ETL

etl - sir 2023

La guía de usuario está enfocada en caracterizar y definir los artefactos y procesos necesario para implementar un método de extracción, transformación y carga (ETL-Extract, Transform, Load) para el proyecto de migración de datos del Sistema de Información Registral.

Contenido

[INTRODUCCIÓN 1](#_Toc127136433)

[DEFINICIÓN DEL AMBIENTE DE TRABAJO 2](#_Toc127136434)

[DISEÑO E IMPLEMENTACIÓN DEL PROCESO ETL 3](#_Toc127136435)

[ETLs 9](#_Toc127136436)

[MODELO ETL 10](#_Toc127136437)

[ANALISIS Y RESULTADOS 21](#_Toc127136438)

[CONSLUSIONES 24](#_Toc127136439)

# INTRODUCCIÓN

La etapa de ETL (Extract, Transform, Load) tiene como objetivo proveer y mejorar el proceso de recopilación, limpieza y carga de datos de diferentes fuentes del Sistema de Información Registral – SIR de la Superintendencia de Notaría y Registro en un sistema de almacenamiento centralizado, mediante el diseño, construcción e implementación del modelo de la base de datos y los flujos de calidad. Es determinado el ambiente de trabajo que permita la extracción de los datos de diferentes formatos y estructuras, la limpieza y transformación de los mismos según las necesidades establecidas, y garantice la carga de datos de manera eficaz. La ejecución de la etapa permitirá mejorar la calidad de los datos, así como reducir los tiempos de procesamiento, contribuyendo así de manera positiva a la operación de la superintendencia.

# DEFINICIÓN DEL AMBIENTE DE TRABAJO

El ambiente de trabajo implementado se determinó a través del análisis y estudio de diferentes plataformas enfocadas en el ‘*Business Intelligence - BI’* y el análisis de datos de código abierto. Los criterios de selección se determinaron según los objetivos del modelo de negocio del proyecto y estipularon en los siguientes conceptos:

* Amplia gama de herramientas para integrar, transformar y analizar datos de diferentes fuentes.
* Ofrezca una interfaz gráfica intuitiva para crear y gestionar procesos ETL.
* Licenciamiento de código abierto.

Acorde a la definición del ambiente de trabajo, se determinó el uso de las siguientes herramientas que permiten el despliegue funcional de la etapa:

## ORACLE SQL DEVELOPER

Entorno de desarrollo integrado y regido bajo licencia gratuita que simplifica la gestión y desarrollo de las bases de datos de Oracle en implementaciones tradicionales y en la nube.

Versión: 22.2.1.234.1810 - 12 de septiembre de 2022.

## SPOON - PENTAHO DATA INEGRATION

Herramienta de integración de datos de código abierto escrita en Java que forma parte de la plataforma Pentaho Data Integration. Spoon despliega una interfaz gráfica de usuario que permite a los usuarios crear, probar y ejecutar flujos de trabajo de integración de datos, así como gestionar la integración de datos en tiempo real.

Versión: 9.3.0.0.428 - 12 de abril de 2022.

Requisitos: Preinstalación del JDK 11.

## jdk 11 – java development ki

Kit de desarrollo de software para la plataforma Java, que ofrece un conjunto de herramientas, bibliotecas y documentación referente a aplicaciones Java. Incluye el compilador Java, la máquina virtual Java (JVM), y bibliotecas Java estándar.

Versión: 11.0.13.

# DISEÑO E IMPLEMENTACIÓN DEL PROCESO ETL

## FASE I

Definición de los requisitos que definen el ETL y la base de datos según lo establecido en la versión 2.7 del modelo de negocio del proyecto. Los campos de información sobre las cuales se debe garantizar la calidad de los datos son:

* Dirección del Inmueble
* Nombre del Propietario
* Identificación del Propietario
* Sexo del Propietario
* Código de Departamento y Municipio (Divipola),
* Área del Predio

## DIRECCIÓN DEL INMUEBLE

Figura 1. Estructura del objeto.

Tabla 1. Campos de información de la entidad dirección del inmueble.

### NOMBRE DEL PROPIETARIO



Tabla 2. Campos de información Entidad dirección del inmueble.

**IDENTIFICACIÓN DEL PROPIETARIO**



Tabla 3. Campos de información Entidad dirección del inmueble.

**SEXO DEL PROPIETARIO**



Tabla 4. Campos de información Entidad dirección del inmueble.

**CÓDIGO DE DEPARTAMENTO Y MUNICIPIO (DIVIPOLA)** 

Tabla 5. Campos de información Entidad dirección del inmueble.

**AREA DEL PREDIO**



Tabla 6. Campos de información Entidad dirección del inmueble.

* Área privada: char [(10)] – Datos de longitud de cadena de 10 bytes de tamaño fijo.
* Área construida: char [(10)] – Datos de longitud de cadena de 10 bytes de tamaño fijo.
* Área total: char [(10)] – Datos de longitud de cadena de 10 bytes de tamaño fijo.
* Coeficiente: numeric (3,0) – Datos numéricos que tiene tres dígitos y cero decimales.

## fase II

Es diseñada la arquitectura de la base de datos a implementar, a partir del planteamiento y desarrollo de los modelos conceptual, lógico y físico, según lo establecido en los criterios de calidad especificados en la fase anterior.

### MODELO CONCEPTUAL

El modelo conceptual aterriza los objetos de información en diferentes entidades de datos y flujos de interacciones. El modelo conceptual diseñado establece las siguientes entidades como artefactos fundamentales de la base de datos a diseñar: 1) Ciudadano, 2) Anotación Ciudadano, 3) Folio, 4) Zona registral, 5) Dirección, 6) Oficina de Origen, 7) Documento, 8) Municipio, 9) Departamento y 10) Vereda.

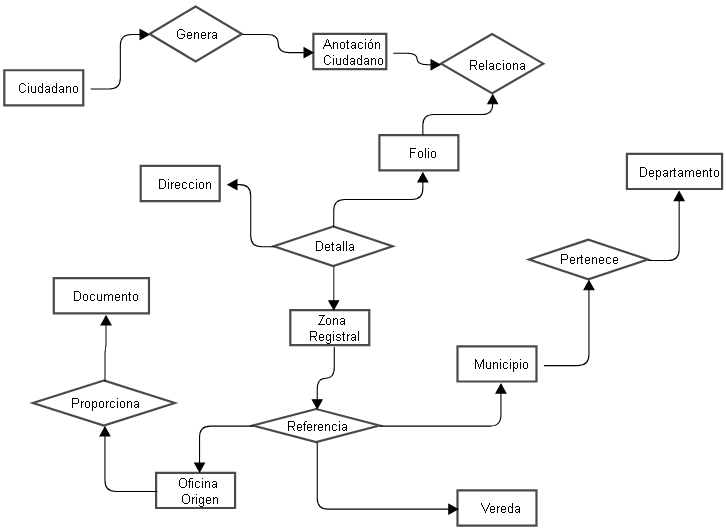


Figura 1. Modelo conceptual

### MODELO LOGICO

A través del modelo son relacionadas y normalizadas las entidades y flujos de interacción, como lo ilustra la Figura 2. A continuación son los especificados los atributos que conforman las entidades del modelo.

