		EDIFICACION	--

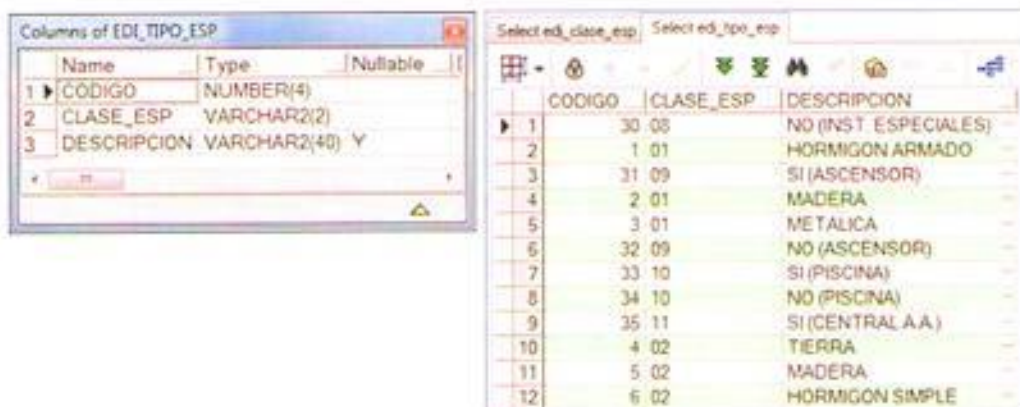
Figura 6. Solicitudes con tipo de trámite no existente en el Sistema

Durante el análisis de la integridad de los datos se identificarán los valores válidos para cada una de columnas en las tablas del nuevo sistema y se procederá a determinar la forma en cómo forzar que se mantenga dicha integridad cuando se ejecute el proceso de migración de los datos desde el sistema anterior, para lo cual se verificará que se cumplan las siguientes restricciones de integridad:

## 1. Integridad de dominio

Como parte de esta restricción se verifica que los datos con los contenidos en las columnas de las tablas del modelo actual se encuentren acorde a las definiciones establecidas en la entidad a la que pertenecen, verificándose las siguientes características:

- a. Tipo de datos; cada columna de una tabla se encuentra definida con un tipo de dato específico como por ejemplo un tipo de dato varchar2; sin embargo, es posible que la información que esta contiene corresponde a un formato de números enteros y no alfanumérico por lo tanto se debe verificar que los datos almacenados en la columna cumplan con el patrón requerido. Por ejemplo, en la Figura 7 se muestra que en el sistema actual el tipo de dato de la columna CLASE\_ESP en la tabla EDI\_TIPO\_ESP es de tipo Varchar2 pero los datos que contiene son de tipo entero.



Name	Type	Nullable
1 CODIGO	NUMBER(4)	
2 CLASE_ESP	VARCHAR2(2)	
3 DESCRIPCION	VARCHAR2(40) Y	

	CODIGO	CLASE_ESP	DESCRIPCION
1	30 08		NO (INST ESPECIALES)
2	1 01		HORMIGON ARMADO
3	31 09		SI (ASCENSOR)
4	2 01		MADERA
5	3 01		METALICA
6	32 09		NO (ASCENSOR)
7	33 10		SI (PISCINA)
8	34 10		NO (PISCINA)
9	35 11		SI (CENTRAL A.A.)
10	4 02		TIERRA
11	5 02		MADERA
12	6 02		HORMIGON SIMPLE

Figura 7. Definición y datos de la tabla EDI\_TIPO\_ESP del sistema actual



- b. Datos requeridos, si la columna ha sido definida para permitir valores nulos pero según las definiciones del sistema dicha columna no debe aceptar este tipo de valores, es necesario proceder con la verificación de los datos. Por ejemplo, en la Figura anterior se identifica que la columna DESCRIPCION en la tabla EDI\_TIPO\_ESP ha sido definida para aceptar valores nulos pero este es un catálogo cuya descripción nunca debe ser nula, razón por la cual es necesario identificar cuáles son los registros que poseen un valor nulo en este campo, Figura 8.



Figura 8. Ejemplo de validación de campos requeridos

- c. Formato, para lo cual se verifica que los datos cumplan las reglas y restricciones establecidas sobre ellos como claves foráneas (FOREING KEY), restricciones de campo (CHECK) y valores por defecto (DEFAULT). Por ejemplo, una restricción del negocio es que una solicitud puede realizar 3 reintrosos, esta información se

almacena en la columna SECUENCIA de la tabla EDI\_DATOS\_SOLIC, por lo tanto la columna antes mencionada solo puede contener valores entre 1 y 3; sin embargo, al verificar dichos valores se identificaron registros con 4 y 5 los cuales representan inconsistencias, Figura 9.

The screenshot shows a SQL query window with the following text:

```

select tramite, secuencia, count(*) inconsistencias
from edi_datos_solic
group by tramite, secuencia
having secuencia > 3
  
```

Below the query, a table displays the results:

	TRAMITE	SECUENCIA	INCONSISTENCIAS
1	SF	4	1
2	SMP	4	1
3	SRC	4	1
4	SRC	5	1

Figura 9. Ejemplo de validación de restricciones de campos

## 2. Integridad de entidad

La cual está marcada por la definición de la restricción de una clave primaria (PRIMARY KEY) o mediante un índice marcado como único (UNIQUE) en la tabla, para lo cual se verificará que los datos a ser migrados en la columna correspondiente a dicha clave o que indexada de esa forma sean únicos para cada uno de los registros. En la Figura 10, se muestra la verificación de registros únicos en la tabla EDI\_REG\_CONST.

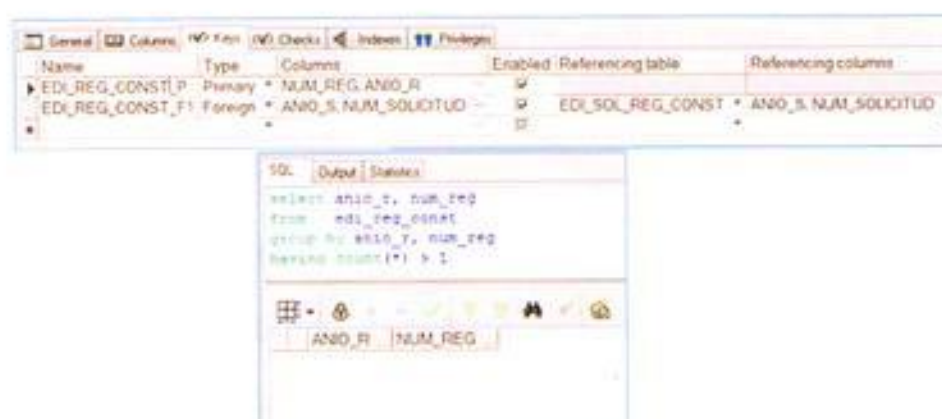


Figura 10. Ejemplo de validación de integridad de entidad

### 3. Integridad referencial

Para lo cual se deberá asegurar que los valores a ser almacenados en las columnas que son claves de las tablas que en el nuevo esquema manejan una relación tipo maestro/detalle se encuentren consistentes. Es necesaria realizar esta verificación ya que puede que las tablas tengan este tipo de relación pero no se haya implementado una integridad referencial de forma que los datos pueden haber sido corrompidos por operaciones como:

- Creación de una fila en la tabla detalle en la que la clave foránea no corresponde a la clave primaria de la tabla maestro.
- Actualización de la clave foránea del detalle que no coincide con ninguna clave primaria.

- Eliminación de una fila en la tabla maestro sin eliminar las filas en la tabla detalle, quedando filas en el detalle que no corresponden a ningún registro maestro.
- Actualización de la clave primaria en el registro maestro, pero sin actualizar la clave foránea de los registros del detalle, quedando las filas del detalle sin su correspondiente registro maestro.

Por ejemplo, al verificar este tipo de restricción se identifica que las tablas EDI\_CLASE\_ESP y EDI\_TIPO\_ESP no presentan físicamente una integridad referencial la cual está dada a través de los campos EDI\_CLASE\_ESP.CODIGO y EDI\_TIPO\_ESP.CLASE\_ESP, Figura 11. Sin embargo, no se identifican inconsistencias a nivel de datos.

EDI_CLASE_ESP	
PK	CODIGO
	DESCRIPCION

EDI_TIPO_ESP	
PK	CODIGO
	CLASE_ESP DESCRIPCION

Figura 11. Ejemplo de validación de integridad referencial

Este mismo escenario se presenta al analizar las tablas EDI\_TIPO\_OBRAM y EDI\_OBRA\_MENOR en la cual no existe la relación entre EDI\_TIPO\_OBRAM.TIPO\_OBRA y EDI\_OBRA\_MENOR.TIPO\_OBRA. En este caso se identifica como



insistencia que la tabla EDI\_TIPO\_OBRAM se encuentra vacía ocasionando que un JOIN entre las tablas antes mencionadas, no produzca ningún resultado ya que no existen elementos de coincidencia, Figura 12.

EDI_TIPO_OBRAM	
PK	TIPO_OBRA
	DESCRIPCION DESC_CORTA

EDI_OBRA_MENOR	
PK	ANIO_OBRA
PK	NUM_OBRA
	FECHA_EMISION
	ESTADO
	APE_PROPIETARIO
	NOM_PROPIETARIO
	UBICACION
	TIPO_OBRA
	CED_RUC
	OBSERVACION
	FECHA
	USUARIO
	AREA
	ANIO_SOL
	NUM_SOL
	FECHA_SOL

SQL	Output	Statistics								
<pre>select * from edi_obra_menor a,      edi_tipo_obram b where a.tipo_obra = b.tipo_obra</pre>										
										
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 25%;">ANIO_OBRA</th> <th style="width: 25%;">NUM_OBRA</th> <th style="width: 25%;">FECHA_EMISION</th> <th style="width: 25%;">ESTADO</th> </tr> </thead> <tbody> <tr> <td> </td> <td> </td> <td> </td> <td> </td> </tr> </tbody> </table>			ANIO_OBRA	NUM_OBRA	FECHA_EMISION	ESTADO				
ANIO_OBRA	NUM_OBRA	FECHA_EMISION	ESTADO							

Figura 12. Ejemplo de validación de datos de integridad referencial

#### 4. Integridad definida por el usuario

Como parte de este tipo de restricción se verificarán todas aquellas restricciones que no corresponden a las anteriores y que se encuentran relacionadas con reglas de negocio. Por ejemplo, en la Figura 13 se analiza la integridad del cálculo de la liquidación a cobrar por una Solicitud de Inspección Final, la cual según las reglas de negocio corresponde al 2%% del presupuesto total de la obra, detectándose 2778 registros inconsistentes.

```
-- 2778 registros inconsistentes con errores en el cálculo de la liquidación
select *
from   edi_solic_inspe a
where  round(a.presup_total * 0,002, 2) <> a.dos_mil;
```

	ANIO_S	NUM_SOLICITUD	TRAMITE	FECHA_ING	APE_PROFES	NOM_PROFES
1	1997	169	SIF	13/01/1997	ARQ CAMPOS BERMEO	LUIS
2	1997	270	SIF	15/01/1997	ARQ MUNOZ	VICENTE
3	1997	297	SIF	15/01/1997	ING RODRIGUEZ BALDEON	HECTOR
4	1997	76	SIF	17/01/1997	ING CORDOVA YEROVI	EDUARDO
5	1997	362	SIF	20/01/1997	ARQ AVELLANEDA ANDRADE	MARCOS
6	1997	420	SIF	20/01/1997	ING BARROS PAZMINO	PEDRO VICENTE
7	1997	455	SIF	21/01/1997	ARQ MANCHENO DE CEDENO	ROSAURA

Figura 13. Ejemplo de validación de integridad definida por el usuario.

Como resultado de este análisis se obtendrán un conjunto de valores inválidos que deberán ser corregidos previamente a la migración o que serán transformados durante la misma para de esta manera preservar la integridad de los datos almacenados en la mayor medida posible. En el Anexo F, se encuentra el resultado de este análisis aplicado al Sistema de Edificaciones.

### 3.2. Análisis de consistencia

Una vez analizada la integridad existente entre los datos almacenados en el sistema, se realizará un análisis de la consistencia de los mismos, es decir, una comparación de los valores presentes en cada campo con respecto a los valores presentes en los campos que guardan relación con los mismos.

Como parte del proceso se analiza si existe o no redundancia de datos, procediendo luego con la eliminación de todos aquellos datos redundantes que han sido identificados. Al eliminar o controlar las redundancias de datos se reduce en gran medida el riesgo de que haya inconsistencias.

De esta forma si un dato está almacenado una sola vez en una tabla, cualquier actualización se debe realizar sólo una vez en la misma, y se asegura que estará disponible para todos los usuarios inmediatamente, mientras que si un dato está duplicado y se lo utiliza para realizar consultas que manejan JOIN entre tablas producirán más registros que los que debe, lo cual degenera en más inconsistencias.

