



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

INFORME DE ACTIVIDADES

CONSULTAS GUBERNAMENTALES PARA VUCEM

PARA OBTENER EL TÍTULO DE
INGENIERO EN COMPUTACIÓN

PRESENTA

PAOLA IVETTE GUTIÉRREZ GRAJALES

ASESOR DE INFORME

M. C. ALEJANDRO VELÁZQUEZ MENA

CIUDAD UNIVERSITARIA SEPTIEMBRE 2015



AGRADECIMIENTOS	4
INTRODUCCIÓN	5
CAPITULO I	6
ORGANIZACIÓN DE LA EMPRESA	6
1.1 <i>Visión</i>	6
1.2 <i>Organigrama de la empresa</i>	6
1.3 <i>Organigrama del área</i>	7
1.4 <i>Cánones de Softtek</i>	7
CAPITULO II	9
PARTICIPACIÓN EN PROYECTOS.....	9
CAPITULO III	10
INTELIGENCIA DE NEGOCIOS.....	10
3.1 <i>VUCEM (Ventanilla Única del Comercio Exterior Mexicano)</i>	11
3.2 <i>Definición del Problema</i>	12
3.3 <i>Requerimiento del cliente</i>	12
3.3.1 <i>Generación de consulta de información</i>	12
3.4 <i>Desarrollo de BI</i>	13
3.4.1 <i>ETL</i>	14
3.4.2 <i>Área de Staging</i>	17
3.4.3 <i>ODS</i>	17
3.4.4 <i>Datamart</i>	18
3.5 <i>IBM Websphere Datastage</i>	20
3.6 <i>Unidad de negocio 1</i>	22
3.6.1 <i>ODS Unidad de negocio 1</i>	22
3.6.1.1 <i>Extracciones Unidad de negocio 1</i>	22
3.6.1.2 <i>Transformaciones Unidad de negocio 1</i>	22
3.6.1.3 <i>Cargas Unidad de negocio 1</i>	24
3.6.1.4 <i>Secuencias ODS Unidad de negocio 1</i>	26
3.6.2 <i>DataMart Unidad de negocio 1</i>	28
3.6.2.1 <i>Transformaciones Unidad de negocio 1</i>	28
3.6.2.2 <i>Cargas Unidad de negocio 1</i>	31
3.6.2.3 <i>Secuencias Datamart Unidad de negocio 1</i>	33
3.7 <i>Unidad de negocio 2</i>	35
3.7.1 <i>ODS Unidad de negocio 2</i>	35
3.7.1.1 <i>Extracciones Unidad de negocio 2</i>	35
3.7.1.2 <i>Transformaciones Unidad de negocio 2</i>	35
3.7.1.3 <i>Cargas Unidad de negocio 2</i>	36
3.7.1.4 <i>Secuencia ODS Unidad de negocio 2</i>	38
3.7.2 <i>DataMart Unidad de negocio 2</i>	38
3.7.2.1 <i>Transformaciones Unidad de negocio 2</i>	38
3.7.2.2 <i>Cargas Unidad de negocio 2</i>	40
3.7.2.3 <i>Secuencias Datamart Unidad de negocio 2</i>	40
3.8 <i>Unidad de negocio 3</i>	41
3.8.1 <i>ODS Unidad de negocio 3</i>	41
3.8.1.1 <i>Extracciones Unidad de negocio 3</i>	41
3.8.1.2 <i>Transformaciones Unidad de negocio 3</i>	41
3.8.1.3 <i>Cargas Unidad de negocio 3</i>	44
3.8.1.4 <i>Secuencia ODS Unidad de negocio 3</i>	47
3.8.2 <i>DataMart Unidad de negocio 3</i>	47

3.8.2.1 Transformaciones Unidad de negocio 3	47
3.8.2.2 Cargas Unidad de negocio 3	51
3.8.2.3 Secuencias Datamart Unidad de negocio 3	53
3.9 <i>Testing</i>	56
4.0 <i>Documentación de los procesos</i>	57
CAPITULO IV	59
ANÁLISIS DE RESULTADOS.....	59
CONCLUSIONES	62
ANEXOS	64
BIBLIOGRAFÍA	84

Agradecimientos

Con todo mi cariño y mi amor para las personas que hicieron todo lo posible para que yo pudiera lograr mis sueños, por motivarme y darme la mano cuando sentía que el camino se terminaba, a ustedes por siempre mi corazón y mi agradecimiento.

Gracias Charles, Genito y Sergio.

Agradezco enormemente la dedicación y apoyo de cada uno de mis sinodales.

Al M.C Eduardo Espinosa Ávila porque sin su forma de ver la carrera y la vida tan "sencilla", no me hubiera podido dar cuenta que en realidad así es, no hay que complicársela, solo es cuestión de actitud y solo un poco de aptitud en cada una de las cosas que realice. Gracias.

A la M.I Tanya Itzel Arteaga Ricci por todo el apoyo brindado desde el inicio en mi primer trabajo en el área de la carrera, hasta ahora, el cual ha sido incondicional. Gracias por la motivación y el optimismo de cada día.

Al Ing. Jorge Alberto Campos Rodríguez por tener la paciencia y dedicación con cada uno de nosotros sus alumnos, además de que sus enseñanzas han sido fundamentales para mi desarrollo profesional. Sin duda ud. ha sido uno de los mejores profesores en la facultad de ingeniería que he tenido. Gracias.