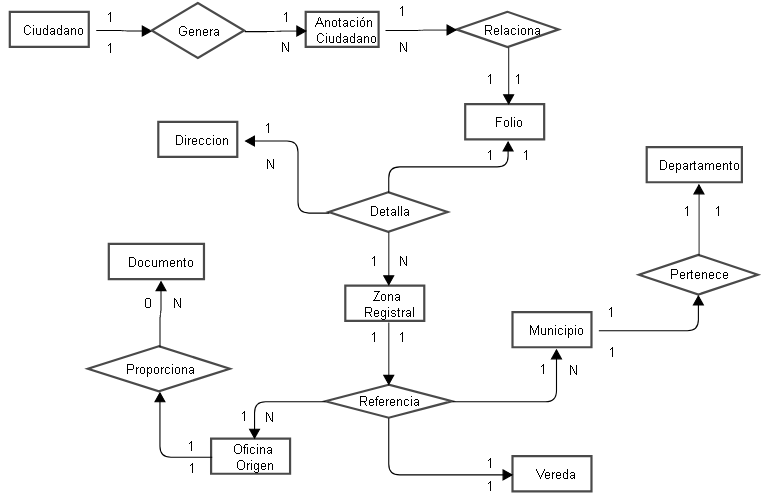


Figura 2. Modelo lógico.

#### Entidad Ciudadano

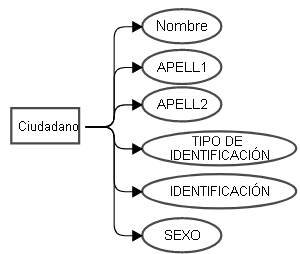


Figura 3. Entidad Ciudadano.

* Nombre: char [(65)] – Datos de longitud de cadena de 65 bytes de tamaño fijo.
* Apell1: char [(30)] – Datos de longitud de cadena de 30 bytes de tamaño fijo.
* Apell2: char [(30)] – Datos de longitud de cadena de 30 bytes de tamaño fijo.
* Tipo de identificación: char [(3)] – Datos de longitud de cadena de 2 bytes de tamaño fijo.
* Identificación: char [(12)] – Datos de longitud de cadena de 12 bytes de tamaño fijo.
* Sexo: char [(1)] – Datos de longitud de cadena de 1 bytes de tamaño fijo.

#### Entidad Anotación - Ciudadano

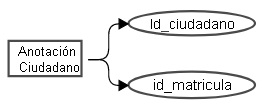


Figura 4. Entidad Anotación - Ciudadano.

* id\_ciudadano: char [(12)] – Datos de longitud de cadena de 12 bytes de tamaño fijo.
* id\_matricula: char [(10)] – Datos de longitud de cadena de 10 bytes de tamaño fijo.

#### Entidad Folio

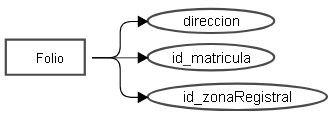


Figura 5. Entidad Folio.

* direccion: char [(n)] – Datos de longitud de cadena de n bytes de tamaño fijo.
* id\_matricula: char [(10)] – Datos de longitud de cadena de 10 bytes de tamaño fijo.
* id\_zonaRegistral: char [(n)] – Datos de longitud de cadena de n bytes de tamaño fijo.

#### Entidad Dirección

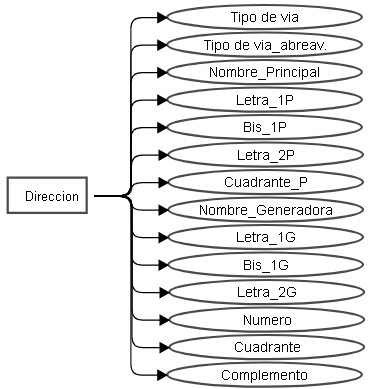


Figura 6. Entidad Dirección.

* TipoDeVia: char [(100)] – Datos de longitud de cadena de 100 bytes de tamaño fijo.
* TipoDeViaAbrev: char [(2)] – Datos de longitud de cadena de 2 bytes de tamaño fijo.
* Nombre\_Principal: char [(45)] – Datos de longitud de cadena de 45 bytes de tamaño fijo.
* Letra\_1P: char [(2)] – Datos de longitud de cadena de 2 bytes de tamaño fijo.
* Bis: char [(3)] – Datos de longitud de cadena de 2 bytes de tamaño fijo.
* Letra\_2P: char [(2)] – Datos de longitud de cadena de 2 bytes de tamaño fijo.
* Cuadrante\_P: char [(5)] – Datos de longitud de cadena de 5 bytes de tamaño fijo.
* Nombre\_generadora: char [(45)] – Datos de longitud de cadena de 45 bytes de tamaño fijo.
* Letra\_1G: char [(2)] – Datos de longitud de cadena de 2 bytes de tamaño fijo.
* Bis\_1G: char [(3)] – Datos de longitud de cadena de 2 bytes de tamaño fijo.
* Letra\_2G: char [(2)] – Datos de longitud de cadena de 2 bytes de tamaño fijo.
* Numero: numeric (4,0) – Datos numéricos que tiene cuatro dígitos y cero decimales.
* Cuadrante: char [(5)] – Datos de longitud de cadena de 5 bytes de tamaño fijo.
* Complemento: char [(45)] – Datos de longitud de cadena de 45 bytes de tamaño fijo.

#### Entidad Zona Registral

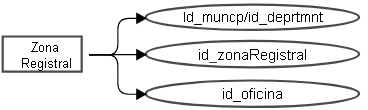


Figura 7. Zona Registral

* Id\_muncp/id\_deprtmnt: numeric (5,0) – Datos numéricos que tiene cinco dígitos y cero decimales.
* id\_zonaRegistral: char [(n)] – Datos de longitud de cadena de n bytes de tamaño fijo.
* id\_oficina: char [(n)] – Datos de longitud de cadena de n bytes de tamaño fijo.

#### Entidad Zona Registral



Figura 7. Zona Registral.

* ID\_Circulo: char [(3)] – Datos de longitud de cadena de tres bytes de tamaño fijo.

#### Entidad Documento

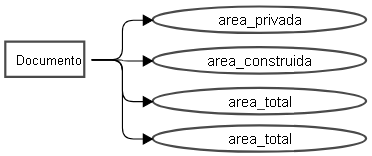
DIRE

Figura 8. Documento.

* Area\_Privada: numeric (10,0) – Dato numérico de diez dígitos y cero decimales.
* Area\_Construida: numeric (10,0) – Dato numérico de diez dígitos y cero decimales.
* Area\_Total: numeric (10,0) – Dato numérico de diez dígitos y cero decimales.
* Coeficiente: numeric (3,0) – Dato numérico de tres dígitos y cero decimales.

#### Entidad Municipio



Figura 8. Documento.

* cplobla: numeric (3,0) – Datos numéricos que tiene tres dígitos y cero decimales.
* Lmunic: char [(60)] – Datos de longitud de cadena de 60 bytes de tamaño fijo.

#### Entidad Departamento



Figura 8. Documento.

* cplobla: numeric (2,0) – Datos numéricos que tiene tres dígitos y cero decimales.
* ldepar: char [(40)] – Datos de longitud de cadena de 40 bytes de tamaño fijo.

# ETLs

Para el desarrollo del proceso ETL se definen las reglas de validación acorde al modelo de negocio del proyecto, el modelo de flujo de transformación y representación gráfica de las estructuras desarrolladas.

### Reglas de validación

Las reglas son definidas de acuerdo al análisis del modelo y las utilidades disponibles en la versión de la herramienta:

Conjunto de caracteres de control:

* Nueva línea (LF o "\n").
* Retorno de carro (CR o "\r").
* Tabulador (HT o "\t").
* Avance de página (FF o "\f").
* Sustitución (SUB o "\0").
* Retroceso (BS o "\b").
* Alerta (BEL o "\a").

Conjunto de caracteres especiales identificados:

* [A-Z] Alfabeto en mayúsculas.
* [a-z] Alfabeto en minúsculas.
* [0-9] Números naturales.
* [Ñ] Letra eñe en mayúsculas.
* [ñ] Letra eñe en minúsculas.
* [áéíóú] Vocales tildadas en minúscula.
* [ÁÉÍÓÚ] Vocales tildadas en mayúsculas.