Durante este proceso, se deberá establecer qué tipos de inconsistencias no pueden ser manejadas por los procesos de limpieza y extracción, es decir que no pueden ser resueltas automáticamente y deberán ser analizadas manualmente para de esta forma corrija o elimine la información que genera la inconsistencia.

Por ejemplo, en el Sistema de Edificaciones es necesario relacionar las solicitudes con los solicitantes pero en la nueva plataforma esta relación se maneja a través de un identificador único que no existe en la plataforma



anterior, de tal manera que si existe una Solicitud de Edificaciones relacionada a un Contribuyente a través de su número de cédula y se encuentre mal escrito, al momento que el proceso intente buscar cual es el identificador que le corresponde respecto al Sistema de Ciudadanos, este dará como salida un valor nulo y no se podrá resolver esta inconsistencia automáticamente dando como resultado un rechazo en el proceso de migración de dicho registro. En la Figura 14, se muestra un análisis realizado a la tabla EDI\_DATOS\_SOLIC, en la cual se identifica que 1834 registros serán rechazados durante el proceso de migración por presentar inconsistente el número de cédula del Contribuyente.

```
-- * 1834 registros inconsistentes, longitud cedulas de solicitantes menores a 8 dígitos
select *
from edi_datos_solis a
where length(ISO_char(cedula_sol)) < 8;
```

	TRAMITE	ANIO	NUM_SOLICITUD	SECUENCIA	APELLIDO_SOL	NOMBRE_SOL
1	SRC	1996	306	1	CONST. INMOB. ARGOS C.A.	--
2	SRC	1996	307	1	CONST. INMOB. ARGOS C.A.	--
3	SRC	1996	2257	1	CIA. UNYSIS	--
4	SNE	1996	2276	1	ING. ALBARRACIN PENARANDA	EZEQUIEL
5	SNE	1996	2344	1	ARD. BETANCOURT MORAN	FRANCISCO
6	SNE	1996	2243	1	ARD. GALLEGOS	WILSON
7	SRC	1996	1079	1	HUAYAMABE SALVATIERRA	ENRIQUE
8	SRC	1996	128	1	GUERRERO SANCHEZ	LUIS O.
9	SRC	1996	81	1	TUMBACO HOLGUN	VICENTA PIEDAD
10	SRC	1996	138	1	TE-R-CE ECUATORIANA S.A.	ELMAR SCHULZ MAASS

Figura 14. Análisis de consistencia de datos de la tabla EDI\_DATOS\_SOLIC

Por este motivo se desarrollarán herramientas SQL (Figura 15) para identificar tanto las redundancias de datos en tablas como también aquellas inconsistencias entre valores almacenados.

```

-----
--:          edi_sol_reg_const
-----
-- Análisis de integridad:
-- * No presenta Tiene integridad referencial, no se relaciona con tabla edi_subzone, edi_tipo_uso
-- * Si presenta integridad de entidad
-- * Si presenta integridad de dominio

-- 409 registros inconsistentes, campo cédula menor a 8 dígitos
select *
from   edi_sol_reg_const a
where  length(to_char(a.ced_profes)) < 8

-- 162 registros inconsistentes, campo cédula nulo
select *
from   edi_sol_reg_const a
where  a.ced_profes is null;

-- 38780 registros inconsistentes, campo ape_profes contiene la profesión
select count(*)
from   edi_sol_reg_const a
where  a.ape_profes like 'ARQ.%' or a.ape_profes like 'ING.%';

-- 54 registros inconsistentes, campo reg_profes nulo
select count(*)
from   edi_sol_reg_const a
where  a.reg_profes is null;

-- 28530 registros inconsistentes, no se encuentran zonas o subzonas específicas
select *
from   edi_sol_reg_const a

```

Figura 15. Herramienta SQL para identificación de inconsistencias de datos.

Como resultado este análisis se obtendrá un conjunto de registros que mantienen información inconsistente y que debe ser corregida o eliminada en caso de que pueda serlo, lo cual debe de realizarse previo a la migración para de esta manera preservar la consistencia de los datos almacenados. En el Anexo F, se encuentra el resultado de este análisis aplicado al Sistema de Edificaciones.

### 3.3. Análisis de nulidad

Luego de realizar el análisis de la consistencia que existen entre los datos almacenados en la base de datos, se dará inicio a un análisis de nulidad de los datos.

Como primer paso para realizar este análisis se identificará por cada una de las tablas cuales son las columnas que contienen valores nulos. Para este propósito se deberá desarrollar una herramienta SQL (Figura 16).

```
begin
  c1 integer;
  i_inserted_rows integer;
  sql_str VARCHAR(1024);

  cursor c_tablas is
  SELECT TABLE_NAME FROM all_tables where owner = 'EMPRESACION' and TABLE_NAME <> Upper('AnálisisNulidad');

  cursor c_columnas (p_TableName in varchar2) is
  SELECT column_name FROM all_tab_columns where Table_name = p_TableName;
begin
  c1 := dbms_sql.open_cursor;
  for tabla IN c_tablas loop
    for columna IN c_columnas (tabla.TABLE_NAME) loop
      begin
        sql_str := 'insert into AnalisisNulidad('||tabla.||' columna, regNull) select ''' ||
          tabla.TABLE_NAME || ''',''' || columna.column_name || ''', count(*) regNull from ' ||
          tabla.TABLE_NAME || ' where ' || columna.column_name || ' is null';

        execute immediate sql_str;

        sql_str := 'insert into AnalisisNulidad('||tabla.||' columna, regNotNull) select ''' ||
          tabla.TABLE_NAME || ''',''' || columna.column_name || ''', count(*) regNull from ' ||
          tabla.TABLE_NAME || ' where ' || columna.column_name || ' is not null';

        dbms_sql.parse(c1, sql_str, dbms_sql.native);
        i_inserted_rows := dbms_sql.execute(c1);
      exception
        when others then
          dbms_output.put_line (tabla.TABLE_NAME);
      end;
    end loop;
  end loop;
  dbms_sql.close_cursor(c1);
  commit;
end;
```

Figura 16. Herramienta para análisis de nulidad

Como resultado se tendrá un listado de las tablas con cada una de sus columnas y cuantos registros contienen valores nulos, lo cual permite identificar columnas que nunca han sido utilizadas. A continuación, en la



Figura 17 se muestran los resultados de este análisis aplicado al Sistema de Edificaciones:

```

SQL> select tabla, columna, sum(case when nvl(ano) = 0 then 1 else 0 end) no_nulos, sum(case when nvl(ano) = 0 then 1 else 0 end) total
  2 from analisis@ltdad
  3 group by tabla, columna
  4 /

```

TABLA	COLUMNA	NO_NULOS	NULOS	TOTAL
EDI_DONA	CODIGO	17	0	17
EDI_DONA	DESCRIPCION	17	0	17
EDI_PLIEZO	ANEO_F	679	0	679
EDI_PLIEZO	ESTADO	679	0	679
EDI_PLIEZO	ANIO_SOL	679	0	679
EDI_PLIEZO	USUARIO	679	0	679
EDI_PLIEZO	FECHA_ING	679	0	679
EDI_PLIEZO	FECHA_OTR	679	0	679
EDI_PLIEZO	ACTIVIDADES	679	0	679
EDI_PLIEZO	OBSERVACION	510	169	679
EDI_PARAM	ANEO	55	0	55
EDI_PARAM	VALORO1	31	24	55
EDI_PARAM	VALORO2	31	24	55
EDI_PARAM	TRABITE	55	0	55
EDI_PARAM	TERMINAL	55	0	55
EDI_PARAM	FEI_DESDE	24	31	55
EDI_PARAM	FEI_HASTA	0	55	55
EDI_PARAM	ANIO_SOLICITUD	55	0	55
EDI_AVALUO	ANEO	27975	0	27975
EDI_AVALUO	TRABITE	27979	0	27979

Figura 17. Resultados de ejecución de herramienta para análisis de nulidad

Luego de identificar las columnas que contienen valores nulos, se ubicará en el documento de mapeo de campos cuales de estas columnas en el nuevo modelo de base de datos serán migradas a columnas que no soportan valores nulos. Estas inconsistencias deben ser corregidas llenando los campos con información previa a la ejecución del proceso de carga de datos en la nueva estructura de tal forma que no se produzcan rechazos de registros con estos campos en nulo. En el Anexo F, se encuentra el resultado de este análisis aplicado al Sistema de Edificaciones.

### 3.4. Criterios de calidad de datos

Conforme un sistema se mantiene en uso, los problemas que se presentan en la calidad de sus datos se van incrementando gradualmente hasta el

punto en que estos comienzan a ser notorios ocasionando que la información consultada por las aplicaciones comience a reflejar valores inexactos, incompletos y hasta ambiguos, los cuales pueden impactar negativamente a la organización ya sea técnicamente al existir problemas en las bases de datos, económicamente produciendo la pérdida de cliente ó pérdidas financieras y legalmente afectando a sus cliente o terceros producto que la baja calidad de sus datos la cual se refleja a manera de errores en los datos almacenados por el sistema.

Con el propósito de medir la calidad de los datos se establecerán un conjunto de dimensiones de calidad (DQ), esto con el afán de simplificar el estudio de la misma como lo propone ISO9126[1]; estas DQ permitirán establecer la calidad de los datos desde un punto de vista determinado para lo cual se definirán criterios de aceptación para cada una de ellas basándose en el punto de vista de los usuarios a los que les corresponde ejecutar cierto tipo de tareas sobre los datos.

Los criterios de aceptación que se establezcan para las DQ estarán definidos en base a los posibles rangos de medición que se puedan aplicar. Así, si luego de realizar la medición del nivel de calidad se obtienen resultados que no cumplen con los rangos de aceptación predefinidos, se considerará al dato como defectuoso y no deberá permitirse su utilización



por el proceso hasta que este sea revisado y corregido. En este punto es necesario no confundir el criterio de calidad de datos con "cero errores", ya que lo que se busca es una adecuación de los datos para su utilización en la organización.

A partir de las funciones del sistema, se realiza una clasificando de las mismas en relación con los criterios para medir la calidad de los datos. En la Tabla 1, se detallan algunas de las funciones a evaluar del Sistema de Edificaciones y como estas se relacionan con algunos de los criterios de calidad.

		<b>Criterios</b>		
<b>Funciones</b>	<b>Reglas</b>	<b>Exactitud</b>	<b>Complejitud</b>	<b>Consistencia</b>
		Solicitudes	Tasa de tramite de edificaciones Tener por lo menos un solicitante	x
Solicitantes	Solicitante debe existir en Sistema Ciudadano			x
Responsables Técnicos	El responsable técnico debe de existir en el Sistema Ciudadano		x	x
Registro de Construcción	Tener una solicitud de registro de Tener un responsable tecnico	x	x	x

**Tabla 1. Matriz de relación de Dimensiones de DQ y Funcionalidades de los datos**

### 3.5. Métricas de calidad

Una vez identificadas las funciones del sistema y su relación con los criterios de calidad, se establecerá el tipo de medida a ser utilizada. Para la metodología que se está desarrollando se utilizarán las siguientes funciones de medición:

- Ratio: basada en porcentajes, como se muestra en la Ecuación 1.

$$QD = 1 - \frac{\# \text{ de datos que satisfacen el criterio}}{\# \text{ total de datos}}$$

Ecuación 1. Función Ratio para Medición de Calidad de Datos.

Para aplicar esta fórmula y establecer la calidad de los datos se realizará un conteo de todas las unidades y otro conteo de sólo aquellas unidades de datos que satisfacen el criterio evaluado. Esta métrica será aplicada a las reglas de negocio que deben de cumplir los criterios de exactitud y consistencia.

- Operación Mínimo, como se muestra en la Ecuación 2.

$$QD = \frac{\min(\text{responsable técnico por Registro de Construcción})}{1}$$

Ecuación 2. Función Min para Medición de Calidad de Datos.

Esta métrica será aplicada a las reglas de negocio que deben de cumplir con un número mínimo de registros definidos. Por ejemplo para el sistema de Control de Edificaciones, como regla de negocio se ha definido que una solicitud debe tener mínimo un solicitante y una solicitud de registro de construcción debe tener mínimo un responsable técnico, por lo tanto el valor de aceptación es 1.

Para registrar los resultados producto de la evaluación de estos criterios se utilizará la plantilla que se muestra en la Figura 18. Los resultados de la evaluación de estos criterios aplicados a un trámite ejemplo del sistema, se encuentran en el Anexo G.

Métrica 1:	
# total de datos:	
# datos satisfactorios:	
QD Medida:	
Rango de aceptación:	

Figura 18. Formato de documentación de evaluación de métricas.

### 3.6. Herramienta de migración

En esta sección se definirán cuales son las herramientas que se utilizarán para poner en ejecución la estrategia de migración descrita en el Capítulo 2. Para esto, el primer paso a dar es identificar la plataforma bajo la cual se encuentran almacenados actualmente los datos.