A la Ing. Gabriela Betzabé Lizarraga Ramírez porque sin el conocimiento que me brindó hubiera sido más complicado entender cada proceso y desarrollo realizado en el trabajo de éste informe. Gracias.

Quiero agradecer a todos mis maestros que en este andar por la vida, influyeron con sus lecciones y experiencias en formarme como una persona de bien y preparada para los retos que pone la vida, a todos y cada uno de ellos, gracias.

Por último quiero agradecer a todos mis amigos y amigas con los que día a día nos apoyábamos para ser mejores cada día.

Estoy segura que mis metas planteadas darán fruto en el futuro y por ende me esforzaré cada día para ser mejor en todo lugar sin olvidar el respeto que engrandece a la persona.

Introducción

El informe consta de 4 capítulos en los cuales explico mi experiencia profesional hasta ahora, siendo más específica en el proyecto donde mi desarrollo profesional ha sido mayor, sobre todo en el área de bases de datos, el cual fue el módulo que escogí en la universidad.

El capítulo I es una breve contextualización de Softtek, la empresa en la que adquirí experiencia profesional. Describo los inicios de Softtek como empresa, su visión, como también el panorama general de la empresa y su área de inteligencia de negocios. Así mismo, explico parte de la cultura Softtek en la que se basa, los cuales son sus cánones de acción, ya que en ellos está expresado una guía para mejorar como individuos, pero sobre todo, son los que determinan el accionar como equipo y permite que la organización evolucione.

El capítulo II trata de toda la experiencia profesional que he adquirido, indicando los diversos puestos en los que me he desarrollado laboralmente y participado activamente, describiendo las funciones que desempeño, además de hacer mención de las herramientas que más se utilizan en cada función.

En el capítulo III es en donde explico mi experiencia profesional, en la que adquirí mayor conocimiento, desarrollado en el área de inteligencia de negocios (con siglas BI, del inglés *Business Intelligence*), teniendo como clientes a 18 unidades de negocio cada una con requisitos específicos, ya que cada una hace una labor distinta y por lo mismo, la idea o estructura del negocio varía. Participé construyendo parte del almacén operacional de datos (con siglas ODS, del inglés *Operational Data Store*) y construyendo almacenes de datos, los cuales son específicos para cada unidad de negocio (del inglés *Datamarts*) de 3 de esas unidades de negocio.

Por último, en el capítulo IV anexo detalles de los resultados y conclusiones a los que llegué al desempeñar tanto mis funciones laborales como profesionales, además de los beneficios que obtendrá el cliente y la experiencia que obtuve con el proyecto.

CAPITULO I

Organización de la empresa

1.0 Softtek

Softtek es una empresa global que provee soluciones y servicios de tecnología, que ayuda a los clientes a prosperar en la actual era digital.

Fué fundada en Diciembre de 1982 en Monterrey por Gerardo López, el cual reunió a varios ejecutivos (entre ellos a la actual directora general Blanca Treviño) para formar un equipo profesional que se dedicara a proveer servicios de TI (del acrónimo *Tecnologías de la Información*).

Softtek empezó como un paraguas de unidades de negocio separadas, es decir, cada accionista enfocaba sus esfuerzos hacia objetivos diferentes y una de las ventajas de ésto es que cada proyecto era analizado y llevado a cabo de forma individual y, si fallaba, sólo afectaba a la unidad de negocio en cuestión.

Actualmente Softtek tiene una presencia mundial, ya que tiene centros de desarrollo en América, Europa y Asia con aproximadamente 9000 colaboradores alrededor del mundo.

1.1 Visión

Trascender como proveedor global líder en soluciones de TI y procesos de negocio, generando relaciones mutuamente benéficas, de largo plazo y cimentadas en una base de confianza ganada, además de construir nuestro futuro siendo una empresa sólida y socialmente responsable. Proporciona servicios innovadores y de alta calidad, impulsados por la pasión de nuestra cultura centrada en el elemento humano.