Conjunto de utilidades de verificación disponibles en la herramienta:

* Trim Type: Permite quitar espacios en blancos, ya sea a la derecha, a la izquierda o a ambos lados de la cadena.
* Lower/Upper: Nos da la opción de poner la cadena, en mayúscula o minúscula sostenida.
* Padding, Pad Char y Pad Length: Con padding podemos elegir si queremos rellenar la cadena a la izquierda o a la derecha con el carácter definido en Pad Char, hasta la longitud definida en Pad Length. Ejemplo: si tenemos la cadena '11', definimos un Pad Cahr '0' y un Pad Length 5, el resultado será 00011.
* Escape: una función muy útil para escapar caracteres especiales según un lenguaje de programación o formato de archivo. Por ejemplo, en XML, las letras tildadas, las cambiará a su respectivo código.
* Digits: Deja solo los caracteres 0-9 si elegimos la opción only, o los elimina por completo si dejamos la opción remove.
* Remove Special character: permite elegir entre una lista de caracteres especiales a remover, por ejemplo, los retornos de carro, espacios en blanco o tabuladores.

# MODELO ETL

El proceso de desarrollo del ETL es realizado de acuerdo a la metodología de desarrollo de software Incremental, dado la orientación progresiva del proyecto, la agregación de funcionalidades y ejecución de modelos de forma secuencial en escenarios individuales y/o grupales.

La definición e implementación del ETL es determinada por etapas de ejecución, fases verificación en un proceso secuencial de ejecución y artefactos de fase.

Las etapas definen la manera en que se deben ejecutar los ETL en el proceso de migración y las fases de verificación representan la funcionalidad del ETL a través de artefactos de la misma, los artefactos son establecidos bajo la siguiente nomenclatura:

Extracción: EX iniciales de la entidad y orden de ejecución. Ej. EXCDN00 🡪 Extracción de la entidad ciudadano, orden de ejecución 00.

Transformación: TR iniciales de la entidad y orden de ejecución. Ej. TRCDN00 🡪 Transformación de la entidad ciudadano, orden de ejecución 00.

Carga: LD iniciales de la entidad y orden de ejecución. Ej. LDCDN00 🡪 Transformación de la entidad ciudadano, orden de ejecución 00.

Generación de archivo de texto: TO iniciales de la entidad y estado de ejecución. Ej. TOCDDN\_OK 🡪 Generación de archivo txt con los atributos y registros normalizados que fueron migrados a la SIR\_NE\_CIUDADANO.

**1° ETAPA DE RESPALDO Y CARGA DEL TABLOIDE:**

Es definido el gestor de base de datos PostgreSQL, como sistema de gestión de la base de dato local que ejercerá la función de almacén operacional de datos local, dado el enfoque relacional orientado a objetos y licencia de código abierto, publicado bajo la licencia PostgreSQL.

Las propiedades del ODS son las siguientes:

Data base: bd\_repo

Usuario: postgres

Clase: admin123

Direccion: localhost

**ETL 1:** ETL\_Respaldo: Ejecución del ETL enfocado en realizar una copia de los datos de información contenidos en las entidades a normalizar a través de la migración.

|  |  |  |
| --- | --- | --- |
| **PROCESO** | **OBJETIVO** | **IMPLEMENTACIÓN** |
| ENTIDAD | CIDUADANO | SIR\_NE\_CIUDADANO |
| EXCDN00 | Extraer todos los atributos de la entidad ciudadano del modelo SIR. | Selecciona todos los atributos y registros que conforman la entidad SIR\_NE\_CIUDADANO. |
| TRCDN00 | Establecer los atributos que conforman la entidad. | Selecciona los atributos requeridos por el modelo físico del SIR\_NE\_CIUDADANO y verifica la correcta normalización (nombre, tipo y longitud). |
| LDCDN00 | Construir en el ODS local una copia exacta de la entidad y registros del SIR\_NE\_CIUDADANO. | Carga los registros de información determinados en el paso anterior a una tabla del ODS local denominada SIR\_NE\_CIUDADANO\_RES. |
| ENTIDAD | DIRECCIÓN | SIR\_NE\_DIRECCION |
| EXDIR00 | Extraer todos los atributos de la entidad dirección del modelo SIR. | Selecciona todos los atributos y registros que conforman la entidad SIR\_NE\_DIRECCION |
| TRDIR00 | Establecer los atributos que conforman la entidad. | Selecciona los atributos requeridos por el modelo físico del SIR\_NE\_DIRECCION y verifica la correcta normalización (nombre, tipo y longitud). |
| LDDIR00 | Construir en el ODS local una copia exacta de la entidad y registros del SIR\_NE\_DIRECCION. | Carga los registros de información determinados en el paso anterior a una tabla del ODS local denominada SIR\_NE\_DIRECCION\_RES. |
| ENTIDAD | DEPARTAMENTO | SIR\_NE\_DEPARTAMENTO |
| EXDEP00 | Extraer todos los atributos de la entidad departamento del modelo SIR. | Selecciona todos los atributos y registros que conforman la entidad SIR\_NE\_DEPARTAMENTO |
| TRDEP00 | Establecer los atributos que conforman la entidad. | Selecciona los atributos requeridos por el modelo físico del SIR\_NE\_DEPARTAMENTO y verifica la correcta normalización (nombre, tipo y longitud). |
| LDDEP00 | Construir en el ODS local una copia exacta de la entidad y registros del SIR\_NE\_DEPARTAMENTO. | Carga los registros de información determinados en el paso anterior a una tabla del ODS local denominada SIR\_NE\_DEPARTAMENTO\_RES. |
| ENTIDAD | MUNICIPIO | SIR\_NE\_MUNICIPIO |
| EXMUN00 | Extraer todos los atributos de la entidad MUNICIPIO del modelo SIR. | Selecciona todos los atributos y registros que conforman la entidad SIR\_NE\_MUNICIPIO |
| TRMUN00 | Establecer los atributos que conforman la entidad. | Selecciona los atributos requeridos por el modelo físico del SIR\_NE\_MUNICIPIO y verifica la correcta normalización (nombre, tipo y longitud). |
| LDMUN00 | Construir en el ODS local una copia exacta de la entidad y registros del SIR\_NE\_MUNICIPIO. | Carga los registros de información determinados en el paso anterior a una tabla del ODS local denominada SIR\_NE\_MUNICIPIO \_RES. |
| ENTIDAD | FOLIO | SIR\_NE\_FOLIO |
| EXFOL00 | Extraer todos los atributos de la entidad FOLIO del modelo SIR. | Selecciona todos los atributos y registros que conforman la entidad SIR\_NE\_FOLIO. |
| TRFOL00 | Establecer los atributos que conforman la entidad. | Selecciona los atributos requeridos por el modelo físico del SIR\_NE\_FOLIO y verifica la correcta normalización (nombre, tipo y longitud). |
| LDFOL00 | Construir en el ODS local una copia exacta de la entidad y registros del SIR\_NE\_FOLIO. | Carga los registros de información determinados en el paso anterior a una tabla del ODS local denominada SIR\_NE\_FOLIO \_RES. |
| ENTIDAD | ZONA REGISTRAL | SIR\_NE\_ZONA\_REGISTRAL |
| EXZNR00 | Extraer todos los atributos de la entidad Zona Registral del modelo SIR. | Selecciona todos los atributos y registros que conforman la entidad SIR\_NE\_ZONA\_REGISTRAL. |
| TRZNR00 | Establecer los atributos que conforman la entidad. | Selecciona los atributos requeridos por el modelo físico del SIR\_NE\_ZONA\_REGISTRAL y verifica la correcta normalización (nombre, tipo y longitud). |
| LDZNR00 | Construir en el ODS local una copia exacta de la entidad y registros del SIR\_NE\_ZONA\_REGISTRAL. | Carga los registros de información determinados en el paso anterior a una tabla del ODS local denominada SIR\_NE\_ZONA\_REGISTRAL\_RES. |
| ENTIDAD | ZONA VEREDA | SIR\_NE\_VEREDA |
| EXVER00 | Extraer todos los atributos de la entidad Vereda del modelo SIR. | Selecciona todos los atributos y registros que conforman la entidad SIR\_NE\_VEREDA. |
| TRVER00 | Establecer los atributos que conforman la entidad. | Selecciona los atributos requeridos por el modelo físico del SIR\_NE\_VEREDA y verifica la correcta normalización (nombre, tipo y longitud). |
| LDVER00 | Construir en el ODS local una copia exacta de la entidad y registros del SIR\_NE\_VEREDA. | Carga los registros de información determinados en el paso anterior a una tabla del ODS local denominada SIR\_NE\_VEREDA \_RES. |

Tabla 1. Flujo ETL respaldo.