En el caso del Sistema de Edificaciones, este se almacenan actualmente en una base de datos Oracle versión 7.3.4.0.1, la misma que se encuentra bajo un Sistema Operativo Solaris en un servidor Sun. La plataforma destino a cual serán migrados es una base de datos SQL Server 2008 versión Enterprise, la misma que se encuentra bajo un Sistema Operativo Windows 2008 Server con un clúster de 32 GB de memoria.

Debido a que existe una incompatibilidad entre el motor SQL 2008 y la versión que se tiene de Oracle, no es posible realizar el paso directo de los datos razón por la se utilizará como herramienta de migración Microsoft Integration Services 2008.

Integration Services es una herramienta gráfica (Figura 19) que permite realizar la limpieza, extracción, transformación de datos y posteriormente su cargar en el servidor destino. Como características de esta herramienta se tienen:

- Flexibilidad en la producción de reportes de elementos de datos y reglas.
- Generación de migración de código o scripts directamente a partir de reglas de mapeo.



- Detección de violaciones de integridad

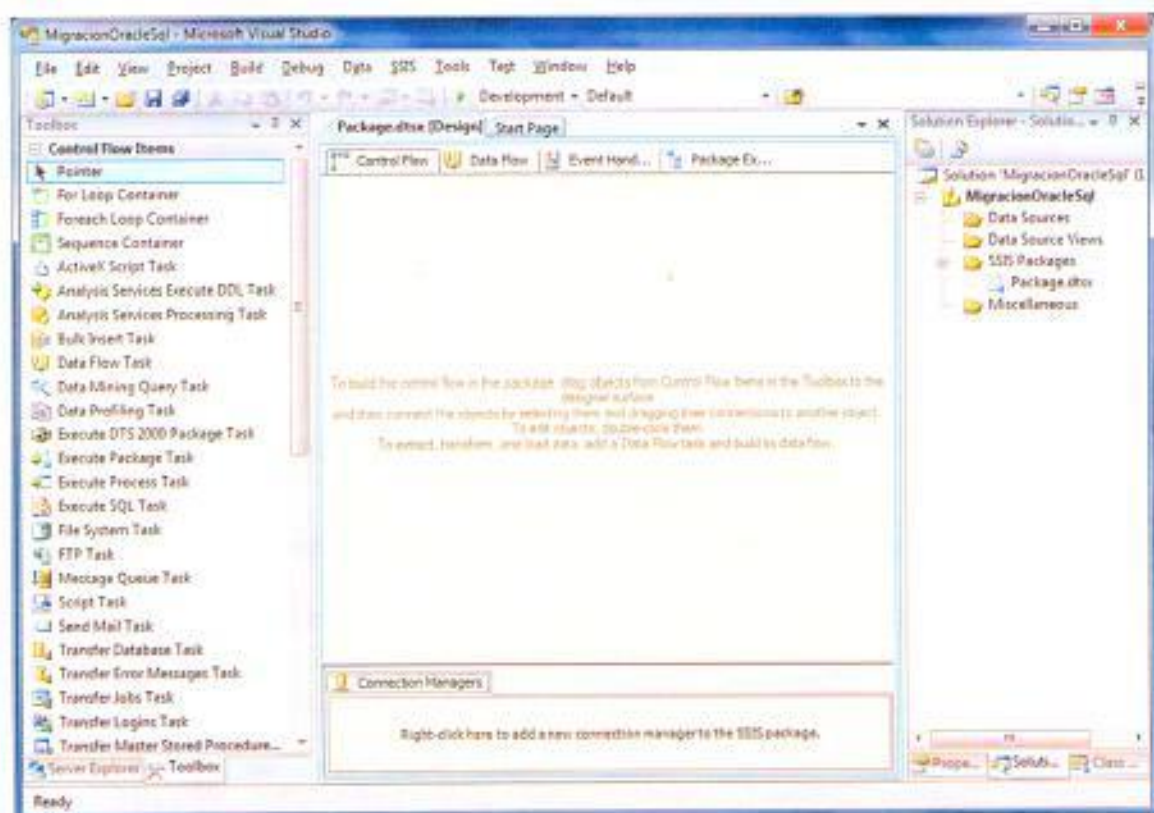


Figura 19. Microsoft Integration Services.

Como soporte adicional, se requiere el uso de scripts en PL/SQL y Transact-SQL que serán invocadas como tareas por el Microsoft Integration Services.



# CAPÍTULO 4

## 4. TRANSFORMACIÓN Y LIMPIEZA DE DATOS

### 4.1. Proceso de Extracción de datos

Para realizar el proceso de migración de datos, el primer proceso necesario a realizar es la extracción de datos desde el sistema origen, figura 20.

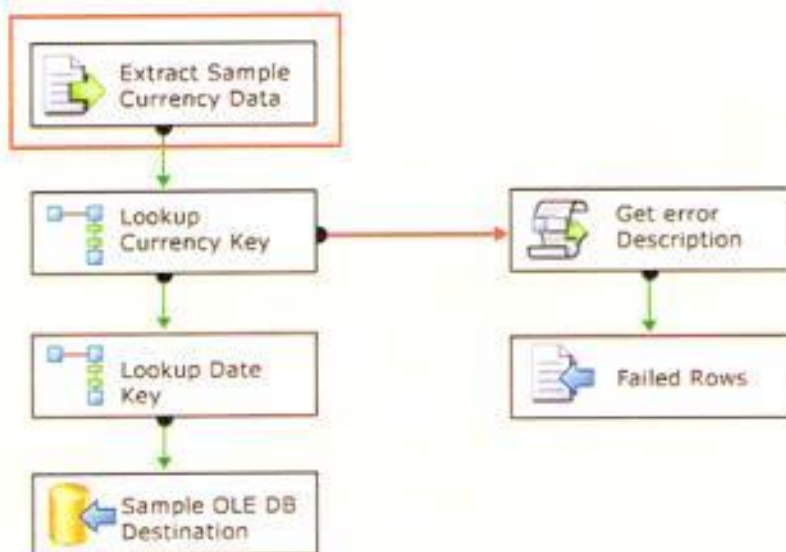


Figura 20. Proceso ETL – Extracción de datos.

Debido a que cada sistema de una organización puede manejar una organización diferente de los datos o estos pueden encontrarse bajo distintos formatos (texto plano, bases relaciones y no relaciones u otras estructuras) es necesario realizar un proceso de transformación para convertir los datos al formato correcto para iniciar el proceso de transformación. En el caso del Sistema de Control de Edificaciones, la extracción de datos se reduce a consultas en SQL, debido a que los datos se encuentran ya bajo un SGBDR[2]. Sin embargo, para sistemas que mantienen los datos en bases de datos relaciones, es importante que como parte del proceso de extracción se realice un chequeo que verifique si los datos cumplen la estructura que se esperaba y de no darse así rechazarlos.

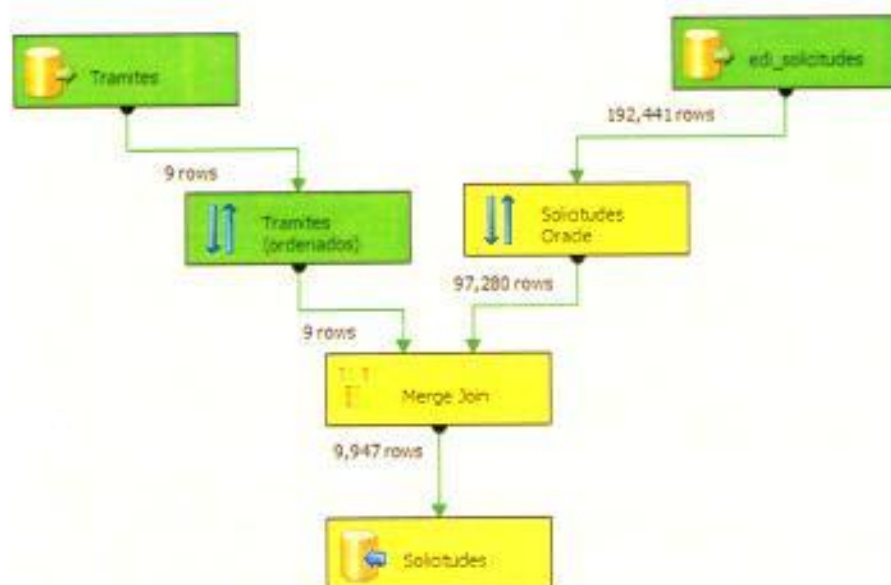


Figura 21. Proceso ETL – Extracción de datos de solicitudes.

Cuando se realice el proceso de extracción de datos siempre se debe tomar en cuenta que esta cause el mínimo impacto al sistema origen de los datos, por este motivo se debe buscar que dicha operación sea realizada en horarios en los cuales el impacto sea mínimo o nulo. Si los datos a ser extraídos son muchos, existe la posibilidad que el rendimiento del sistema origen disminuya o que colapse, interrumpiendo así su uso normal.

#### 4.2. Proceso de Limpieza de datos

Este proceso está diseñado para corregir o remover la información incorrecta (Fig. 21), con formato inapropiado o duplicado en la base de datos. Para esto es necesario definir un conjunto de reglas y algoritmos que permitan establecer que registros no cumplen con un formato en particular.

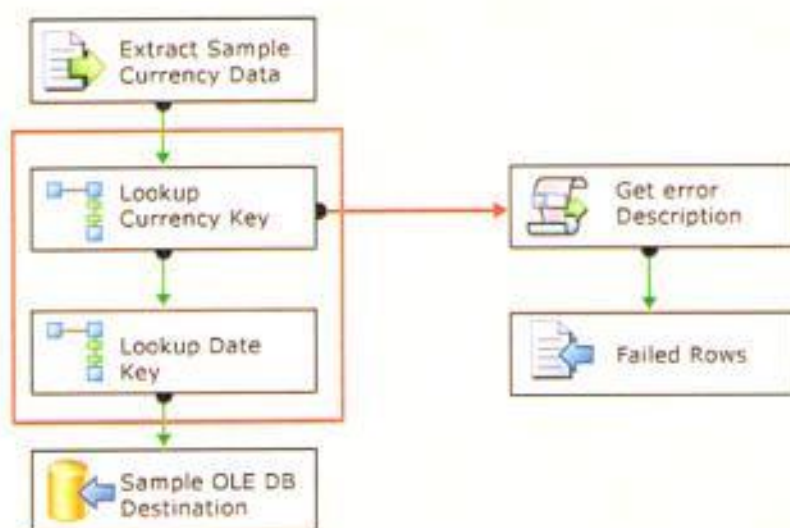


Figura 22. Proceso ETL – Carga de datos.

Por lo general una herramienta de limpieza de datos incluye programas que son capaces de corregir un número específico de tipos de errores como completar números telefónicos o encontrar registros duplicados. La utilización de una herramienta de limpieza de datos puede ahorrar un tiempo significativo al administrador de la base de datos y puede ser menos costoso que arreglarlo a mano.

Estas herramientas de limpieza de datos también pueden realizarse mediante scripts SQL que busquen inconsistencias en formatos.

A continuación, se detallan ciertas reglas que deberán ser aplicadas para el proceso de migración de datos del Sistema de Control de Edificaciones:

- Los campos que almacenan números de cédulas o RUC deben cumplir con las siguientes reglas:
  - o Contener únicamente números
    - Por ejemplo, si contiene el carácter '-' este deberá eliminarse
  - o Tener 10 dígitos en caso de ser una cedula



- Si el número guardado tiene 9 dígitos se deberá concatenar un 0 a la izquierda para que cumpla con la longitud de una cédula.
  - Tener 13 dígitos en caso de ser una cedula
    - Si el número guardado tiene 12 dígitos se deberá concatenar un 0 a la izquierda para que cumpla con la longitud de un RUC.
  - Identificar registros con cédulas y RUCs válidos
    - Ejecutar algoritmo validador implementado por el Servicio de Rentas Internas
- Los campos que almacenan números de registro de arquitecto deben cumplir con las siguientes reglas:
  - Contener únicamente números
  - Tener una longitud máxima de 6 dígitos
- Para los campos que almacenan fecha y hora:
  - Truncar el campo en caso de que en el nuevo modelo solo se almacene la fecha
  - Utilizar "dd/mm/yyyy hh24:mi:ss" como formato de conversión de fechas para fechas con año comprendido entre 50 y 99
  - Utilizar "dd/mm/rrrr hh24:mi:ss" como formato de conversión de fechas para fechas con año comprendido entre 0 y 49
- Para los campos que guardan un solo carácter:
  - Si la longitud del campo es igual a 1, actualizar el valor del campo con null



- Para las solicitudes:
  - o Verificar que se encuentren en un estado correcto acorde al sistema.
- Los campos que almacenan partes de un código predial:
  - o Contener únicamente números
    - Por ejemplo, si contiene letras estas deberán eliminarse

Adicionalmente, a las reglas de formato se deberán aplicar reglas que permitan identificar los registros duplicados en las tablas. Así mismo, se deben aplicar las reglas que son parte del negocio; por ejemplo, en el Sistema de Edificaciones, se deberán identificar las solicitudes que contienen predios que consten en solicitudes de cualquier tipo en trámite.