1.2 Organigrama de la empresa

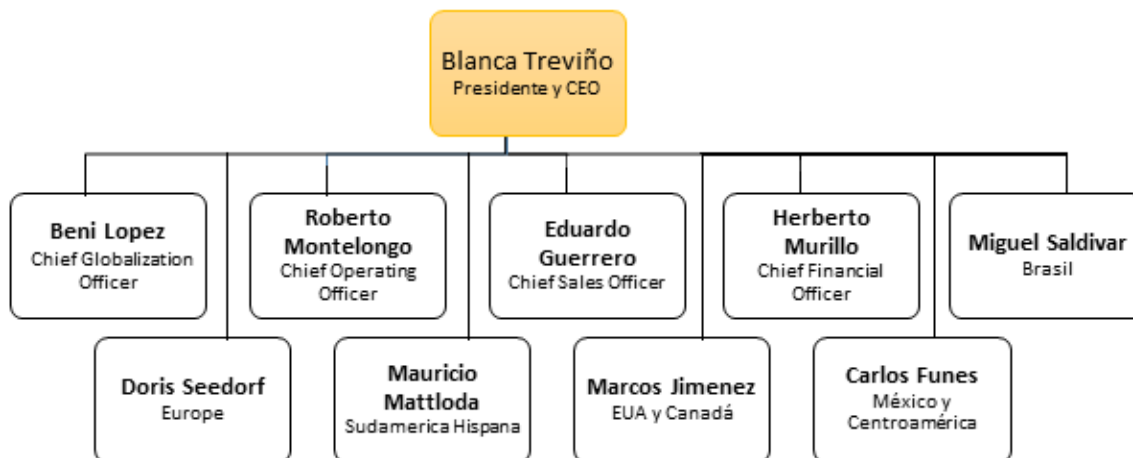


Figura 1.0

1.3 Organigrama del área

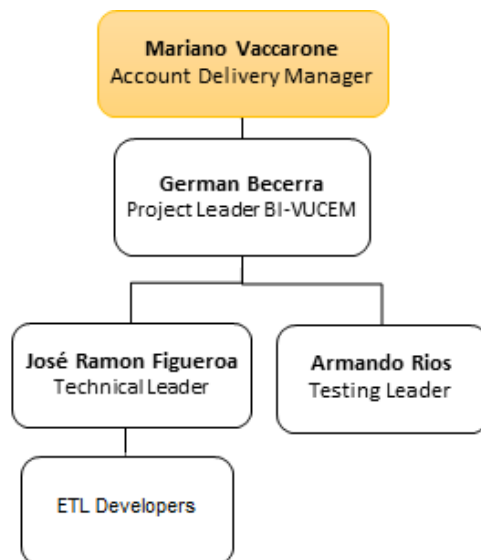


Figura 1.1

1.4 Cánones de Softtek

En Softtek contamos con un código de ética y conducta, en la que los colaboradores de la organización nos basamos todos los días para lograr una conducta profesional de una manera integral, transparente, honesta y responsable. La cultura Softtek se basa en los 11 cánones de acción que constituyen una parte esencial de nuestro código. El accionar como equipo bajo los mismos cánones permite que la organización evolucione.

Apertura.- Soy abierto cuando interactúo con las personas sin retener ni distorsionar lo que pienso y siento. Cuanto más consciente soy de lo que pasa conmigo, más claramente puedo identificar mis preferencias y temores; de esta forma genero relaciones más efectivas.

Autodeterminación.- Me auto determino cuando me siento responsable de lo que me sucede y reconozco la capacidad que tengo de elegir cada una de mis acciones, reacciones, sentimientos y auto concepto. Desde esta perspectiva, transformo mi vida, no culpo a otros y asumo mi responsabilidad para diseñar la persona que quiero ser.

Compromiso.- Soy comprometido cuando cumplo todas las promesas que hago, encargándome de hacer lo necesario para que ocurra lo que prometí.

Confianza.- Confío en las personas cuando creo que actuarán de buena fe. Cuando confío en una persona, le doy la oportunidad de creer en sí misma y superarse; así obtengo mayor tranquilidad y fortaleza en nuestra relación. La confianza se otorga, no se gana.

Entusiasmo.- Soy entusiasta cuando estoy en búsqueda de retos continuos, cada vez más ambiciosos.

Flexibilidad.- Soy flexible cuando me adapto a diferentes ambientes y personas. Ser flexible me da la posibilidad de dialogar, crear alternativas y adecuarlas al contexto presente.

Perseverancia.-Soy perseverante cuando estoy seguro de que alcanzaré mis metas y tengo la energía para permanecer en el empeño hasta que tarde o temprano lo consiga.

Respeto.- Respeto a otras personas cuando me relaciono con ellas de forma efectiva, aun cuando tengan creencias, preferencias y comportamientos diferentes a los míos. En la medida en que las respeto, dejo de imponer mis puntos de vista y aprendo más de las distintas creencias y percepciones que existen a mí alrededor.

Sinergia.- Logro sinergia cuando uno mis esfuerzos a los de un grupo para multiplicar nuestras fortalezas y alcanzar grandes metas, más allá de la suma de los esfuerzos individuales.

Sociedad.- Soy socio cuando mi visión personal coincide con la de Softtek y actúo como dueño y parte de la organización, buscando crear un patrimonio sólido, a través de compartirlo y multiplicarlo con otras personas. Ser socio me transforma en emprendedor y visionario, apostando e invirtiendo en Softtek con la firme idea de trascender.

Visión.- Construyo mi visión cuando identifico mis sueños y estoy dispuesto a entregar lo mejor de mí para hacerlos realidad. Cuando soy claro en mi visión y mis acciones, trasciendo, encuentro la fuerza para movilizarme, supero la adversidad y contribuyo de forma positiva a la sociedad.

CAPITULO II

Participación en proyectos

Grupo modelo - Softtek - Actual

Actualmente estoy en el área de inteligencia de negocios (con siglas BI, en inglés *Business Intelligence*) e IM (en inglés *Integration Management*). Soy analista y desarrolladora de ETL Extracción, Transformación y Carga (del inglés *Extract, Transform, Load data*) donde estoy desarrollando con herramientas como SAPBODS, Tableau, por mencionar algunas.

SAPBODS (Del inglés *SAP Business Objects Data Services*). - Es una herramienta de ETL adquirido por SAP y usada para la integración de diversos sistemas: extrayendo, transformado y cargando información útil. Está estrechamente integrado con los sistemas SAP y por lo tanto es una buena herramienta para migrar datos de los sistemas legados a los sistemas SAP con facilidad, con menor desarrollo y esfuerzo, depuración eficaz y capacidad de monitoreo.

Tableau.- Es una herramienta que ayuda a las personas a ver y comprender datos, también para que analicen, visualicen y compartan información rápidamente, el cual se utiliza para generar reportes.

Soporte a producción - Softtek

Desarrollé una modificación para una pantalla de la cuenta de VUCEM (del acrónimo *Ventanilla Única del Comercio Exterior Mexicano*) utilizando los lenguajes de programación java, jquery, javascript, para que la página permitiera hacer la consulta de réplicas con la cual cuando el usuario no estaba de acuerdo con la observación del trámite que solicitaba, podía dar detalles del por que de su trámite y con ello se volvía a evaluar su situación.