**Modelo visual ETL 1:** ETL\_Respaldo: Ejecución del ETL enfocado en realizar una copia de los datos de información contenidos en las entidades a normalizar a través de la migración.

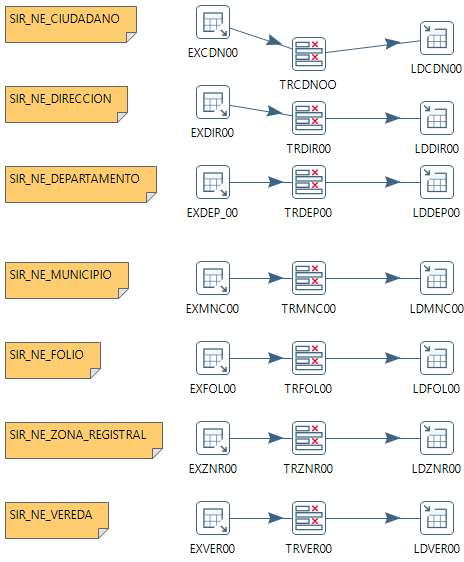


Figura 9. ETL1 respaldo.

**Archivo del ETL 1:** ETL\_RESPALDO.ktr

**ETL 2:** ETL\_LOADODS: Ejecución del ETL enfocado en realizar una copia de los datos de información contenidos en el tabloide de calidad de datos en formato CSV, referente a las entidades a normalizadas.

|  |  |  |
| --- | --- | --- |
| **PROCESO** | **OBJETIVO** | **IMPLEMENTACIÓN** |
| TABLOIDE | FORMATO | CSV |
| EXCSV00 | Extraer todos los atributos normalizados de las entidades contenidas en el tabloide csv. | Selecciona todos los atributos y registros que conforman el tabloide EjemploNormalizacion\_FuncionalV1.csv. |
| TRCSV00 | Establecer los atributos que conforman el tabloide. | Selecciona los atributos requeridos por el modelo físico del SIR y verifica la correcta normalización (nombre, tipo y longitud). |
| TRCSV01 | Establecer los atributos que conforman el tabloide. | Selecciona los atributos requeridos por el modelo físico del SIR y verifica la correcta normalización (nombre, tipo y longitud). |
| LDCSV00 | Construir en el ODS local una copia exacta de la TABLOIDE CSV. | Carga los registros de información determinados en el paso anterior a una tabla del ODS local denominada GRUPO\_NORMALIZADO. |

Tabla 2. Flujo ETL\_LOADODS.

**Modelo visual ETL 2:** ETL\_LOADODS: Ejecución del ETL enfocado en realizar una copia de los datos de información contenidos en el tabloide normalizado.

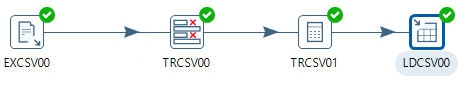


Figura 10. ETL 2 - LoadODS.

**Archivo del ETL 2:** ETL\_LOADODS.ktr

**2° ETAPA - MIGRACIÓN DE INFORMACIÓN:**

Es definido el flujo de información de los atributos normalizados que van a migrar a las entidades del modelo SIR, definidos por la fase de selección y fase de hilos de individual por entidad.

**ETL PRINCIPAL:** ETL\_PRINCIPAL\_VF: Ejecución del ETL principal enfocado en realizar la migración de los datos de información normalizados, que contienen los atributos de las entidades a normalizaren el modelo SIR.

|  |  |  |
| --- | --- | --- |
| **PROCESO** | **OBJETIVO** | **IMPLEMENTACIÓN** |
| FASE SELECTOR |  | Versión Estable |
| EXSEL00 | Extraer todos los atributos normalizados del ODS. | Selecciona todos los atributos y registros que conforman el tabloide normalizado que están alojado en el ODS Local (Grupo Normalizado). |
| TRSEL00 | Establecer los atributos normalizados. | Selecciona los atributos requeridos por el modelo físico del SIR y verifica la correcta normalización (nombre, tipo y longitud). |
| TRSEL01 | Determinar el círculo de trabajo que se desea migrar. | Filtrar los atributos y registros normalizados acorde al círculo de trabajo a migrar. |
| TRSEL02 | Establecer la tabla temporal de trabajo que contiene los atributos y registros normalizados de trabajo según el circulo. | Selecciona los atributos y registros normalizados y los aloja en una tabla temporal de trabajo, a fin de iniciar los hilos individuales por entidad. |

Tabla 3. Flujo ETL\_PRINCIPAL\_VF – Fase Selector.

**Modelo visual fase selector:** ETL\_PRINCIPAL\_VF: Fase enfocada en preparar el bloque de información normalizada según el círculo de trabajo determinado.

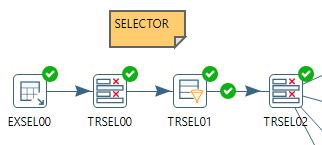


Figura 11. ETL\_PRINCIPAL\_VF – Fase Selector.

|  |  |  |
| --- | --- | --- |
| **PROCESO** | **OBJETIVO** | **IMPLEMENTACIÓN** |
| FASE HILO | INDIVIDUAL | ENTIDAD CIUDADANO (Versión estable) |
| TRHC00 | Inicializar la entidad Ciudadano. | Selección de los atributos normalizados de la entidad Ciudadano acorde al SIR\_NE\_CIUDADANO. |
| TRHC01 | Verificar errores tipográficos. | Establecimiento de los criterios que identifican los posibles errores tipográficos que pueden presentarse y deben ser corregidos, acorde a las reglas del modelo de negocio establecidas para los campos normalizados. |
| TRHC02 | Eliminar caracteres de control. | Parametrización de los caracteres de control que deben eliminarse en los atributos normalizados. |
| TRHC03 | Establecer el criterio de diferenciación de los atributos normalizados y aplicación de utilidades de verificación. | Parametrización de formato “Mayúsculas” como criterio de diferenciación de los atributos normalizados y verficación de espacios al inicio o final del registro. |
| TRHC04 | Ordenar los objetos de la entidad Ciudadano. | Identificación del atributo ID\_Ciudadano como llave primaria, a fin de realizar un ordenamiento ascendente. |
| TRHC05 | Verificación de no existencia de objetos vacíos. | Filtro de los objetos vacío con base al atributo normalizado ID\_Ciudadano (condición ID\_CIUDADANO IS NOT NULL). |
| TRHC07 | Verificación de no existencia de objetos duplicados. | Eliminación de los objetos duplicados con base al atributo normalizado ID\_Ciudadano. |
| TRHC07 | Ordenar los objetos de la entidad Ciudadano. | Identificación del atributo ID\_Ciudadano como llave primaria, a fin de realizar un ordenamiento ascendente. |
| TRHC08 | Validación de los atributos que conforman la entidad Ciudadano | Mapeo de los atributos que conforman la entidad SIR\_NE\_CIUDADANO. |
| LDCDN00 | Carga de los registros normalizados acorde a los atributos de la entidad Ciudadano. | Actualización de los registros normalizados de la entidad ciudadano en la entidad SIR\_NE\_CIUDADANO del modelo SIR, acorde a la igualdad de la llave primaria (ID\_CIUDADANO). |
| LDC\_OK | Visualización de los registros actualizados en la entidad Ciudadano. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_CIUDADANO en una tabla temporal. |
| TOCCDN\_OK | Generación de informe de los registros actualizados en la entidad Ciudadano. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_CIUDADANO en un archivo txt. (La ruta del texto de salida es determinada por el usuario: F\DOCUMENTOS\_SALIDA\SIR\_NE\_CIUDADANO\_OK.txt) |
| LDC\_NOK | Visualización de los registros No actualizados en la entidad Ciudadano. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_CIUDADANO en una tabla temporal. |
| TOCCDN\_OK | Generación de informe de los registros No actualizados en la entidad Ciudadano. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_CIUDADANO en un archivo txt. (La ruta del texto de salida es determinada por el usuario: F\DOCUMENTOS\_SALIDA\SIR\_NE\_CIUDADANO\_NOK.txt) |