#### **4.3. Proceso de Transformación de datos**

En esta fase, se aplican ciertas reglas de negocio o funciones sobre los datos que han sido extraídos previamente y convertirlos en los datos que finalmente serán cargados a la nueva base de datos.

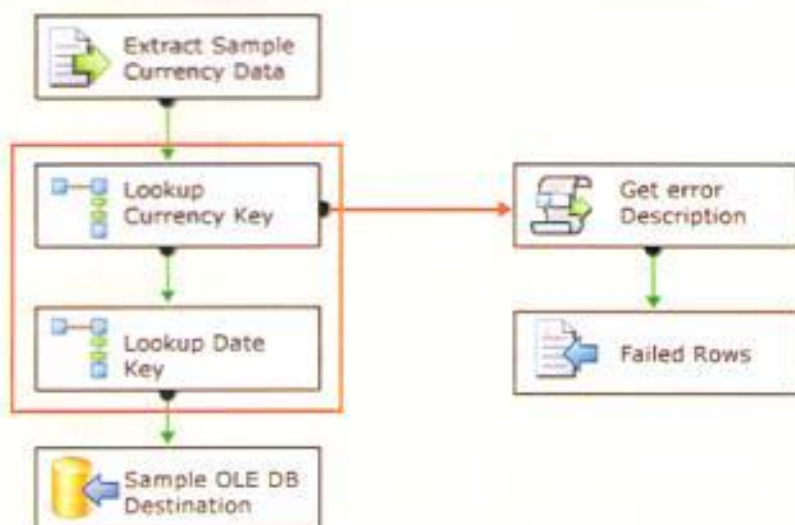


Figura 23. Proceso ETL – Transformación de datos.

Algunas de las siguientes transformaciones pueden ser aplicadas durante este proceso:

- Seleccionar las columnas a ser cargadas
  - o Columnas que durante el análisis de nulidad se identificó que únicamente contiene valores nulos no deben ser cargadas
- Traducir códigos
  - o Por ejemplo el Sistema de Control de Edificaciones en la base de datos actual, lleva el seguimiento de la solicitud manejando estados:
    - I para Solicitud Ingresada
    - S para Solicitud Aprobada por Supervisor
    - C para Solicitudes con Certificado Impreso, etc.

Sin embargo, para la nueva base de datos los estados serán identificadores numéricos que se encontraran registrados en el Sistema de Seguimiento diseñado por la empresa.

- Codificar valores libres
  - o Codificar la información referencial en tablas que se manejen como catálogos, por ejemplo para el Sistema de Control de Edificaciones los códigos de tipos de solicitudes serán reemplazados por números, es decir que por SRC (Solicitud de Registro de Construcción) será reemplazado por 1.
- Generar valores calculados
  - o Por ejemplo, para el Sistema de Control de Edificaciones, el valor a pagar por la liquidación de una solicitud de Inspección final es el 1%% del valor total de la construcción.
- Unir datos de múltiples fuentes
  - o Por ejemplo, para el Sistema de Control de Edificaciones, el seguimiento (histórico) de cada solicitud será migrado al Sistema de Seguimiento diseñado por la empresa; el código generado por este último sistema deberá asociarse con la solicitud para mantener la historia del trámite.
- Generación de campos clave en el destino.
  - o Por ejemplo, en la base de datos actual del Sistema de Control de Edificaciones, la clave primaria de la tabla que contiene la solicitud está compuesta por Año y Número de solicitud, pero en

el nuevo modelo la clave no será una clave compuesta y será un campo de tipo auto-numérico.

- Pivotar
  - o Convertir filas en columnas o columnas de una tabla en registros.
- Dividir una columna en varias o unificar columnas
  - o Por ejemplo, en el Sistema de Control de Edificaciones, el campo que guarda el nombre y el campo que guarda el apellido del solicitante deben unificarse en uno solo.
- Completar información
  - o Por ejemplo, el nuevo modelo de datos del Sistema de Control de Edificaciones contempla campos de auditoria que no soportan valores nulos, razón por la cual se deberán establecer valores por defectos que serán aplicados a los registros migrados.

Una vez que se han definido estas reglas y se tienen los datos en el formato esperado, se está listo para iniciar el proceso de carga. Sin embargo, durante este proceso se identificarán aquellos datos que no cumplen con las reglas establecidas.

#### **4.4. Administración de errores y revisión de información**

La aplicación de las reglas de validación y transformación de datos siempre presentará un escenario exitoso, en el cual los datos pasan la transformación y siguen a la siguiente etapa que es la etapa de carga; y el



escenario fallido, en el cual un registro no cumple con las reglas establecidas y debe ser reprocesado, Fig. 23.

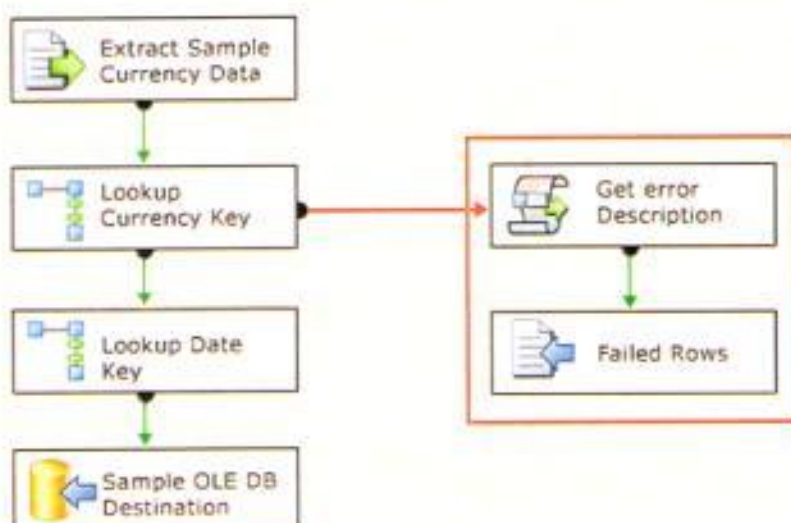


Figura 24. Proceso ETL – Falla en transformación de datos.

Antes de iniciar el proceso es necesario establecer un conjunto de políticas de tratamiento de las excepciones que pudieran ocasionarse durante la transformación de datos. Estas políticas deberán definir por ejemplo:

- Condiciones bajo las cuales se rechaza un registro por completo
- Valores por defectos (nulo o centinela) para campos erróneos en caso de que no impliquen el rechazo de un registro.

Por ejemplo en el caso del Sistema de Control de Edificaciones, se establecen como políticas de rechazo de registros:

- Predios que no existen en el Sistema de Catastro
- Solicitudes sin predios
- Solicitudes sin solicitantes
- Solicitudes cuyos solicitantes no tienen un número de cédula válido.

Otra política que puede ser aplicada es trabajar con el usuario responsable de la aplicación para que de esta forma sea el quien complete y corrija cierto conjunto de información que no puede administrado.

#### **4.5. Proceso de carga de datos**

En esta fase del proceso de migración de datos es cuando los datos de la fase de transformación son finalmente cargados en el sistema de destino.

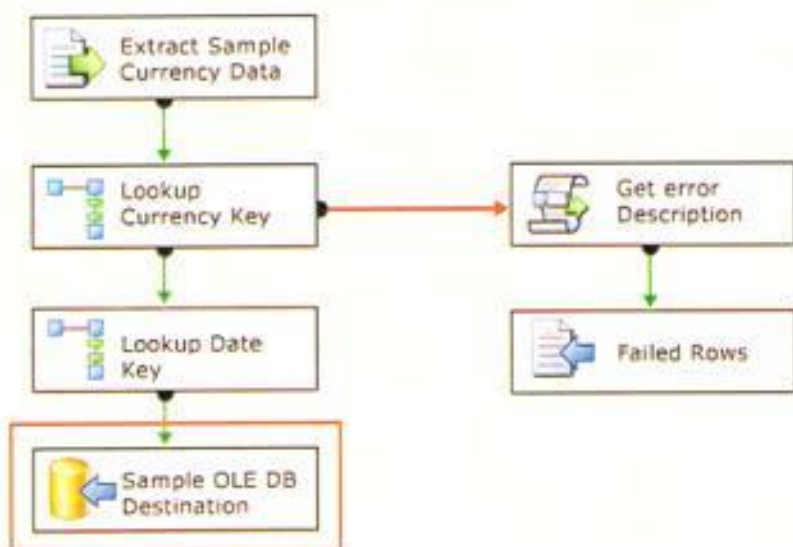


Figura 25. Proceso ETL – Carga de datos.

Para el caso de una migración de datos, este proceso consiste simplemente en insertar en las tablas los registros que provienen del proceso de transformación. Sin embargo, este proceso puede abarcar un conjunto de acciones adicionales, por ejemplo actualización de campos asociados a otros sistemas referenciales; en el caso del Sistema de Control de Edificaciones los solicitantes deberán asociarse con los ciudadanos registrados en el Sistema de Administración de Ciudadanos, mientras que la historia de la solicitud cargada en el Sistema de Seguimiento también deberá ser asociada.

Debido a que esta fase interactúa directamente con la base de datos de destino, es necesario tener en consideración que se aplicaran las distintas restricciones que se han definido sobre la misma (claves primarias, integridad referencial, campos que no soportan valores nulos, restricciones de rango de valores, triggers) contribuyendo de esta forma para que se garantice la integridad y calidad de los datos que han sido migrados.



# CAPÍTULO 5

## 5.PRUEBAS

### 5.1. Plan de Pruebas

Debido a que la información es uno de activos intangibles más importantes de cualquier empresa, el factor más crítico durante la migración de los datos es la realización de pruebas para validar que los datos se acoplen a la nueva estructura en que serán almacenados; esta comprobación también servirá para determinar si las aplicaciones continuaran funcionando correctamente.

Es necesario que mediante un conjunto de pruebas y escenarios se asegure la disponibilidad de la información ya que la nueva estructura de los datos puede resultar incompatible con la que existía. Por lo tanto, se debe verificar que todas las consultas continúen siendo ejecutadas exitosamente y que objetos como vistas, triggers y procedimientos almacenados estén correctamente compilados, de no ser ese caso deben ser modificados.

Uno de los aspectos importantes en el desarrollo del plan de pruebas es contar con el apoyo y coordinación de la gerencia responsable de los datos, ya que de esta forma se tendrá poder de decisión sobre cómo proceder en los escenarios errados que se detecten. Adicionalmente, se debe contar con un equipo que tenga la experiencia y competencia requerida para el procesamiento y personal que conozca que aplicaciones interactúan con los datos que están siendo afectados incluyendo el funcionamiento de las mismas.

En el plan de pruebas es donde se definen el alcance las pruebas, los responsables, las situaciones bajo las cuales el plan debe ser detenido, etc. El plan de pruebas del Sistema de Control de Edificaciones se encuentra en el Anexo H.

### **5.1.1. Definición de escenarios**

Como parte del plan de pruebas se deben definir cuáles son los casos o escenarios de prueba que serán verificados durante la etapa de pruebas del proceso de migración.

Estos escenarios deben contemplar todos aquellos posibles escenarios que violen las reglas y procesos de migración para de esta manera poder solucionar problemas como:

- Bases de datos fuente y destino bloqueadas por problemas en los procesos de migración.
- Procesos externos que no han sido deshabilitados y que acceden a la base de datos fuente.
- Registros rechazados por no cumplir las reglas de migración.
- Proceso de migración detenido por valores que no cumplen formatos requeridos.
- Baja calidad de datos, como por ejemplo el porcentaje de registros migrados exitosamente de una tabla.
- Diferencias de registros entre base fuente, base destino y repositorio de registros errados.
- Diferencias en valores luego ejecutar conciliaciones sobre los datos de la base fuente y destino. Por ejemplo, el proceso de migración pudo haber migrado exitosamente todos los registros de la tabla pero existe la posibilidad de que por diferencias en los modelos los valores de ciertas columnas hayan sido truncados.

Además, de considerarse los escenarios que violan las reglas establecidas, deben considerarse también aquellos escenarios que no las violan para de

esta manera poder darle seguimiento a todo el proceso implementado de tal forma que se confirmen los resultados esperados a nivel de base de datos.

**5.1.2. Definición de criterios de calidad**

Una vez establecidos los escenarios de migración o casos de prueba, se procede a la definición de las métricas de calidad para determinar la efectividad y eficiencia de la migración de datos realizada.

Con este propósito se utilizaran dos métricas, para las cuales se evaluarán varios criterios de aceptación. A continuación, se establecen un conjunto de criterios por métrica aplicada:

Métrica	Criterio
Ratio	Registros migrados
	Tablas migradas
	Registros migrados por tecnología
	Registros migrados por aplicación
	Datos con problemas de calidad
	Errores reconciliados
	Datos limpiados
Máximo	Personalizaciones en línea
	Errores de migración
	Impacto de migración en el tamaño de la base de datos
	Tiempo fuera de línea

Tabla 2. Criterios de calidad aplicados al proceso de migración de datos



Una vez ejecutadas las pruebas, se verifican los criterios de aceptación; si estos no satisfacen los parámetros de aceptación definidos, el proceso debe ser detenido y reversadas las acciones ejecutadas sobre la base de datos destino.