Generé y resolví algunos reportes que los usuarios solicitaban así como requerimientos de usuarios que tuvieran problemas en la VUCEM, ya fuera de generación de formato, modificación de algún dato erróneo con sus datos, asignación de roles, etc.

Documentación - Softtek

Realicé ingeniería inversa para diferentes procesos, entre ellos los servicios web de DGN (del acrónimo *Dirección General de Normas*), para el proceso de privados que englobaba al proceso de aéreos, de tránsitos y de recintos fiscalizados, todo se hacía con base a UML (del inglés *Unified Modeling Language*) y utilizando herramientas como StarUML, UMLab, Visual Paradigm, Enterprise Architect.

CAPITULO III

Inteligencia de Negocios

El término de inteligencia de negocios (BI, en inglés *Business Intelligence*) es un agregado de aplicaciones y herramientas enfocadas al procesamiento de los datos o información, de la cual se puede obtener conocimiento relevante para así poder hacer una toma de decisiones más acertada.

En 1989, el analista Howard Dresner popularizó *Business Intelligence* como un término general para describir un conjunto de conceptos y métodos para mejorar la toma de decisiones empresariales mediante el uso de sistemas de soporte a toma de decisiones basados en hechos.

El objetivo principal de BI es sustentar continuamente a las empresas de hoy en día, para mejorar su competitividad, desarrollo, crecimiento y maximizar su desempeño con el fin de crear estrategias funcionales, seguirlas, medir su ejecución y comparar el resultado frente a diversos escenarios y el entorno.

Para lograrlo es necesario que se establezca e identifique la necesidad del negocio, para que se brinde la información oportuna, de valor y en el contexto adecuado, todo esto puede ser para acelerar un proceso, reducir costos o mejorar productividad de algún área en particular.

Para la construcción de BI es importante la selección de la herramienta correcta, que vaya de la mano de acuerdo al negocio, ya que a pesar de que hay muchas herramientas de BI no todas tienen el mismo desempeño, enfoque o simplicidad de uso.

Otro punto muy importante para generar conocimiento valioso para el negocio es estandarizar, validar y asegurar la calidad de los datos, además del diseño del modelo a generar. Los sistemas de BI pueden llegar a tener un diseño muy diferente a un sistema transaccional, inclusive dependiendo del negocio puede llegar a estar desnormalizado, esto con el propósito de mejorar el desempeño de alguna consulta no planeada.

Por último pero no menos importante, es indispensable tener a un buen equipo de trabajo, para evitar re trabajo y por lo tanto poder hacer un buen desarrollo en tiempo y forma de cualquier desarrollo de BI.

En la actualidad hay mucha diversidad de software para el desarrollo de BI, que brinda a la empresa la habilidad de analizar de una forma rápida por qué pasan las cosas y enfocarse a patrones y amenazas.

La inteligencia de negocios puede ser utilizada en las diferentes áreas de la empresa, tales como, ventas, *marketing*, finanzas entre otras.

El desarrollo y utilización de la inteligencia de negocios va en aumento día a día, inclusive se prevé que en un futuro no muy lejano toda empresa u organización decidirá abandonar los métodos

tradicionales de manejo de información, para utilizar sistemas BI, con el fin de contar con una herramienta administrativa que ofrezca los elementos necesarios para establecer una estrategia adecuada que responda a los cambios del mercado en tiempo real.

3.1 VUCEM (Ventanilla Única del Comercio Exterior Mexicano)

La VUCEM es un sistema que permite la digitalización, registro y entrega de la información estandarizada (esto último se refiere a la forma de estructura de la información, de manera que sea entendible, amigable para el usuario y que cumpla con los estándares solicitados) a una única unidad de negocio para el cumplimiento de todos los requerimientos relacionados con la importación, exportación y tránsito de mercancías, contribuyendo al comercio exterior.

Algunos de los objetivos de la VUCEM son:

- Reducir tiempos en procesos administrativos, así como en las plataformas de revisión de las aduanas del país.
- Facilitar la búsqueda y acceso de información desde ubicaciones remotas a un repositorio central de información.
- Incidir en la competitividad de la economía del país con procesos de comercio exterior más ágiles, fáciles y seguros.
- Eliminar considerables gastos en mensajería y flete.
- Eliminar el papel, al convertir los procesos manuales en electrónicos utilizando documentación digitalizada.

En términos prácticos, la Ventanilla Única tiene como objetivo agilizar y simplificar los flujos de información entre el comercio y el gobierno y aportar beneficios significativos para todas las partes involucradas en el comercio transfronterizo.

Al unificar de manera digital todos los trámites de las unidades de negocio que pasan a través de las aduanas, nace la necesidad de los usuarios de las unidades de negocio de consultar la información que actualmente está entrando por la VUCEM por lo que para omitir el manejo de reportes y generación de tickets, así como para facilitar el acceso a la información para su análisis por parte de las unidades de negocio, es que nació en la VUCEM el proyecto de consultas gubernamentales (BI).

El proyecto de consultas gubernamentales (BI) de la VUCEM es principalmente un sistema reportador dinámico en el que cada usuario de las unidades de negocio puede acceder a la aplicación de la herramienta Cognos Query Studio sin tener que depender del proveedor de manera que lo pueden hacer en cualquier momento y con los campos e información que el usuario decida.