Tabla 4. Flujo ETL\_PRINCIPAL\_VF – Fase Hilo Ciudadano.

**Modelo visual fase Hilo Individual Ciudadano:** ETL\_PRINCIPAL\_VF: Fase enfocada en migrar el bloque de información normalizada según el círculo de trabajo determinado para la entidad SIR\_NE\_CIUDADANO.

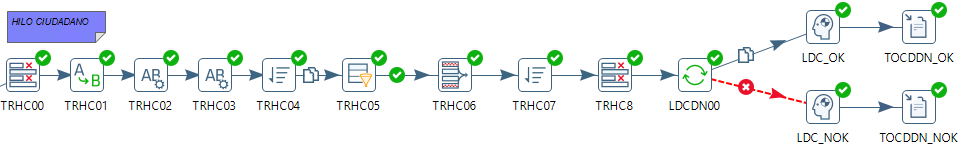


Figura 12. ETL\_PRINCIPAL\_VF – Fase Hilo Individual Entidad Ciudadano.

|  |  |  |
| --- | --- | --- |
| **PROCESO** | **OBJETIVO** | **IMPLEMENTACIÓN** |
| FASE HILO | INDIVIDUAL | ENTIDAD DIRECCIÓN (Versión beta) |
| TRHD00 | Inicializar la entidad Dirección. | Selección de los atributos normalizados de la entidad Dirección acorde al SIR\_NE\_DIRECCION. |
| TRHD01 | Eliminar caracteres de control. | Parametrización de los caracteres de control que deben eliminarse en los atributos normalizados. |
| TRHD02 | Establecer el criterio de diferenciación de los atributos normalizados y aplicación de utilidades de verificación. | Parametrización de formato “Mayúsculas” como criterio de diferenciación de los atributos normalizados y verficación de espacios al inicio o final del registro. |
| TRHD03 | Ordenar los objetos de la entidad Dirección. | Identificación del atributo ID\_MATRICULA como llave foránea, a fin de realizar un ordenamiento ascendente. |
| TRHD04 | Establece un atributo de control e integración. | Genera un atributo de control compuesto por el ID\_MATRICULA\_ID\_CIUDADANO. |
| TRHD05 | Establecer los atributos normalizados. | Selecciona los atributos requeridos por el modelo físico del SIR y verifica la correcta normalización (nombre, tipo y longitud). |
| TRHD06 | Ordenar los objetos de la entidad dirección. | Identificación del atributo ID\_Ciudadano como llave primaria, a fin de realizar un ordenamiento ascendente. |
| TRHD07 | Verificación de no existencia de objetos vacíos. | Filtro de los objetos vacío con base al atributo normalizado ID\_MATRICULA (condición ID\_MATRICULA IS NOT NULL). |
| TRHD08 | Verificación de no existencia de objetos vacíos. | Filtro de los objetos vacío con base al atributo normalizado ID\_CIUDADANO (condición ID\_CIUDADANO IS NOT NULL). |
| TRHD09 | Verificación de no existencia de objetos duplicados. | Eliminación de los objetos duplicados con base al atributo de control e integración (ID\_MATRICULA\_ID\_CIUDADANO). |
| TRHD10 | Identificación del segundo atributo primario no contenido en el tabloide de información del ODS. | Generación del atributo ID\_DIRECCION (llave foránea) no contenido en el tabloide normalizado cargado al ODL acorde al modelo físico de la entidad SIR\_NE\_DIRECCION (ID\_DIRECCION = 1). |
| TRHD11 | Validación de los atributos que conforman la entidad SIR\_NE\_DIRECCION. | Mapeo de los atributos que conforman la entidad SIR\_NE\_DIRECCION. |
| LDDC00 | Carga de los registros normalizados acorde a los atributos de la entidad SIR\_NE\_DIRECCION. | Actualización de los registros normalizados de la entidad ciudadano en la entidad SIR\_NE\_DIRECCION del modelo SIR, acorde a la igualdad de la llave primaria y foránea (ID\_DIRECCION y ID\_MATRICULA). |
| LDDC\_OK | Visualización de los registros actualizados en la entidad Dirección. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_DIRECCION en una tabla temporal. |
| TOFC\_OK | Generación de informe de los registros actualizados en la entidad Dirección. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_DIRECCION en un archivo txt. (La ruta del texto de salida es determinada por el usuario: F\DOCUMENTOS\_SALIDA\ SIR\_NE\_DIRECCION \_OK.txt) |
| LDDC\_NOK | Visualización de los registros No actualizados en la entidad Dirección. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_DIRECCION en una tabla temporal. |
| TODC\_NOK | Generación de informe de los registros No actualizados en la entidad Dirección. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_DIRECCION en un archivo txt. (La ruta del texto de salida es determinada por el usuario: F\DOCUMENTOS\_SALIDA\ SIR\_NE\_DIRECCION\_NOK.txt) |

Tabla 6. Flujo ETL\_PRINCIPAL\_VF – Fase Hilo Dirección.

**Modelo visual fase Hilo Individual Dirección:** ETL\_PRINCIPAL\_VF: Fase enfocada en migrar el bloque de información normalizada según el círculo de trabajo determinado para la entidad SIR\_NE\_DIRECCION.

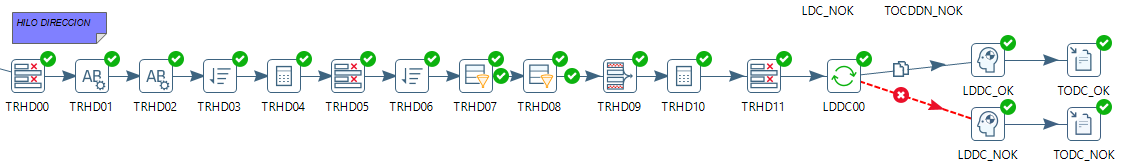


Figura 13. ETL\_PRINCIPAL\_VF – Fase Hilo Individual Entidad Dirección.