## **5.2. Pruebas unitarias de Extracción, Limpieza, Transformación y Carga**

Las pruebas unitarias comprenden las verificaciones asociadas a cada componente del proceso de migración. Su realización tiene como objetivo verificar la funcionalidad y estructura de cada componente individual; es decir, si un componente tiene la tarea de limpieza de un campo en este punto se puede verificar que ese proceso creado se encuentra funcionando efectivamente.

Durante este tipo de pruebas, además de verificar que cada componente del proceso de Extracción, Transformación y Carga funcione correctamente por separado, también se verifica que funcionen correctamente de forma integrada entre todos los procesos involucrados en llevar un registro desde una tabla en la base de datos fuente a la tabla correspondiente en la base de datos destino, es decir, que se verifica el correcto ensamblaje entre los distintos componentes del proceso.

Para el Sistema de Control de Edificaciones, mediante este tipo de pruebas se puede verificar que al migrar una solicitud de algún tipo de trámite desde el modelo actual, se genere correctamente toda la información relacionada a la misma en el nuevo modelo, es decir, la información que está relacionada a los predios afectados, seguimiento de la solicitud, modificaciones al registro de construcción si las hubiese, cantidad de registros que se deben crear en cada tabla asociada a la solicitud.

Entonces durante las pruebas unitarias se verificarán los siguientes escenarios:

- Extracción correcta de datos
- Cuantos registros debían ser creados en la base de datos destino para un registro fuente.
  - o Se crearon la cantidad correcta de registros y si no se crearon por qué no se crearon.
- Los datos fueron cargados dentro de los campos correctos, es decir, validar el mapeo de los campos
- Los datos cargados fueron cargados íntegramente con toda su información asociada o existen campos que no se migraron y se han perdido
- Los datos se encuentran correctamente formateados
- Cumplimiento de reglas de negocio

- Se necesita ejecutar alguna tarea pos-migración que sea requerida
- Consulta del registro por parte de las aplicaciones que los consume

La meta de la ejecución de este tipo de pruebas es confirmar que un registro puede ser migrado efectivamente mediante el proceso de migración establecido.

### **5.3. Pruebas de volumen de Extracción, Limpieza, Transformación y Carga**

Una vez realizadas las pruebas unitarias es necesario realizar pruebas de volumen para de esta forma verificar la respuesta del conjunto de componentes de los paquetes ETL frente a una gran cantidad de datos ya sea la cantidad real o una muestra.

Mediante este tipo de pruebas se determina que partes del proceso ETL pueden causar problemas de bloqueo y disminución del performance ocasionando que el proceso de migración pueda tomar mucho tiempo cuando se lo haga en el ambiente de producción.

De esta forma se busca que el tiempo de la ejecución de la migración de datos sea el menor posible. Sin embargo, este es un parámetro que

depende del tipo de empresa y sistema para el cual se está realizando el proceso.

#### **5.4. Análisis y clasificación de errores**

Durante la ejecución de pruebas unitarias y pruebas de cargas del conjunto de procesos ETL diseñados para el proceso de migración de datos, se procede a realizar un análisis de los errores detectados, así como su clasificación como leves y graves.

Errores de formato, por ejemplo fechas mal migradas, y errores de nulidad pueden ser categorizados como errores leves que luego de un simple análisis y corrección pueden ser superados. Sin embargo, errores en los cuales no se tratan de un solo campo sino de fallas en registros completos y generaciones erróneas o falta de generación de información se los debe clasificar como errores graves.

Es necesario realizar la clasificación de los errores para determinar el tiempo de corrección de los mismos y de esta forma poder realizar las correcciones necesarias dentro de la ventana de tiempo que se tiene disponible para la ejecución y conclusión de la migración de los datos.



## CONCLUSIONES Y RECOMENDACIONES

Numerosos son los desafíos para poner en marcha un proceso de migración de datos entre un sistema existente y uno nuevo dentro de cualquier organización. Este proceso llega a ser tan complejo que involucrar a personal de distintas áreas así como muchas tecnologías para poder llegar a este fin.

En este trabajo de tesis se ha presentado una solución al problema de migración de los datos en las aplicaciones para adecuarse a los nuevos requisitos que surgen con el paso del tiempo. Luego de haberse llevado a cabo todos los objetivos de su realización, se puede llegar a las siguientes conclusiones:

1. Un proceso de migración de datos involucra al personal de todas las áreas, tanto a los dueños de la información como al personal de Sistemas que realiza la migración y el personal de Producción encargado del control de todos los procesos.
2. Utilizar una metodología de migración con la cual se lleven los lineamientos paso a paso de los procesos necesarios para pasar los datos de una fuente a otra.

3. Coordinar correctamente los tiempos disponibles para cada una de las fases.
4. Establecer correctamente los tiempos para la ventana de migración ya que todas estas estructuras objetivo tienen requisitos diferentes de transformación de datos, y distintas latencias. Tener en cuenta que las transformaciones implicadas en los procesos ETL pueden ser muy complejas.

A continuación se detallan las recomendaciones que deberían tomarse en consideración:

1. Establecer reuniones con todo el personal de las distintas áreas involucradas.
2. Capacitar al personal de sistema en la creación y ejecución de procesos ETL ya que estos procesos pueden ser muy complejos y un sistema ETL mal diseñado puede provocar importantes problemas operativos.
3. Seguir la metodología de migración
4. Llevar a cabo la documentación de los procesos acorde a la metodología de migración
5. Realizar pruebas en ambientes de Testing.
6. Realizar pruebas de cargas que incluyan la carga estimada que pudieran tener los servidores de producción ya que el tiempo disponible para realizar la extracción de los sistemas de origen podría cambiar

entre ambientes, lo que implicaría que la misma cantidad de datos tendría que ser procesada en menos tiempo.

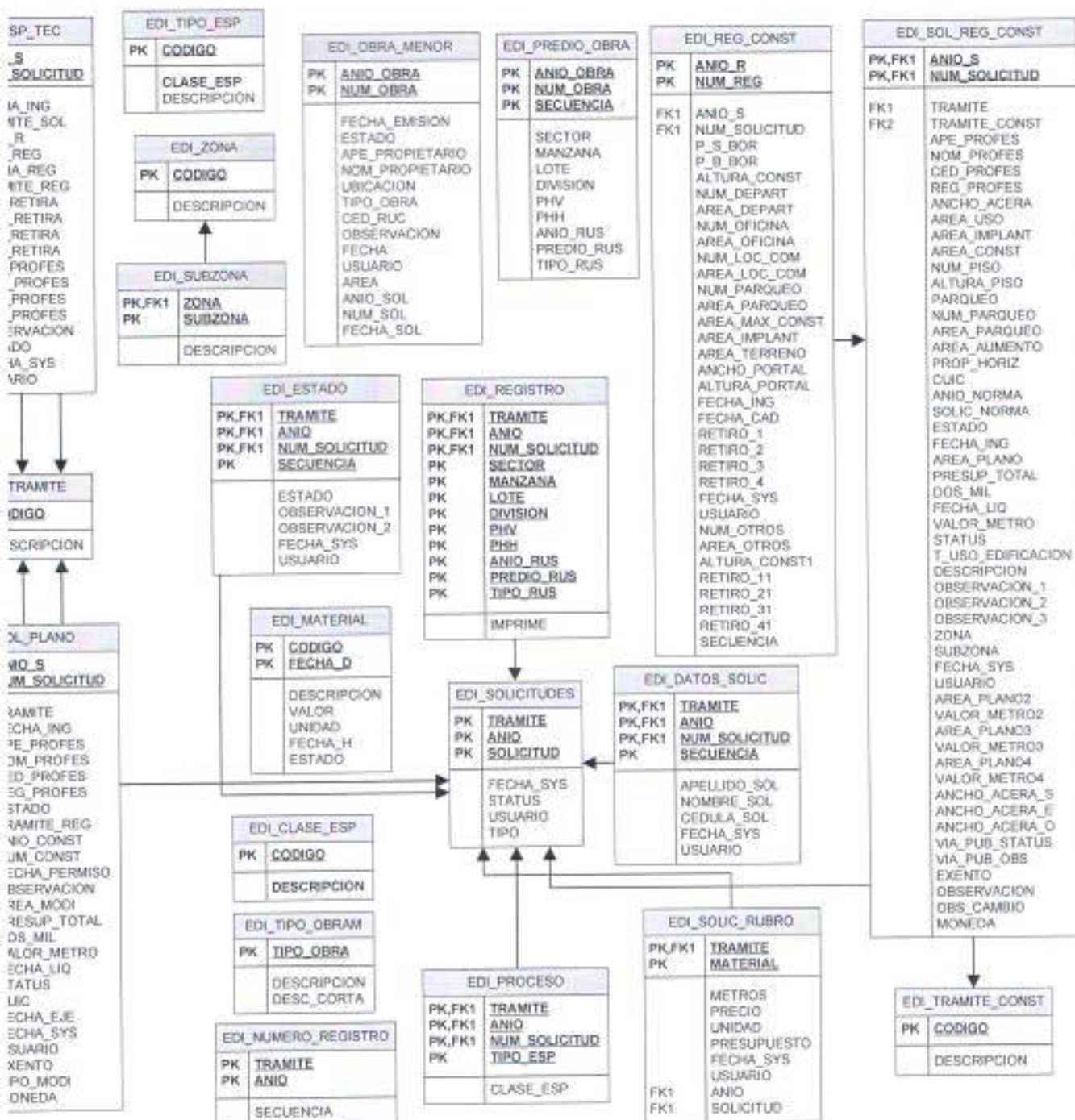
7. Realizar un examen completo de la validez de los datos (Data profiling) del sistema de origen durante el análisis para identificar las condiciones necesarias para que los datos puedan ser tratados adecuadamente por las reglas de transformación especificadas.