3.2 Definición del Problema

Analizar y construir los diferentes procesos automatizados correspondientes a cada una de las unidades de negocio contempladas en el proyecto, de acuerdo a las necesidades y el negocio que el cliente solicita.

3.3 Requerimiento del cliente.

3.3.1 Generación de consulta de información.

1) Generar una consulta con los datos de negocio respectivos de cada trámite de las unidades de negocio para obtener la información requerida. Esta debe incluir:

- Información propia del Negocio
- Datos de Solicitud
- Datos del Capturista Gubernamental
- Datos del Dictamen
- Datos de Observación
- Datos de la Resolución
- Datos del Trámite
- Datos del Funcionario

2) La consulta que se genere por parte del usuario puede contener todos o una parte de los datos definidos en los documentos de referencia respectivos a cada unidad de negocio, según los indique el usuario.

3) La información de los datos de negocio seleccionados en la consulta, podrá ser descargada en un archivo, el cual contendrá la información correspondiente a la consulta, así como una referencia al nombre de negocio al que pertenecen los datos.

4) Las consultas que el usuario realiza con mayor frecuencia, podrán ser almacenadas en VU (*Ventanilla Única*).

3.3.2 Configuración de Usuarios de Consulta

Creación de un módulo el cual permita la configuración de los usuarios de consulta a nivel de unidad de negocio y trámite.

Los campos donde se almacenará la información de los usuarios y su correspondiente configuración, será la que se encuentra definida en el Módulo de Administración de Usuarios de la Ventanilla Única.

Debe existir un único usuario que podrá tener acceso para hacer consultas de la información de cualquier trámite en cualquier unidad de negocio. Los demás usuarios serán asignados al nivel de la unidad de negocio, así como especificar a qué trámites tendrán acceso a consultar.

3.4 Desarrollo de BI

Un modelo de inteligencia de negocios busca presentar la información de una manera estándar, sencilla y sobre todo intuitiva para los usuarios, además de que permite accesos a la información mucho más rápida por parte de los manejadores de bases de datos, es un modelo que difiere bastante del modelo entidad-relación y que se enfoca en el objetivo del negocio.

En la siguiente tabla se puede ver una breve comparación entre un modelo entidad relación y uno dimensional, éste último es el que se utiliza para el desarrollo de BI.

Modelo Entidad Relación	Modelo Dimensional
<ul style="list-style-type: none"> • Enfocado a eliminar redundancia, coordinar actualizaciones y repetir el mismo tipo de operaciones muchas veces en el día. • Altamente normalizadas para soportar actualizaciones consistentes y mantenimiento de la integridad referencial • Tiempos de respuesta en segundos o inferior. • Almacenan pocos datos derivados. • Pocos datos agregados 	<ul style="list-style-type: none"> • Altamente desnormalizada ya que se requiere disminución de tiempos en la obtención de grandes cantidades de datos • Tiempos de respuesta aceptables pueden ser segundos, minutos, horas • Gran cantidad de datos derivados • Agregación: Varios niveles de datos precalculados.

El proyecto de consultas gubernamentales (BI) para VUCEM contempla un total de 18 unidades de negocio, de las cuales cada una tiene sus propias reglas de negocio. El esquema general del proyecto se muestra en la figura 1.3

El proceso general del desarrollo de las consultas gubernamentales fue de la siguiente manera:

- 1) Se tiene una fuente de información la cual es la base de datos transaccional de la Ventanilla Única.
- 2) Se hace el proceso de extracción, transformación y carga de datos (ETL) de acuerdo al tipo de trámite de cada unidad de negocio,
- 3) Se unifica la información de todos los trámites de todas las unidades de negocio en un almacén de datos operacional (ODS).
- 4) Se realiza el proceso de extracción, transformación y carga de datos de acuerdo a las reglas de cada negocio y se cargan a un almacén de datos específico de cada unidad de negocio (*Datamart*).

- 5) Toda la información específica de cada unidad de negocio, puede ser consultada de acuerdo a lo que el cliente solicite, en la herramienta de Cognos Query Studio la cual es accesible desde cualquier computadora con los usuarios necesarios.

Un diagrama simple que ejemplifica el proceso general del desarrollo del proyecto se puede observar en la Figura 1.2.