|  |  |  |
| --- | --- | --- |
| **PROCESO** | **OBJETIVO** | **IMPLEMENTACIÓN** |
| FASE HILO | INDIVIDUAL | ENTIDAD FOLIO (AREA DEL PREDIO) (Versión beta) |
| TRHAP00 | Inicializar la entidad Folio. | Selección de los atributos normalizados de la entidad Área Predio acorde al SIR\_NE\_FOLIO. |
| TRHAP01 | Eliminar caracteres de control. | Parametrización de los caracteres de control que deben eliminarse en los atributos normalizados. |
| TRHAP02 | Establecer el criterio de diferenciación de los atributos normalizados y aplicación de utilidades de verificación. | Parametrización de formato “Mayúsculas” como criterio de diferenciación de los atributos normalizados y verficación de espacios al inicio o final del registro. |
| TRHAP03 | Ordenar los objetos de la entidad Folio. | Identificación del atributo ID\_MATRICULA como llave primaria, a fin de realizar un ordenamiento ascendente. |
| TRHAP04 | Verificación de no existencia de objetos duplicados. | Eliminación de los objetos duplicados con base al atributo de control e integración (ID\_MATRICULA). |
| TRHAP05 | Ordenar los objetos de la entidad Folio. | Identificación del atributo ID\_MATRICULA como llave primaria, a fin de realizar un ordenamiento ascendente. |
| LDHAP00 | Carga de los registros normalizados acorde a los atributos de la entidad SIR\_NE\_FOLIO. | Actualización de los registros normalizados de la entidad ciudadano en la entidad SIR\_NE\_FOLIO del modelo SIR, acorde a la igualdad de la llave primaria (ID\_MATRICULA). |
| LDAP\_OK | Visualización de los registros actualizados en la entidad Folio. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_FOLIO en una tabla temporal. |
| TOAR\_OK | Generación de informe de los registros actualizados en la entidad Folio. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_FOLIO en un archivo txt. (La ruta del texto de salida es determinada por el usuario:  F\DOCUMENTOS\_SALIDA\ SIR\_NE\_FOLIO\_OK.txt) |
| LDAP\_NOK | Visualización de los registros No actualizados en la entidad Folio. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_FOLIO en una tabla temporal. |
| TOAP\_NOK | Generación de informe de los registros No actualizados en la entidad Folio. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_DIRECCION en un archivo txt. (La ruta del texto de salida es determinada por el usuario: F\DOCUMENTOS\_SALIDA\ SIR\_NE\_ FOLIO\_NOK.txt). |

Tabla 7. Flujo ETL\_PRINCIPAL\_VF – Fase Hilo Folio.

**Modelo visual fase Hilo Individual Folio:** ETL\_PRINCIPAL\_VF: Fase enfocada en migrar el bloque de información normalizada según el círculo de trabajo determinado para la entidad SIR\_NE\_FOLIO.

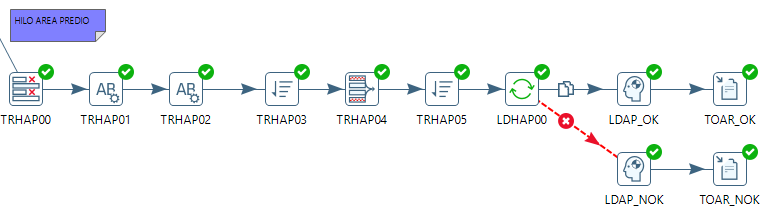


Figura 14. ETL\_PRINCIPAL\_VF – Fase Hilo Individual Entidad Folio (Área del Predio).

|  |  |  |
| --- | --- | --- |
| **PROCESO** | **OBJETIVO** | **IMPLEMENTACIÓN** |
| FASE HILO | INDIVIDUAL | ENTIDAD DVIPOLA (Versión estable) |
| TRHDV00 | Inicializar la entidad DIVIPOLA. | Selección de los atributos normalizados de la entidad DIVIPOLA acorde al SIR\_NE\_MUNICIPIO y SIR\_NE\_DEPARTAMENTO. |
| TRHDV01 | Eliminar caracteres de control. | Parametrización de los caracteres de control que deben eliminarse en los atributos normalizados. |
| TRHDV02 | Identificación de los atributos de identificación del Municipio y Departamento. | Parametrización de la longitud de caracteres que constituyen el ID\_MUNICIPIO y ID\_DEPARTAMENTO. |
| TRHDV03 | Establecer el criterio de diferenciación de los atributos normalizados y aplicación de utilidades de verificación. | Parametrización de formato “Mayúsculas” como criterio de diferenciación de los atributos normalizados y verficación de espacios al inicio o final del registro. |
| TRHDV04 | Establecer los atributos normalizados. | Selecciona los atributos requeridos por el modelo físico del SIR y verifica la correcta normalización (nombre, tipo y longitud). |
| TRHDV05 | Ordenar los objetos de la entidad Divipola. | Identificación del atributo ID\_MUNICIPIO como llave primaria, a fin de realizar un ordenamiento ascendente. |
| TRHDV06 | Verificación de no existencia de objetos vacíos. | Filtro de los objetos vacío con base al atributo normalizado ID\_MUNICIPIO (condición ID\_MUNICIPIO IS NOT NULL). |
| TRHDV07 | Verificación de no existencia de objetos vacíos. | Filtro de los objetos vacío con base al atributo normalizado ID\_DEPARTAMENTO (condición ID\_DEPARTAMENTO IS NOT NULL). |
| TRHDV08 | Ordenar los objetos de la entidad Divipola. | Identificación del atributo ID\_MUNICIPIO como llave primaria, a fin de realizar un ordenamiento ascendente. |
| TRHDV09 | Verificación de no existencia de objetos duplicados. | Eliminación de los objetos duplicados con base al atributo normalizado ID\_MUNICIPIO. |
| TRHDV10 | Ordenar los objetos de la entidad Divipola. | Identificación del atributo ID\_MUNICIPIO como llave primaria, a fin de realizar un ordenamiento ascendente. |
| TRHDV11 | Validación de los atributos que conforman la entidad DIVIPOLA. | Mapeo de los atributos que conforman la entidad DIVIPOLA. |
| LDMNC00 | Carga de los registros normalizados acorde a los atributos de la entidad SIR\_NE\_MUNICIPIO. | Actualización de los registros normalizados de la entidad ciudadano en la entidad SIR\_NE\_MUNICIPIO del modelo SIR, acorde a la igualdad de la llave primaria (ID\_MUNICIPIO). |
| LDMNC\_OK | Visualización de los registros actualizados en la entidad MUNICIPIO. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_MUNICIPIO en una tabla temporal. |
| TOCCDN\_OK | Generación de informe de los registros actualizados en la entidad MUNICIPIO. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_MUNICIPIO en un archivo txt. (La ruta del texto de salida es determinada por el usuario:  F\DOCUMENTOS\_SALIDA\SIR\_NE\_SIR\_NE\_MUNICIPIO\_OK.txt) |
| LDMNC\_NOK | Visualización de los registros No actualizados en la entidad Municipio. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_MUNICIPIO en una tabla temporal. |
| TOMNC\_NOK | Generación de informe de los registros No actualizados en la entidad Municipio. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_MUNICIPIO en un archivo txt. (La ruta del texto de salida es determinada por el usuario: F\DOCUMENTOS\_SALIDA\ SIR\_NE\_MUNICIPIO \_NOK.txt) |
| LDDEP00 | Carga de los registros normalizados acorde a los atributos de la entidad Departamento. | Actualización de los registros normalizados de la entidad ciudadano en la entidad SIR\_NE\_DEPARTAMENTO del modelo SIR, acorde a la igualdad de la llave primaria (ID\_ DEPARTAMENTO). |
| LDDEP\_OK | Visualización de los registros actualizados en la entidad Departamento. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_DEPARTAMENTO en una tabla temporal. |
| TODEP\_OK | Generación de informe de los registros actualizados en la entidad Departamento. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_DEPARTAMENTO en un archivo txt. (La ruta del texto de salida es determinada por el usuario:  F\DOCUMENTOS\_SALIDA\ SIR\_NE\_DEPARTAMENTO\_OK.txt) |
| LDDEP\_NOK | Visualización de los registros No actualizados en la entidad Departamento. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_DEPARTAMENTO en una tabla temporal. |
| TODEP\_NOK | Generación de informe de los registros No actualizados en la entidad Departamento. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_DEPARTAMENTO en un archivo txt. (La ruta del texto de salida es determinada por el usuario:  F\DOCUMENTOS\_SALIDA\SIR\_NE\_MUNICIPIO \_NOK.txt) |
| LDZNR00 | Carga de los registros normalizados acorde a los atributos de la entidad Zona Registral. | Actualización de los registros normalizados de la entidad ciudadano en la entidad SIR\_NE\_ZONA\_REGISTRAL del modelo SIR, acorde a la igualdad de la llave primaria y foránea:  (ID\_MUNICIPIO y ID\_ DEPARTAMENTO). |
| LDZNR\_OK | Visualización de los registros actualizados en la entidad Zona Registral. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_ZONA\_REGISTRAL en una tabla temporal. |
| TOZNR\_OK | Generación de informe de los registros actualizados en la entidad Zona Registral. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_ZONA\_REGISTRAL en un archivo txt. (La ruta del texto de salida es determinada por el usuario:  F\DOCUMENTOS\_SALIDA\ SIR\_NE\_ZONA\_REGISTRAL\_OK.txt) |
| LDZNR\_NOK | Visualización de los registros No actualizados en la Zona Registral. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_ZONA\_REGISTRAL en una tabla temporal. |
| TOZNR\_NOK | Generación de informe de los registros No actualizados en la Zona Registral. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_DEPARTAMENTO en un archivo txt. (La ruta del texto de salida es determinada por el usuario:  F\DOCUMENTOS\_SALIDA\SIR\_NE\_ZONA\_REGISTRAL\_NOK.txt) |
| LDVER00 | Carga de los registros normalizados acorde a los atributos de la entidad Vereda. | Actualización de los registros normalizados de la entidad Vereda en la entidad SIR\_NE\_VEREDA del modelo SIR, acorde a la igualdad de la llave primaria: (ID\_MUNICIPIO). |
| LDVER\_OK | Visualización de los registros actualizados en la entidad Vereda. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_VEREDA en una tabla temporal. |
| TOVER\_OK | Generación de informe de los registros actualizados en la entidad Vereda. | Presentación de todos los registros normalizados actualizados en los atributos de la entidad SIR\_NE\_VEREDA en un archivo txt. (La ruta del texto de salida es determinada por el usuario:  F\DOCUMENTOS\_SALIDA\ SIR\_NE\_VEREDA\_OK.txt) |
| LDVER\_NOK | Visualización de los registros No actualizados en la entidad Vereda. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_VEREDA en una tabla temporal. |
| TOVER\_NOK | Generación de informe de los registros No actualizados en la Zona Registral. | Presentación de todos los registros normalizados No actualizados en los atributos de la entidad SIR\_NE\_ VEREDA en un archivo txt. (La ruta del texto de salida es determinada por el usuario:  F\DOCUMENTOS\_SALIDA\SIR\_NE\_ VEREDA\_NOK.txt) |