## APÉNDICES



# Anexo A

## Diagrama Entidad Relación del sistema actual



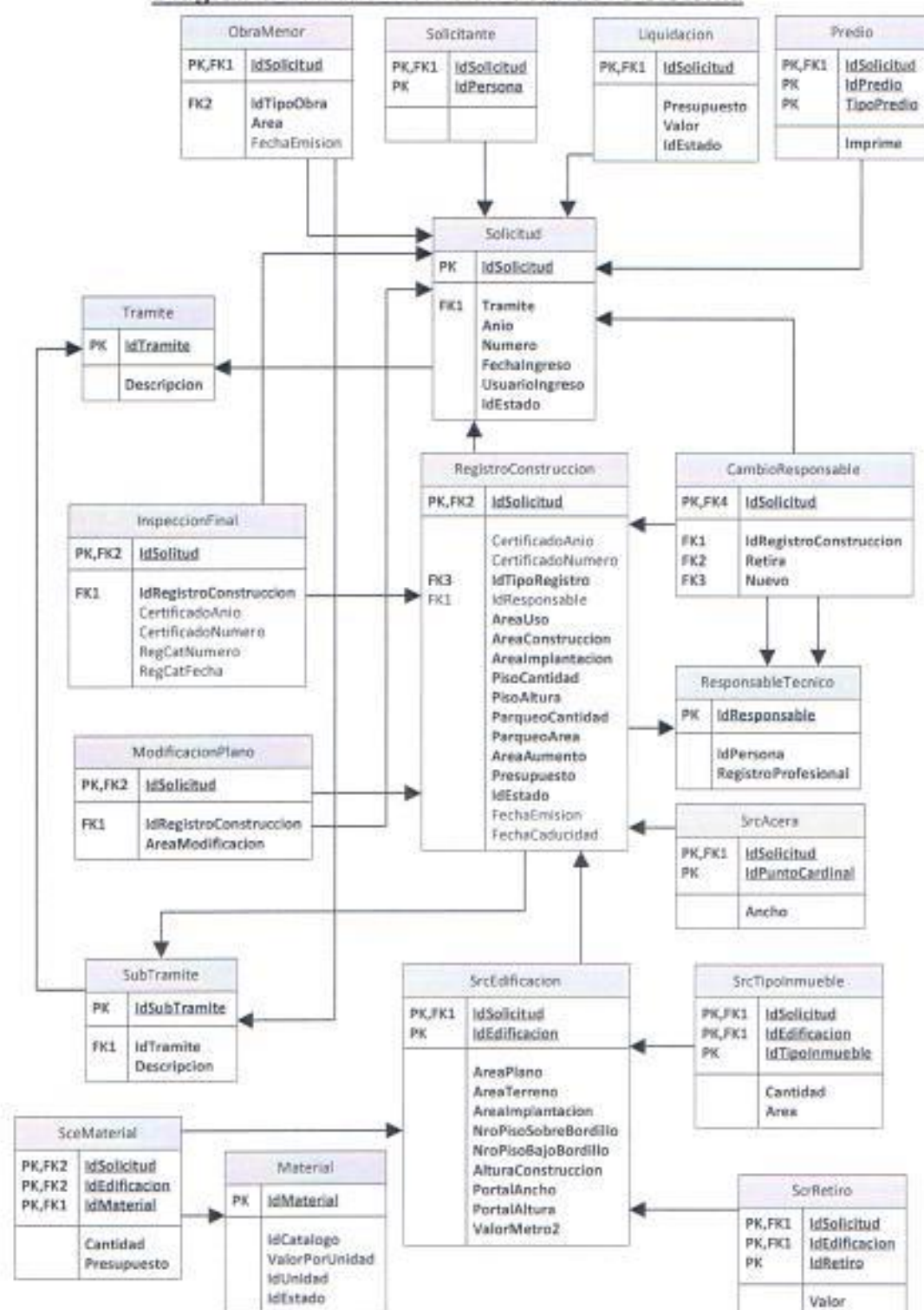
## Anexo B

### Ejemplo de documentación de tabla en el Diccionario de datos

<b>Tabla:</b>	EDI_OBRA_MENOR		
<b>Descripción:</b>	Tabla en la que se registran todos los trámites correspondientes a Obras Menores		
Columna	Tipo de dato	Nulo	Descripción
anio_obra	number(4)		Año de emisión de la Obra Menor
num_obra	number(6)		Número de emisión de la Obra Menor
fecha_emision	date	S	Fecha de emisión de la Obra Menor
estado	char(1)	S	Estado del trámite
ape_propietario	varchar2(30)	S	Apellidos del propietario del terreno
nom_propietario	varchar2(30)	S	Nombre del propietario del terreno
ubicacion	varchar2(100)	S	Dirección donde se ubica el terreno
tipo_obra	number(2)	S	Tipo de Obra Menor: Aumento o Remodelación
ced_ruc	varchar2(13)	S	Número de identificación del solicitante
observacion	varchar2(100)	S	Observación ingresada
fecha	date	S	Fecha de ingreso del trámite
usuario	varchar2(15)	S	Usuario que atiende el trámite
area	number(8,2)	S	Area de construcción en caso de una obra menor de aumento
anio_sol	number(4)	S	Año de la Solicitud de Obra Menor
num_sol	number(6)	S	Número de la Solicitud de Obra Menor
fecha_sol	date	S	Fecha de ingreso de la Solicitud de Obra Menor

# Anexo C

## Diagrama Entidad Relación del nuevo sistema



## Anexo D

### Ejemplo de documentación de mapeo de campos entre modelos

Mapeo de Tabla				
Tabla	SceSolObrasMenores			
Tablas Fuentes	Edi_Obra_Menor	Sistema en Oracle		
Campo	Origen	Campo	Valor por defecto	Unico
IdSolDetalle	--	--	Secuencia	--
IdTipoObraMenor	Edi_Obra_Menor	IdTipoObraMenor	--	--
Area	Edi_Obra_Menor	Area	--	--
RetiroFrontal1	--	--	null	--
RetiroFrontal2	--	--	null	--
RetiroLateral	--	--	null	--
RetiroPosterior	--	--	null	--
FechaMovimiento	Edi_Obra_Menor	Fecha	--	--
UsuarioAuditoria	--	--	MIGRACION	--
PcClienteAuditoria	--	--	MIGRACION	--
TransaccionAuditoria	--	--	MIGRACION	--
FechaIngreso	--	--	Getdate()	--
Observaciones				



# Anexo E

## Volumen de datos

Nombre de Tabla	Nro. de Reg.
EDI_ACTIVIDADES	13
EDI_AUDITORIA	2
EDI_AVALUO	27805
EDI_BLOQUE	8
EDI_CADUCIDAD	13
EDI_CERTIFICADOS	1
EDI_CLASE_ESP	13
EDI_CONDIC_COMPATIB	289
EDI_CONDIC_CONDIC	568
EDI_CONDIC_ESTACION	162
EDI_CONDIC_PERMITI	279
EDI_CONDIC_PROHIB	371
EDI_CONTRAVENCION	5
EDI_DATOS_SOLIC	33870
EDI_DESCUENTO_REGULA	2967
EDI_EDIFIC_SOPORTAL	79
EDI_ESPECIFICACION	1
EDI_ESTACIONAMIENTO	7
EDI_ESTADO	441912
EDI_FLUJO	679
EDI_FRONTAL	3
EDI_IMPRESION	0
EDI_INFORME_CORTE	3
EDI_INFORME_DETALLE	26
EDI_INSP_FINAL	9982
EDI_LINEA_LIN	3
EDI_MATERIAL	76
EDI_MOTI_DESCUENTO	4
EDI_MULTAS_CANCELADAS	0
EDI_MULTAS_REBAJA	0
EDI_NOTIFICACION	2704
EDI_NUMERO_REGISTRO	43
EDI_OPERATIVO	982
EDI_ORDEN_EDIF	902

Nombre de Tabla	Nro. de Reg.
EDI_PARAM	2777
EDI_PARAMETROS	51710
EDI_PREDIO_ABONO	1
EDI_PROCESO	97752
EDI_PROP_HORIZ	153
EDI_REGISTRO	56132
EDI_REGULARIZA	20636
EDI_REG_CONST	14280
EDI_REG_REG	7663
EDI_REG_REPAR	503
EDI_REPORTE_FINAL	8886
EDI_RESP_TEC	2065
EDI_RUBROS	0
EDI_SECTOR_COMISARIA	102
EDI_SOLICITUDES	118429
EDI_SOLIC_INSPE	12454
EDI_SOLIC_REPAR	942
EDI_SOLIC_RUBRO	6524
EDI_SOL_NORMA	4120
EDI_SOL_PLANO	7848
EDI_SOL_REGULA	13223
EDI_SOL_REG_CONST	21244
EDI_SOPORTAL	2
EDI_SUBZONA	110
EDI_TIPO_ESP	88
EDI_TIPO_USO	1
EDI_TRAMITE	9
EDI_TRAMITE_CONST	4
EDI_USO_COMPATIBLE	11
EDI_USO_CONDICIONADO	33
EDI_USO_PERMITIDO	14
EDI_USO_PROHIBIDO	29
EDI_ZONA	17
EDI_ZONA_SECTOR	0

## Anexo F

### RESULTADOS DE ANÁLISIS DE INTEGRIDAD, CONSISTENCIA Y NULIDAD

Tabla	Análisis de integridad		Análisis de Consistencia	Análisis de Nulidad
edi_solicitudes	Dominio	OK	Se identificaron 2.296 solicitudes con valor en campo trámite igual a 'OPE'. Dicho trámite no existe en la tabla edi_tramite.  Se identificaron 96.361 solicitudes con valor en campo trámite igual a 'XXX'. Dicho trámite no existe en la tabla edi_tramite.	OK
Entidad	OK			
Referencial	-			
Negocio	-			
edi_proceso	Dominio	OK	OK	OK
Entidad	OK			
Referencial	Se identifica que no existe la relación con la tabla edi_tipo_esp			
Negocio	-			

edi_datos_solic	Dominio	OK	Se identifican las siguientes inconsistencias: ✓ 4 registros en los cuales el valor del campo secuencia es mayor a 3. ✓ 1834 registros en los cuales la longitud del campo cedula_sol es menor a 9 ✓ 3018 registros en los campos nombre_sol y apellido_sol ✓ 1048 registros con el campo cedula_sol en null	Campo cédula es requerido en el nuevo sistema
	Entidad	Se identifica que el campo cedula_sol debe ser requerido.		
	Referencial	OK		
	Negocio	Se identifica que el campo secuencia almacena valores fuera del rango de valores establecidos (valores entre 1 y 3)		
edi_estado	Dominio	OK	OK	OK
	Entidad	OK		
	Referencial	OK		
	Negocio	-		
edi_sol_reg_con st	Dominio	OK	Se identifican las siguientes inconsistencias:	✓ Se identifica que los campos
	Entidad	OK		

	<p>Referencial Se identifica que no existe la relación con la tabla edi_subzona</p> <p>Negocio Se identifican las siguientes inconsistencias:</p> <ul style="list-style-type: none"> <li>✓ 4 registros constan como pagados (status=P) pero no tienen el código de pago (cuic)</li> <li>✓ 7 registros con fecha_ing &gt; 2009</li> <li>✓ 5 registros con fecha_lig &gt; 2009</li> <li>✓ 5 registros con fecha_sys &gt; 2009</li> <li>✓ Todas las liquidaciones se encuentran en sucres</li> <li>✓ 12428 registros con error en el cálculo de presupuesto</li> <li>✓ 3039 registros con error en la liquidación</li> </ul>	<ul style="list-style-type: none"> <li>✓ 409 registros en los cuales la longitud del campo ced_profes es menor a 9</li> <li>✓ 162 registros con el campo ced_profes en null</li> <li>✓ 38780 registros en los cuales el campo ape_profes contiene la profesión</li> <li>✓ 54 registros con el campo reg_profes en null</li> <li>✓ 29530 registros con zonas o subzonas no existentes en las tablas edi_zona y edi_subzona</li> </ul>	<p>"observacion_3",</p> <p>"moneda",</p> <p>"obs_cambio",</p> <p>"descripcion", y</p> <p>"exento" son nulos en todos los registros de la tabla (39126 registros)</p> <p>✓ El campo cédula es requerido en el nuevo modelo E/R</p>
edi_re_d_o	Dominio OK	Se identifican las siguientes	OK



	Entidad	OK	inconsistencias:	
	Referencial	OK	✓ 2 registros con fecha_sys > 2009	
	Negocio	-	✓ 8 registros con fecha_ing > 2009	
<b>edi_registro</b>	Dominio	OK	Se identifican 2296 registros	Se identifica que el campo "imprime" es null en todos los registros de la tabla (81472 registros)
	Entidad	OK	inconsistentes en los cuales el valor del campo tramite no existente en tabla edi_tramite	
	Referencial	OK		
	Negocio	-		
<b>edi_sol_plano</b>	Dominio	OK	Se identifican las siguientes inconsistencias:	Se identifica que los campos "excento" y "moneda" son nulos en todos los registros de la tabla (12705 registros)
	Entidad	OK		
	Referencial	OK	✓ 421 registros con el campo ced_profes con longitud menor a 9	
	Negocio	Se identifican las siguientes inconsistencias:	✓ 267 registros con el campo ced_profes en null	
		✓ 12705 registros con liquidaciones en sucres	✓ 12578 registros en los cuales el campo ape_profes contiene la profesión	El campo cédula es requerido en el nuevo modelo E/R
		✓ 331 registros con errores en el cálculo de presupuesto	✓ 9 registros con el campo reg_profes en null	
		✓ 370 registros con errores en el cálculo de la liquidación		

edi_solic_repar	<p>Dominio OK</p> <p>Entidad OK</p> <p>Referencial Se identifica que la tabla no está relacionado con la tabla edi_subzona</p> <p>Negocio Se identifican 1080 registros con error en el cálculo de la liquidación</p>	<p>Se identifican las siguientes inconsistencias:</p> <ul style="list-style-type: none"> <li>✓ 246 registros con el campo ced_profes con longitud menor a 9</li> <li>✓ 40 registros con el campo ced_profes en null</li> <li>✓ 1247 registros en los cuales el campo ape_profes contiene la profesión</li> <li>✓ 1 registro con el campo reg_profes en null</li> <li>✓ 341 registros con zonas no existentes</li> <li>✓ 783 registros con subzonas no existentes</li> </ul>	<p>Se identifica que los campos "excento", "moneda" y "observacion" son nulos en todos los registros de la tabla (1282 registros)</p> <p>El campo cédula es requerido en el nuevo modelo E/R</p>
e	<p>Dominio OK</p> <p>Entidad OK</p>	OK	OK

	Referencial	Se identifica que la tabla no se encuentra relacionada con la tabla edi_material		
	Negocio	Se identifican 69 registros con errores en cálculo de presupuesto		
edi_material	Dominio	OK	OK	OK
	Entidad	OK		
	Referencial	-		
	Negocio	-		
edi_reg_repar	Dominio	OK	OK	OK
	Entidad	OK		
	Referencial	-		
	Negocio	OK		
edi_solic_inspe	Dominio	OK	Se identifican las siguientes inconsistencias:	Se identifica que los campos "excento" y "moneda" son nulos en todos los registros de la tabla (21693 registros)
	Entidad	OK		
	Referencial	Se identifica que la tabla no se encuentra relacionada con las tablas edi_solicitudes, edi_reg_const y		
			✓ 308 registros con el campo	