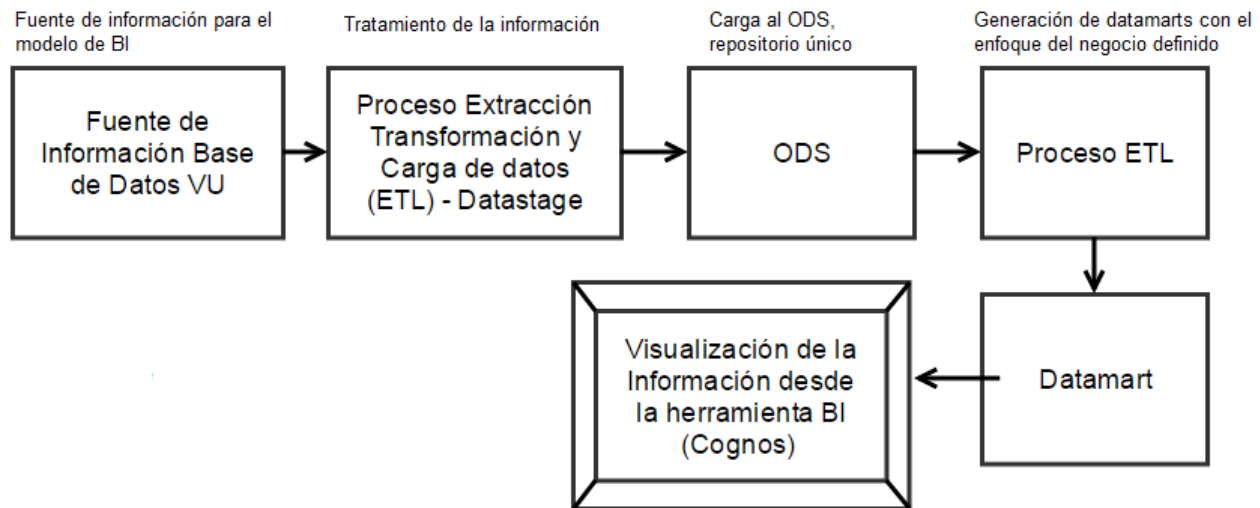


Figura 1.2 Proceso general de BI

Cognos Query Studio es la herramienta de creación de informes o reportes que sirve para crear consultas sencillas. Se pueden ver datos, detalles sobre el elemento de consulta, crear informes, cambiar el aspecto de los informes, trabajar con datos de un informe utilizando filtros, resúmenes y cálculos para comparar y analizar datos.

3.4.1 ETL

El proceso de Extracción, Transformación y Carga, ETL (del inglés *Extract, Transform and Load data*) es aquel en el que se extrae, transforma y carga datos desde uno o varios orígenes hasta uno o varios destinos.

Se puede extraer información de diversos tipos de datos, ya sean archivos excel, archivos planos, de una base de datos, de archivos xml, entre otros.

Las transformaciones que se pueden hacer son innumerables, ya que dependerá de cada requerimiento del negocio. Entre las más destacadas son generar columnas o registros específicos, hacer uniones, juntar información de 2 o más tablas, cambiar el orden de columna a registro, hacer búsquedas específicas, generar llaves primarias y subrogadas, eliminar algún tipo de dato, hacer filtros, entre otros.

La carga de la información transformada, al igual que las extracciones pueden ser en diversos archivos o tablas de bases de datos.



Figura 1.3 Proyecto BI para Vucem

Para desarrollar el ETL necesario, se utilizan herramientas muy específicas. Entre las más destacadas están:

Opensource: *Pentaho Data Integration, Scriptella Open Source ETL Tool, Talend Open Studio, Jitterbit.*

Comerciales: *IBM Websphere Datastage, Informatica Power Center, Oracle Warehouse Builder, Microsoft Integration Services, Ab Initio, Cognos.*

En el caso del proyecto de VUCEM se utilizó *IBM Websphere Datastage*, pero más adelante hablaré sobre ello.

Para poder identificar cada regla de negocio según cada unidad de negocio utilizamos una matriz origen-destino:

MATRIZ ORIGEN-DESTINO

ORIGEN						DESTINO							
ID	Nombre lógico de la tabla	Nombre físico de la tabla	Nombre lógico del campo	Nombre físico del campo	Tipo de dato	Nombre lógico de la tabla	Nombre físico de la tabla	Nombre lógico del campo	Nombre físico del campo	Tipo de dato	Id regla de negocio	Regla de transformación del campo	Regla de transformación de la tabla
1													
2													
				Staging Area	ODS	Modelo Analítico							

Figura 1.4 Matriz Origen Destino

En la matriz origen-destino se pueden identificar 3 procesos principales:

- 1) *Staging* (área de pruebas ó área temporal)
- 2) ODS (almacén operacional de datos, del inglés *Operational Data Store*)
- 3) Modelo Analítico

Cada proceso principal contiene las mismas columnas tanto para origen como destino, las cuales son:

- Nombre lógico de la tabla.- es el nombre con el cual se identifica rápidamente a la tabla.
- Nombre físico de la tabla.- es el nombre con el cual está definida la tabla en la base de datos.
- Nombre lógico del campo.- es el nombre con el cual se identifica rápidamente al campo necesario.
- Nombre físico del campo.- es el nombre con el cual está definido el campo en la base de datos.

- Tipo de Dato.- es el formato ó estructura del campo con el cual está definido.

Para hacer las transformaciones en el caso de ODS se toma como su origen las fuentes de datos y su destino es el ODS. Para el modelo analítico su origen es el ODS y su destino son las dimensiones correspondientes a cada unidad de negocio. En ambas situaciones se tienen las columnas de transformación que indican cómo es que se va a desarrollar el ETL necesario y las cuales son:

- Id de la regla de negocio
- Reglas de transformación del campo
- Regla de transformación de la tabla

Éstas 3 columnas son en las que los desarrolladores nos basamos para poder hacer el ETL y con ello nos cercioramos que al hacer la carga sea la correcta y con los datos requeridos de forma homogénea.

3.4.2 Área de *Staging*

El área de *staging* es un área intermedia de almacenamiento de datos utilizada para el procesamiento de los mismos durante procesos de extracción, transformación y carga de datos.

Las áreas de *staging* de datos son transitorias, es decir, contienen la información de las bases de datos transaccionales, son utilizadas para que no se afecte de manera directa al negocio u ocupe recursos que haga más lenta las transacciones de la base de datos origen.

Su contenido se borrará antes de ejecutar un proceso de ETL o inmediatamente después de haberlo finalizado con éxito. Aunque existen arquitecturas de área *staging* diseñadas para mantener los datos durante largos períodos de tiempo con la finalidad de mantener un archivo de los mismos o para poder resolver problemas detectados a posteriori.