Tabla 8. Flujo ETL\_PRINCIPAL\_VF – Fase Hilo Divipola.

**Modelo visual fase Hilo Individual Divipola:** ETL\_PRINCIPAL\_VF: Fase enfocada en migrar el bloque de información normalizada según el círculo de trabajo determinado para las entidades SIR\_NE\_MUNICIPIO, SIR\_NE\_DEPARTAMENTO, SIR\_NE\_ZONA\_REGISTRAL y SIR\_NE\_VEREDA.

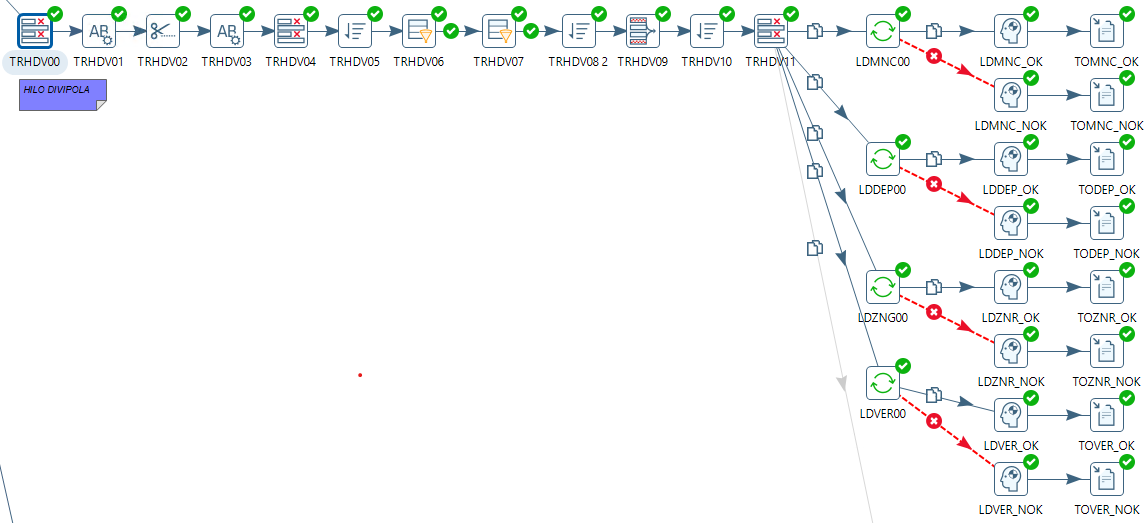


Figura 15. ETL\_PRINCIPAL\_VF – Fase Hilo Individual Entidad Divipola.

**Modelo visual ETL integral:** ETL\_PRINCIPAL\_VF: ETL integrado a todas sus fases de migración en bloque de información normalizada según el círculo de trabajo determinado.

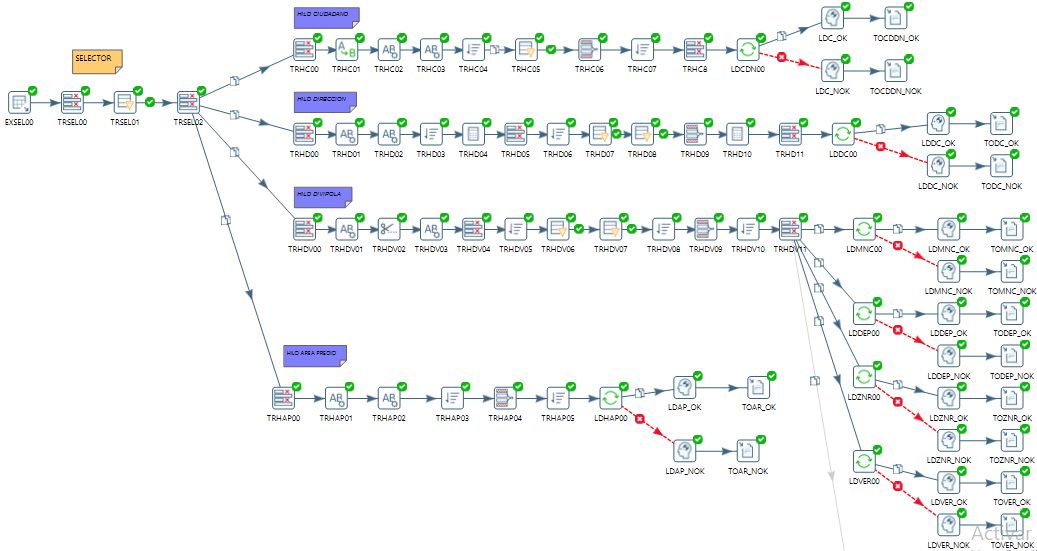


Figura 16. ETL\_PRINCIPAL\_VF – Integración de fases.

**Archivo del ETL PRINCIPAL:** ETL\_PRINCIPAL\_VF.ktr

**Repositorio de desarrollo del proyecto:** https://github.com/keppler22/ETL-PROJECT

# ANALISIS Y RESULTADOS

Para el análisis de las pruebas de desempeño de los ETL son enfocados en métricas de tiempo de ejecución de cada artefacto de fase y la cantidad de registros procesados por la misma.