	<p>edi_subzona</p> <p>Negocio Se identifican las siguientes inconsistencias:</p> <ul style="list-style-type: none"> <li>✓ 2595 registros con errores en el cálculo del presupuesto</li> <li>✓ 2778 registros con errores en el cálculo de la liquidación</li> </ul>	<p>ced_profes en null</p> <ul style="list-style-type: none"> <li>✓ 21503 registros en los cuales el campo ape_profes contiene la profesión</li> <li>✓ 2 registros con el campo reg_profes en null</li> <li>✓ 177 registros con registros de construcción no existentes</li> <li>✓ 16852 registros con subzonas inexistentes</li> <li>✓ 691 registros con zonas inexistentes</li> </ul>	<p>El campo cédula es requerido en el nuevo modelo E/R</p>
edi_insp_final	<p>Dominio OK</p> <p>Entidad OK</p> <p>Referencial Se identifica que la tabla no se encuentra relacionada con las tablas edi_solic_inspe y edi_reg_const</p>	<p>Se identifican las siguientes inconsistencias:</p> <ul style="list-style-type: none"> <li>✓ 168 registros con un registros de construcción no existente</li> <li>✓ 191 registros con longitud del campo ced_retira menor a 9</li> </ul>	<p>El campo cédula es requerido en el nuevo modelo E/R</p>



	Negocio	-	<ul style="list-style-type: none"> <li>✓ 149 registros con el campo ced_retira en null</li> <li>✓ 3545 registros en los cuales el campo Ape_retira contiene la profesión</li> <li>✓ 7 registros con el campo reg_retira en null</li> <li>✓ 45 registros con longitud del campo ced_profes menor a 9</li> <li>✓ 155 registros con el campo ced_profes en null</li> <li>✓ 3562 registros en los cuales el campo ape_profes contiene la profesión</li> <li>✓ 4 registros con el campo reg_profes en null</li> </ul>	
edi_obra_menor	Dominio	OK	Se identifican las siguientes inconsistencias:	El campo cédula es requerido en el nuevo modelo E/R
	Entidad	OK		
	Referencial	Se identifica que la tabla no se encuentra relacionada con la tabla edi_tipo_obra	<ul style="list-style-type: none"> <li>✓ 25082 registros en los cuales no corresponde el tipo de obra menor</li> </ul>	

	Negocio		<ul style="list-style-type: none"> <li>✓ 24044 registros en los cuales los campos ape_propietario, nom_propietario, ced_ruc registran valores nulos</li> <li>✓ 11 registros con números de cédulas(ced_ruc) con longitud menor a 10</li> </ul>	
edi_predio_obra	Dominio OK Entidad OK Referencial - Negocio -		OK	OK
edi_tipo_obram	Dominio OK Entidad OK Referencial - Negocio -		Se identifica que la tabla no contiene datos	OK
edi_zona	Dominio OK Entidad OK Referencial - Negocio -		OK	OK

edi_subzona	<p>Dominio OK</p> <p>Entidad OK</p> <p>Referencial Se identifica que la tabla no se encuentra relacionada con la tabla tabla edi_zona</p> <p>Negocio -</p>	OK	OK
edi_clase_esp	<p>Dominio OK</p> <p>Entidad OK</p> <p>Referencial -</p> <p>Negocio -</p>	OK	OK
edi_tipo_esp	<p>Dominio OK</p> <p>Entidad OK</p> <p>Referencial Se identifica que la tabla no se encuentra relacionada con la tabla tabla edi_clase_esp</p> <p>Negocio -</p>	Se identifican 81 registros inconsistentes en los cuales el campo clase_esp no corresponde a un tipo de código de la tabla edi_clase_esp	OK

## Anexo G

### RESULTADOS DE MÉTRICAS

A continuación, se muestra el registro obtenido durante la medición de una de las posibles métricas a ser aplicadas usando la función Ratio:

<b>Métrica 1:</b>	<b>Solicitudes de registro de construcción migradas exitosamente</b>
<b>Función</b>	$QD = 1 - \frac{\#de\ datos\ que\ satisfacen\ el\ criterio}{\#total\ de\ datos}$
<b># total de datos:</b>	26.696 registros
<b># datos satisfactorios:</b>	26.486 registros
<b># de no satisfactorios:</b>	210 registros
<b>QD Medida:</b>	0.7%
<b>Análisis y rango de aceptación:</b>	10%

A continuación, se muestran los registros obtenidos durante la medición de algunas de las posibles métricas a ser aplicadas usando la función MIN:

<b>Métrica 1:</b>	<b>Solicitudes con tasa de tramite de edificaciones correcto</b>
<b>Función</b>	$QD = \frac{\min(Codigo\ de\ tramite)}{1}$
<b># total de datos:</b>	192.441 registros
<b># datos satisfactorios:</b>	93.784 registros
<b># de no satisfactorios:</b>	98.657 registros
<b>QD Medida:</b>	1
<b>Análisis y rango de aceptación:</b>	93.784 registros
	Según los análisis de consistencia dichos registros



corresponden a solicitudes que nunca terminaron de ingresar al sistema (96.361 registros) ó corresponden a trámites que no serán migrados (OPE-2.296 registros). Cabe mencionar que no se registran novedades en los análisis de integridad y consistencia.

Métrica 2:	Solicitudes con por lo menos un solicitante
Función	$QD = \frac{\min(\text{Solicitante})}{1}$
# total de datos:	93.784 registros
# datos satisfactorios:	46.334 registros
# de no satisfactorios:	47.450 registros
QD Medida:	1
Análisis y rango de aceptación:	93.784 registros Según el análisis realizado durante la ejecución de los ETLs, nulidad y consistencia, los registros que no cumplen el criterio de calidad corresponden a registros con inconsistencias en el número de identificación.

Métrica 3:	Solicitantes registrados en el Sistema Ciudadano (Base de Registro Civil)
Función	$QD = \frac{\min(\text{Solicitante})}{1}$
# total de datos:	52.666 registros
# datos satisfactorios:	47.450 registros
# de no satisfactorios:	5.216
QD Medida:	1
Análisis y rango de aceptación:	52.666 registros Según el análisis realizado durante la ejecución de los ETLs y mediante el análisis de consistencia de

datos, los registros que no satisfacen el criterio corresponden a aquellos cuyo número de identificación exceden los 9 y 12 dígitos de números de cédulas o RUC. Adicionalmente, según el análisis de nulidad se identificaron registros con números de identificación nulos en los solicitantes.

<b>Métrica 4:</b>	<b>Responsables Técnicos registrados en el Sistema Ciudadano (Base de Registro Civil)</b>
<b>Función</b>	$QD = \frac{\min(\text{Solicitante})}{1}$
<b># total de datos:</b>	6.285 registros
<b># datos satisfactorios:</b>	5.881 registros
<b># de no satisfactorios:</b>	404 registros
<b>QD Medida:</b>	1
<b>Análisis y rango de aceptación:</b>	6.285 registros.  Según el análisis realizado durante la ejecución de los ETLs, nulidad y consistencia, los registros que no cumplen el criterio de calidad corresponden a registros con inconsistencias en el número de identificación.

<b>Métrica 5:</b>	<b>Registro de construcción con solicitud</b>
<b>Función</b>	$QD = \frac{\min(\text{Solicitante})}{1}$
<b># total de datos:</b>	26.696 registros
<b># datos satisfactorios:</b>	26.696 registros
<b># de no satisfactorios:</b>	0 registros
<b>QD Medida:</b>	1
<b>Análisis y rango de aceptación:</b>	26.696 registros.



Métrica 6:	Registro de Construcción con responsable Técnico
Función	$QD = \frac{\min(\text{Responsable Técnico})}{1}$
# total de datos:	26.696 registros
# datos satisfactorios:	26.486 registros
# de no satisfactorios:	210 registros
QD Medida:	1
Análisis y rango de aceptación:	6.285 registros. Según el análisis realizado durante la ejecución de los ETLs, nulidad y consistencia, los registros que no cumplen el criterio de calidad corresponden a registros con inconsistencias en el número de identificación.

# Anexo H

## PLAN DE PRUEBAS

### OBJETIVOS

- Validar que los datos se acoplen a la nueva estructura en que serán almacenados
- Asegurar la disponibilidad de la información luego de realizada la migración de datos.

### ALCANCE

El plan de pruebas debe considerar la migración de pre-requisitos y/o tablas referenciales, la creación de nuevas tablas (con y sin información), la migración de tablas existentes en la actual plataforma y finalmente la migración de datos históricos.

### DATOS GENERALES

Como se ha mencionado, la finalidad de ejecutar el plan de pruebas es verificar el resultado obtenido luego de la tarea realizada para comprobar, primero que todas las tareas hayan sido realizadas y segundo verificar la calidad de los resultados para cerciorarse de que todo haya sido realizado exitosamente o en su defecto realizar las correcciones correspondientes. En esta sección del plan se debe identificar, cuándo se ejecutó el plan, quién lo



ejecutó, quién lo supervisó y quién lo aprobó. Además, se puede agregar una descripción general del trabajo realizado.

FECHA	REALIZADO POR	SUPERVISADO POR	APROBADO POR
26JUN2012	ING. JUAN PUEBLO	ING. SUPERVISOR	ING. JORGE RODRIGUEZ
DESCRIPCIÓN			
El presente plan de pruebas ha sido ejecutado para la M.I. Municipalidad de Guayaquil, específicamente para el sistema de edificaciones.			

## FASES

### 1. Fase de migración de pre-requisitos ó tablas referenciales

#### 1.1 Listar todas las tablas referenciales o que son pre-requisitos.

NOMBRE DE LA TABLA	BASE DE DATOS	METODO DE CARGA DE DATOS UTILIZADO	RESULTADO ESPERADO	VERIFICACIÓN
Tipos de Funcionarios	COMUN	Script de inserción de registros	• 4 registros activos	Procesado correctamente
Profesiones	COMUN	Script de inserción de registros	• 3 registros activos	Procesado correctamente

En esta fase la verificación de datos se la realiza comparando cada objeto de la base de datos anterior con la nueva.

### 2. Fase de creación de nuevas tablas

Listar todas las tablas nuevas que van a ser implementadas en el nuevo esquema.

NOMBRE DE LA TABLA	CAMPOS	TIPO DE DATO
CAMBIORESPONSABLE	IDSOLICITUD	INT
	IDREGISTROCONSTRUCCION	INT
	RETIRA	INT
	NUEVO	INT
FUNCIONARIO	IDFUNCIONARIO	INT
	USUARIO	VARCHAR(15)
	NOMBRE	VARCHAR(100)

INSPECCIONFINAL	IDSOLITUD	INT
	IDREGISTROCONSTRUCCION	INT
	CERTIFICADOANIO	SMALLINT
	CERTIFICADONUMERO	SMALLINT
	REGCATNUMERO	INT
LIQUIDACION	REGCATFECHA	INT
	IDSOLICITUD	INT
	PRESUPUESTO	DECIMAL(8, 2)
	VALOR	DECIMAL(15, 2)
MATERIAL	IDESTADO	SMALLINT
	IDMATERIAL	SMALLINT
	IDCATALOGO	SMALLINT
	VALORPORUNIDAD	NUMERIC(8, 2)
	IDUNIDAD	SMALLINT
MODIFICACIONPLANO	IDESTADO	SMALLINT
	IDSOLICITUD	INT
	IDREGISTROCONSTRUCCION	INT
OBRAMENOR	AREAMODIFICACION	NUMERIC(15, 2)
	IDSOLICITUD	INT
	IDTIPOOBRA	SMALLINT
	AREA	NUMERIC(4, 2)
PREDIO	FECHAEMISION	DATETIME
	IDSOLICITUD	INT
	IDPREDIO	INT
	TIPOPREDIO	SMALLINT
REGISTROCONSTRUCCION	IMPRIME	BIT
	IDSOLICITUD	INT
	CERTIFICADOANIO	SMALLINT
	CERTIFICADONUMERO	DECIMAL(18, 0)
	IDTIPOREGISTRO	SMALLINT
	IDRESPONSABLE	INT
	AREAUSO	NUMERIC(15, 2)
	AREACONSTRUCCION	NUMERIC(15, 2)
	AREAIMPLANTACION	NUMERIC(15, 2)
	PISOCANTIDAD	SMALLINT
	PISOALTURA	NUMERIC(5, 2)
	PARQUEOCANTIDAD	SMALLINT
	PARQUEOAREA	NUMERIC(10, 2)
	AREAAUMENTO	NUMERIC(10, 2)
	PRESUPUESTO	DECIMAL(18, 2)
IDESTADO	SMALLINT	
FECHAEMISION	DATETIME	
FECHACADUCIDAD	DATETIME	
RESPONSABLETECNICO	IDRESPONSABLE	INT
	IDPERSONA	INT
	REGISTROPROFESIONAL	VARCHAR(10)
RETIRORESPONSABLE	IDSOLICITUD	INT
	IDRESPONSABLE	INT
	IDREGISTROCONSTRUCCION	INT
SCEMATERIAL	IDSOLICITUD	INT
	IDEDIFICACION	INT
	IDMATERIAL	SMALLINT