La diferencia con respecto al ODS es que en el área de *staging* solo se pasa la información sin sufrir ningún tipo de cambio. En el ODS a pesar de que también se unifica y se pasa la información, si llega a sufrir alguna modificación.

En el caso del proyecto de VUCEM no se utilizó un área similar ya que la solución al problema no lo requería.

3.4.3 ODS

El ODS (del inglés *Operational Data Store*) o almacén operacional de datos, es un contenedor de datos que tiene información relevante, es decir, está diseñado para integrar datos de múltiples fuentes con los que se realizan operaciones adicionales en los propios datos. Es donde se preparan los datos para la realización de otras operaciones o para ser incorporados a un almacén de datos (del inglés *Data Warehouse*) para ser explotados en la realización de informes, reportes o en tablas de consulta.

Dado que los datos se originan a partir de múltiples fuentes, la integración a menudo implica limpiar dichos datos, es decir, hacer que los datos sean homogéneos y uniformes, para resolver redundancias y comprobar la integridad de los mismos según las reglas del negocio.

Para el proyecto de VUCEM fue el área donde las 18 unidades de negocio, cada una con sus respectivos trámites, se unificaban. Cada número de folio de trámite está asociado a un identificador de la solicitud, por medio de ello podríamos saber a qué unidad de negocio correspondía cada trámite.

3.4.4 Datamart

El modelo analítico o *datamarts* son subconjuntos de datos con el propósito de ayudar a que un área específica dentro del negocio pueda tomar mejores decisiones. El datamart es un sistema orientado a la consulta, se puede decir que los datamarts son pequeños almacenes de información centrados en un tema o un área de negocio específico dentro de una organización.

Un *datamart* se compone por dimensiones y hechos, en el proyecto de consultas gubernamentales para VUCEM.

En el caso del proyecto, sólo se tomaron para desarrollo, los puntos 3.4.3 y 3.4.4, ya que para este caso no fue necesario el área de *staging*.

El proceso general de toda unidad de negocio es:

- 1) Identificar los trámites asociados a cada unidad de negocio.
- 2) Realizar la consulta y extracción de los trámites identificados.
- 3) Hacer el proceso de los pasos generales para la carga de los trámites correctos.
 - 3.1) Obtener los trámites con su dictamen, su resolución, su requerimiento, su notificación, sus datos generales y las personas asociadas a la solicitud y se filtra la información obteniendo el máximo (último) dictamen con la última persona asociada y se carga a una tabla temporal 1.
 - 3.2) Se generan las columnas con un tipo de identificador para la resolución (tipo 'r'), para el dictamen (tipo 'd'), para el requerimiento (tipo q), para la solicitud (tipo s) para la notificación de la resolución (tipo n) para la notificación del requerimiento (tipo m), eliminado posibles nulos del dictamen, el resultado se carga a la base .
 - 3.3) Se asocia la tabla temporal 1 con las firmas a partir de la resolución de la notificación y se unen todas las firmas.
 - 3.4) Se cargan las firmas a una tabla temporal 2.
 - 3.5) Se agregan las observaciones y claves de usuario únicas a las firmas y se carga a la temporal 3.
 - 3.6) Se obtienen todos los trámites con todas las características necesarias y dependiendo de cada trámite de cada unidad de negocio.
 - 3.7) Por último se agregan los cargos asociados a los trámites de cada unidad de negocio.

- 4) Realizar las transformaciones necesarias dependiendo de cada unidad de negocio para complementar las reglas de negocio agregando la información suficiente.
- 5) En caso del *datamart*, se generan primero las llaves subrogadas.
- 6) Se cargan las transformaciones a las tablas correspondientes. En el caso del *datamart*, primero se cargan las dimensiones y al final las tablas de hechos.

El proceso general de toda unidad de negocio se resume en la Figura 1.41.

En los *datamarts* solo se hacen extracciones de catálogos necesarios, para el flujo principal se toma como insumo los *datasets* que se cargan en el ODS.

En mi caso me tocó desarrollar 3 de los procesos: la Unidad 1, Unidad 2 y Unidad 3, las cuales explico más adelante a partir del punto 3.6.