El escenario definido está sujeto al bloque de información suministrado por el equipo de análisis de datos, bloque constituido por una muestra de 299 objetos referenciados en los siguientes atributos de las entidades del modelo SIR normalizados: *a)*numero\_inmueble, *b)*tipo\_documento\_propietario, *c)* identificacion\_propietario,*d)*nombre\_propietario,*e)*primer\_apellido, *)*segundo\_apellido, *g)*sexo, )tipo\_via\_abreav, *i)*nombre\_principal, *j)*letra\_1p, *k)*bisp, *l)*letra\_2p, *m)* cuadrante\_p, *n)* nombre\_generadora, *o)* letra\_1g, *p)* bisg, *q)* letra\_2g, *r)* número, *s)* cuadrante, *t)*complemento, *u)*domicilio\_comprimido, *v)*departamento, *w)* municipio, *x)* cod\_postal, *y)* divipola, *z)* departamento y aa)municipio.

Las métricas presentadas a continuación, son muestras extraídas de las utilidades de análisis que dispone la herramienta:

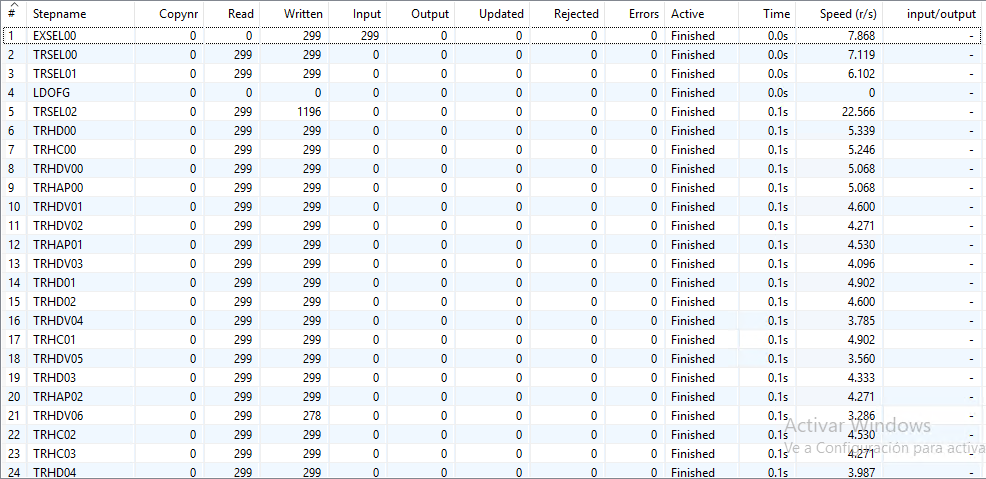


Figura 17. Muestra del análisis de Métricas del ETL\_PRINCIPAL\_VF disponibles en la herramienta.

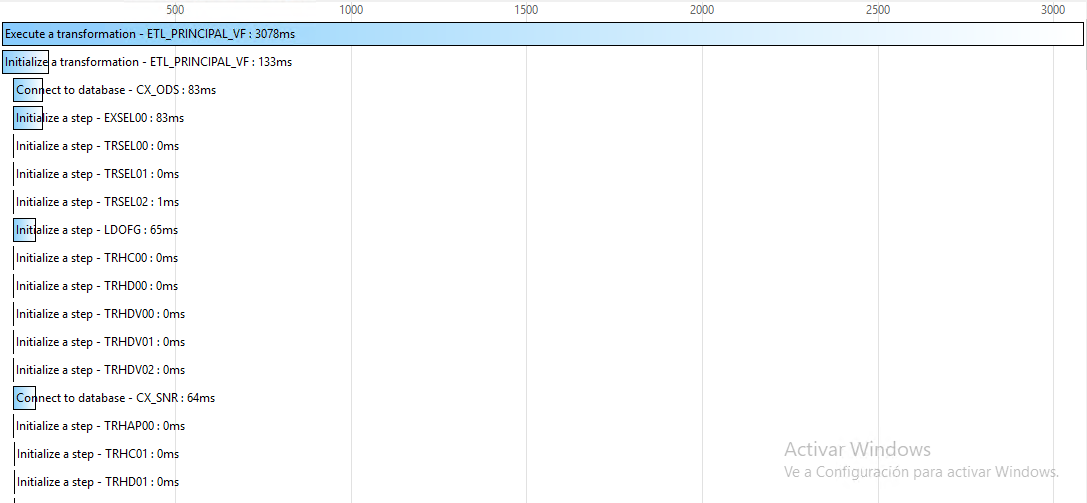


Figura 18. Muestra del análisis de Métricas del ETL\_PRINCIPAL\_VF disponibles en la herramienta.

De las diferentes métricas de desempeño del ETL construido y disponibles en la herramienta se puede inferir los siguiente:

1. Los ETL en etapa de desarrollo estable realizan la ejecución exitosa de las diferentes fases de transformación y verificación de los atributos normalizados de las entidades del modelo, dada la fácil interpretación del modelo lógico del SIR disponible.

Resultado 1: De la muestra de 299 objetos, se evidencio que existen 127 ciudadanos normalizados, de los 48 fueron actualizados dada la condición de igualdad del ID\_CIUDADANO en la entidad SIR\_NE\_CIUDADANO.

Resultado 2: De la muestra de 299 objetos, se evidencio que existen 35 municipios y departamentos normalizados, de los cuales y 33 fueron actualizados dada la condición de igualdad del ID\_MUNICIPIO en las entidades SIR\_NE\_MUNICIPIO, SIR\_NE\_DEPARTAMENTO, SIR\_NE\_ZONA\_REGISTRAL.

1. Los ETL en etapa de desarrollo beta realizan la ejecución exitosa de las diferentes fases de transformación y verificación de los atributos normalizados de las entidades del modelo, pero no realizan la correcta actualización dada la complejidad de interpretación del modelo lógico del SIR disponible.

Resultado 3: De la muestra de 299 objetos, se evidencio que existen 299 direcciones normalizadas, de los cuales 299 no son actualizados dada la condición de igualdad del ID\_MATRICULA en la entidad SIR\_NE\_DIRECCION y diferentes entidades relacionadas en el modelo físico disponible.

# CONCLUSIONES

* La construcción del ETL requerido se desarrolló de manera asertiva bajo la metodología incremental. Evidenciado por los requerimientos establecidos de manera progresiva y la disponibilidad de la información del modelo, los cuales han orientado y permitido una construcción efectiva y flexible del ETL.
* La metodología de desarrollo incremental permite la implementación de características y funcionalidades de manera gradual, lo que ha permitido adaptarse a los cambios en requeridos.
* La metodología permite la evidenciar una mayor visibilidad y control sobre el progreso del proyecto, lo que ha resultado en una ejecución más efectiva y eficiente.
* Para finalizar el despliegue del ETL integral de migración se hace necesario disponer de una abstracción clara del modelo del SIR, desde una perspectiva ejecución funcional del sistema.
* Es necesario disponer del tabloide de información completo, dado que al no contar con todos los atributos fue necesario definirlos dado el análisis de los datos disponibles en la base de pruebas del SIR.
* El ETL de migración realizado logró un avance del 80% en el tiempo disponible para su desarrollo. Aunque aún queda un 20% por completar, los resultados hasta ahora demuestran un buen progreso y una efectiva gestión del tiempo y recursos disponibles. Es importante destacar que el éxito en la implementación de un ETL depende tanto del avance en el desarrollo como de la calidad y precisión de los datos migrados. Por lo tanto, es necesario continuar evaluando y asegurándose de que los datos migrados cumplan con los estándares requeridos antes de continuar con el proceso. En resumen, aunque aún hay un camino por recorrer, los resultados hasta ahora son alentadores y sugieren una buena dirección para el proyecto.