	CANTIDAD	NUMERIC(18, 2)
	PRESUPUESTO	NUMERIC(18, 2)
SRCRETIRO	IDSOLICITUD	INT
	IDEDIFICACION	INT
	IDRETIRO	INT
	VALOR	NUMERIC(5, 2)
SOLICITANTE	IDSOLICITUD	INT
	IDPERSONA	INT
SOLICITUD	IDSOLICITUD	INT
	TRAMITE	SMALLINT
	ANIO	SMALLINT
	NUMERO	INT
	FECHAINGRESO	DATETIME
	USUARIOINGRESO	VARCHAR(15)
	IDESTADO	INT
FUNCIONARIO	IDSOLICITUDFUNCIONARIO	INT
	IDSOLICITUD	INT
	IDFUNCIONARIO	INT
	FECHAINICIO	DATETIME
	FECHAFIN	DATETIME
	ESTADO	CHAR(1)
SRCACERA	IDSOLICITUD	INT
	IDPUNTOCARDINAL	INT
	ANCHO	DECIMAL(5, 2)
SRCEDIFICACION	IDSOLICITUD	INT
	IDEDIFICACION	INT
	AREAPLANO	NUMERIC(8, 2)
	AREATERRENO	NUMERIC(8, 2)
	AREAIMPLANTACION	NUMERIC(8, 2)
	NROPISOSOBREBORDILLO	VARCHAR(100)
	NROPISOBAJOBORDILLO	VARCHAR(100)
	ALTURACONSTRUCCION	NUMERIC(8, 2)
	PORTALANCHO	NUMERIC(5, 2)
	PORTALALTURA	NUMERIC(5, 2)
VALORMETRO2	DECIMAL(15, 2)	
TIPOINMUEBLE	IDSOLICITUD	INT
	IDEDIFICACION	INT
	IDTIPOINMUEBLE	SMALLINT
	CANTIDAD	SMALLINT
	AREA	NUMERIC(8, 2)
SUBTRAMITE	IDSUBTRAMITE	SMALLINT
	IDTRAMITE	SMALLINT
	DESCRIPCION	VARCHAR(50)
TRAMITE	IDTRAMITE	SMALLINT
	SIGLA	VARCHAR(5)
	DESCRIPCION	VARCHAR(250)

En esta fase la verificación de datos se la realiza comparando cada objeto de la base de datos anterior con la nueva.

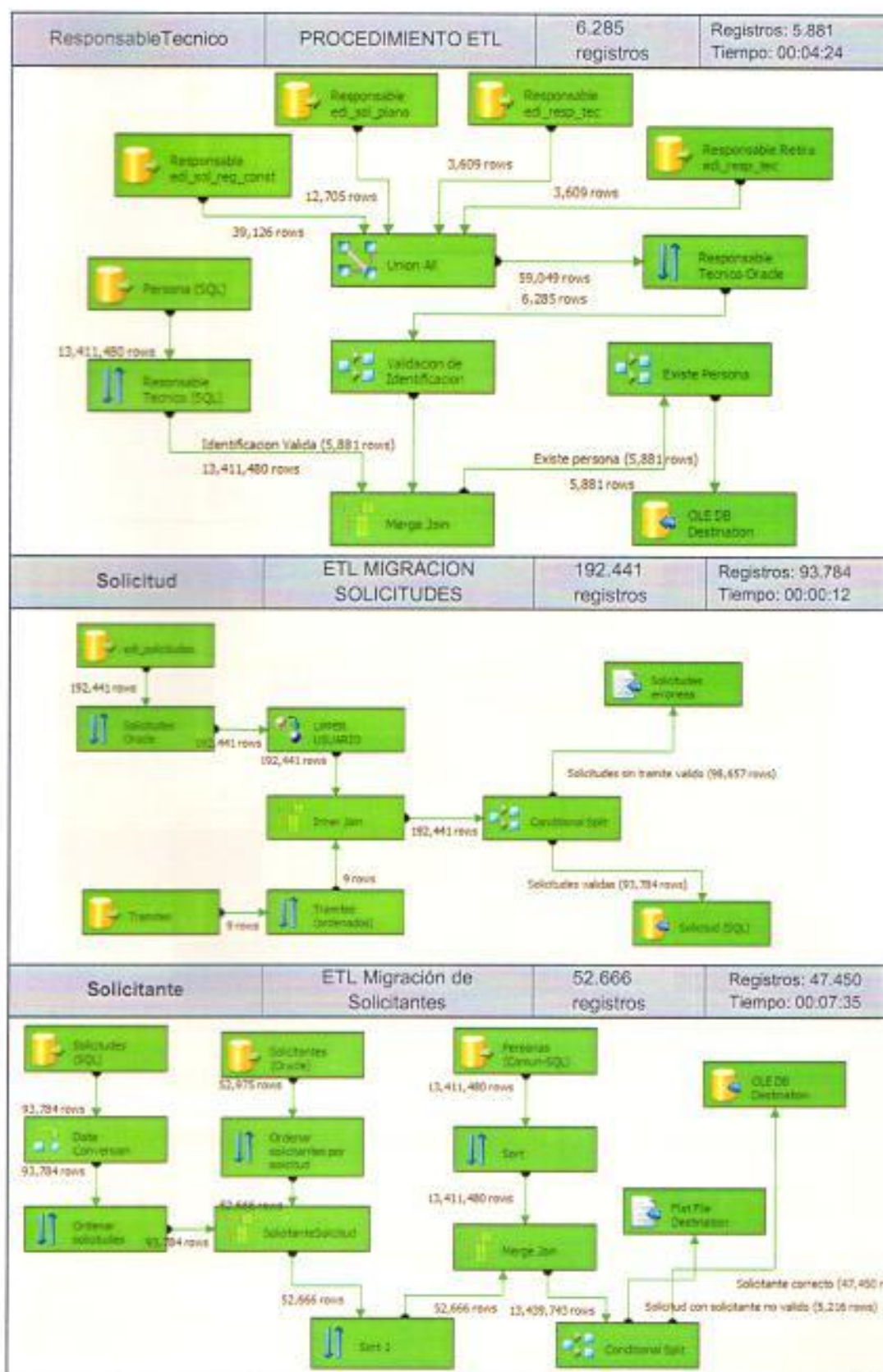


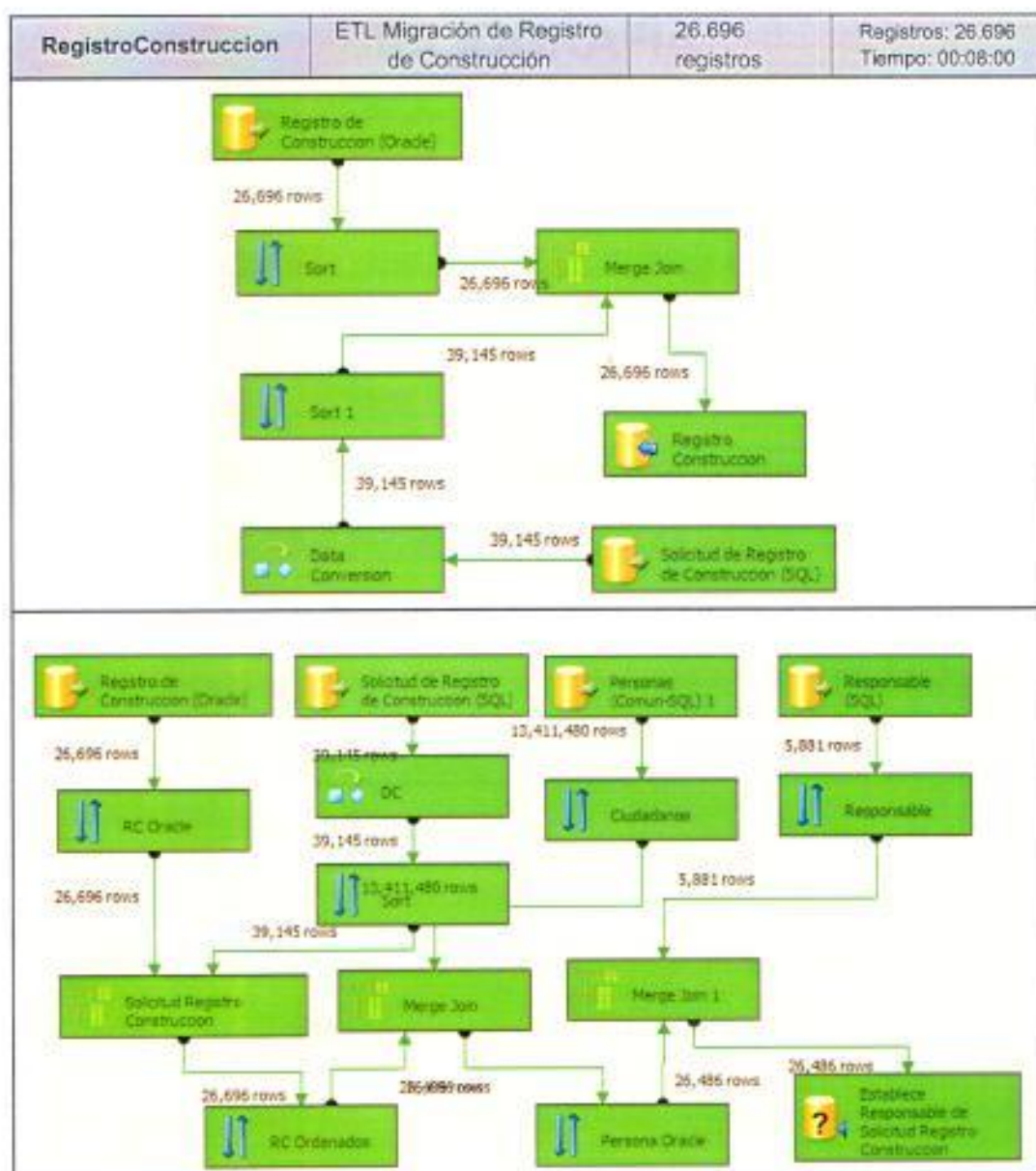
### 3. Migración de tablas existentes en la actual plataforma

A continuación, se muestran algunos ejemplos de los procesos ETL aplicados a la migración del sistema tomado como ejemplo.

NOMBRE DE LA TABLA	MÉTODO DE CARGA DE DATOS UTILIZADO	RESULTADO ESPERADO	VERIFICACIÓN
Tramite	<ul style="list-style-type: none"> <li>- PROCEDIMIENTO ETL</li> <li>- Script de inserción de tramite de Obra Menor</li> </ul>	<ul style="list-style-type: none"> <li>• 10 registros activos</li> </ul>	Procesado correctamente
<div style="text-align: center;">  </div> <pre> -- TRAMITE OBRA MENOR insert into Edificaciones.dbo.Tramite values('SOM','SOLIC. OBRA MENOR') GO                     </pre>			
Subtramite	<ul style="list-style-type: none"> <li>- Script de inserción de subtramites</li> </ul>	<ul style="list-style-type: none"> <li>• 4 registros</li> </ul>	Procesado correctamente
<pre> -- SUBTRAMITES DE REGISTRO DE CONSTRUCCION declare @IdTramite smallint  select      @IdTramite = IdTramite from Edificaciones.dbo.Tramite Where Sigla = 'SRC'  insert into Edificaciones.dbo.Subtramite values(@IdTramite, 'CONSTRUCCION INICIAL')  insert into Edificaciones.dbo.Subtramite values(@IdTramite, 'AUMENTO')  insert into Edificaciones.dbo.Subtramite values(@IdTramite, 'REMODELACION')  insert into Edificaciones.dbo.Subtramite values(@IdTramite, 'AUMENTO Y REMODELACION') GO                     </pre>			







#### 4. Migración de históricos

Crear las tablas del nuevo esquema en la base de datos histórica del nuevo sistema. Cabe indicar que esto depende de la cantidad de información que sea manejada por la empresa.

# BIBLIOGRAFÍA

1. Thomas M. Connolly (Author), Carolyn E. Begg (Author). Sistemas de Bases de Datos: Un Enfoque Practico Para Diseño, Implementación y Gestión (5ta Edición, 2009, Pearson Education)
2. Ismael Caballero, M<sup>a</sup> Ángeles Moraga, Coral Calero. Integración de Aspectos de Calidad de Datos en Sistemas de Información. Universidad de Castilla-La Mancha Grupo Alarcos - Instituto de Tecnologías y Sistemas de la Información P<sup>o</sup> de la Universidad 4, 13071 Ciudad Real.
3. Microsoft, SQL Server (<http://www.microsoft.com/sqlserver/>)
4. Microsoft, Integration Services (<http://www.microsoft.com/sqlserver/en/us/solutions-technologies/enterprise-information-management/integration.aspx>)
5. Oracle, Base de datos Oracle. <http://www.oracle.com>
6. J.Pérez, J.A.Carsi, I.Ramos, V.Anaya, J.Silva. Migración de datos automática a partir de la información de los esquemas conceptuales (Departamento de Sistemas Informáticos y Computación Universidad Politécnica de Valencia Camino de Vera s/n E-46071 Valencia – España).