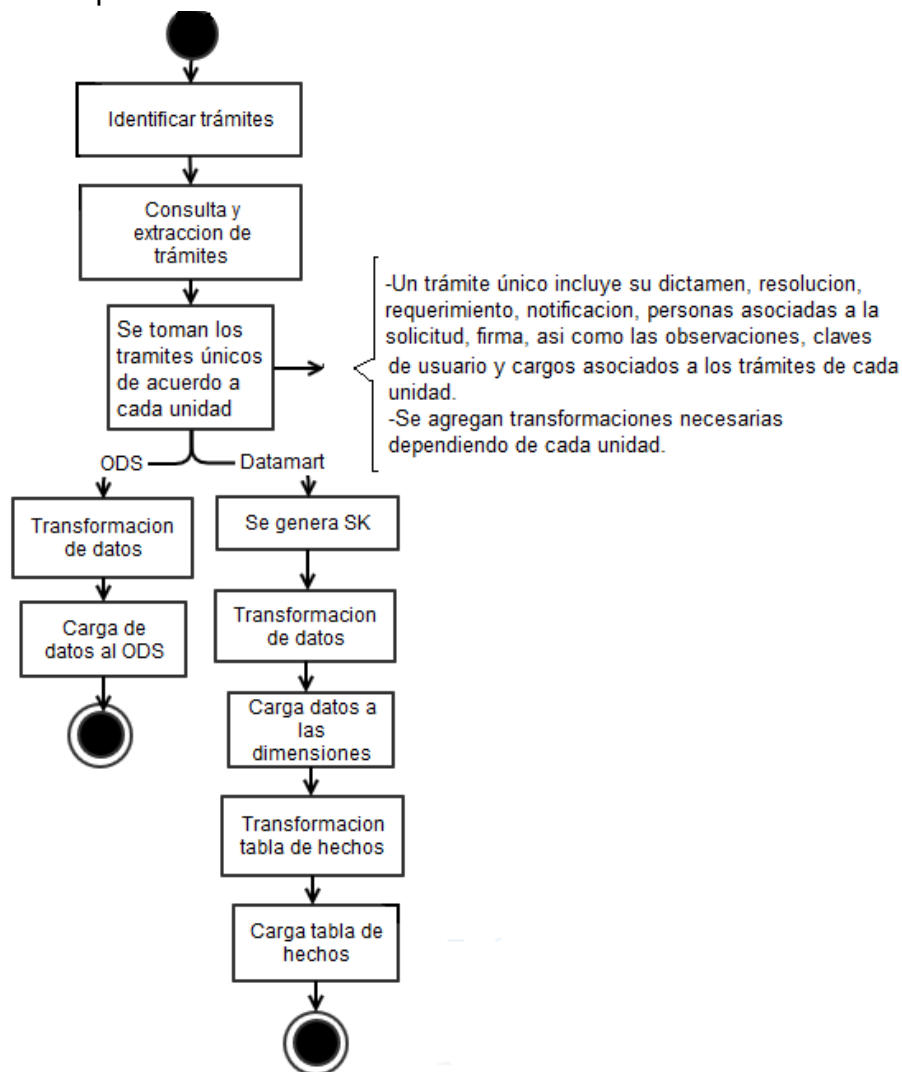


Figura 1.41 Proceso general de ETL de las unidades de negocio

3.5 IBM Websphere Datastage

IBM Websphere Datastage es una herramienta de desarrollo de ETL la cual utilizamos en el proyecto para hacer la construcción tanto de la parte de ODS como de *Datamarts*.

La unidad mínima de ejecución de un proceso automatizado es el *job*.

El *stage* es la representación gráfica de las operaciones que se pueden hacer en un *job*.

Datastage se compone de tres interfaces para la construcción de los procesos ETL, ejecución, monitoreo y administración de los mismos los cuales son: Administrador, Director y Diseñador.

El administrador maneja detalles de licenciamiento, crea actualiza proyectos y usuarios, administra la configuración de las variables de ambiente para todo el proyecto.

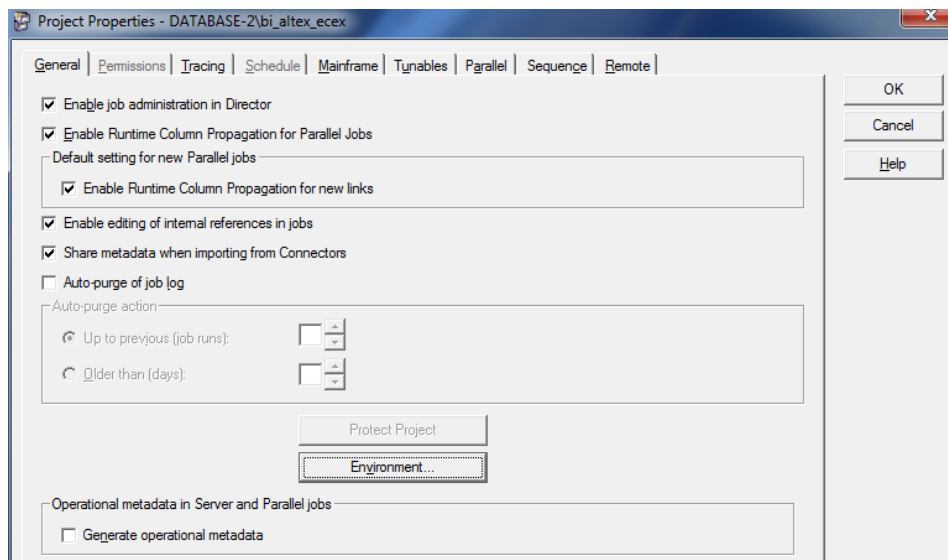


Figura 1.5 Pantalla Inicial de administrador de proyecto

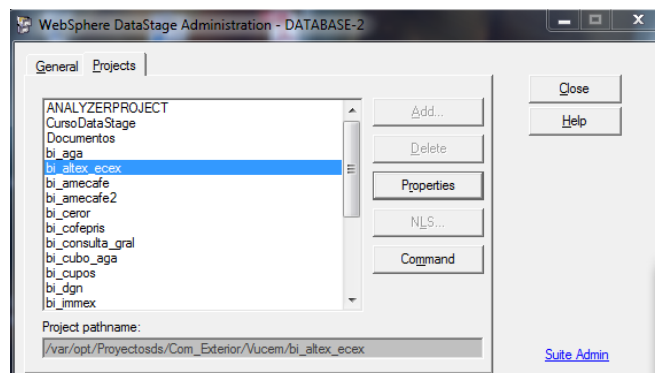


Figura 1.6 Administración de proyecto

El diseñador es en el que se crea las aplicaciones conocidas como *jobs*, además permite crear, modificar, compilar, y ejecutar *jobs*, también monitorear el desempeño del *job*.

El director valida ejecuta y monitorea cada uno de los *jobs*.

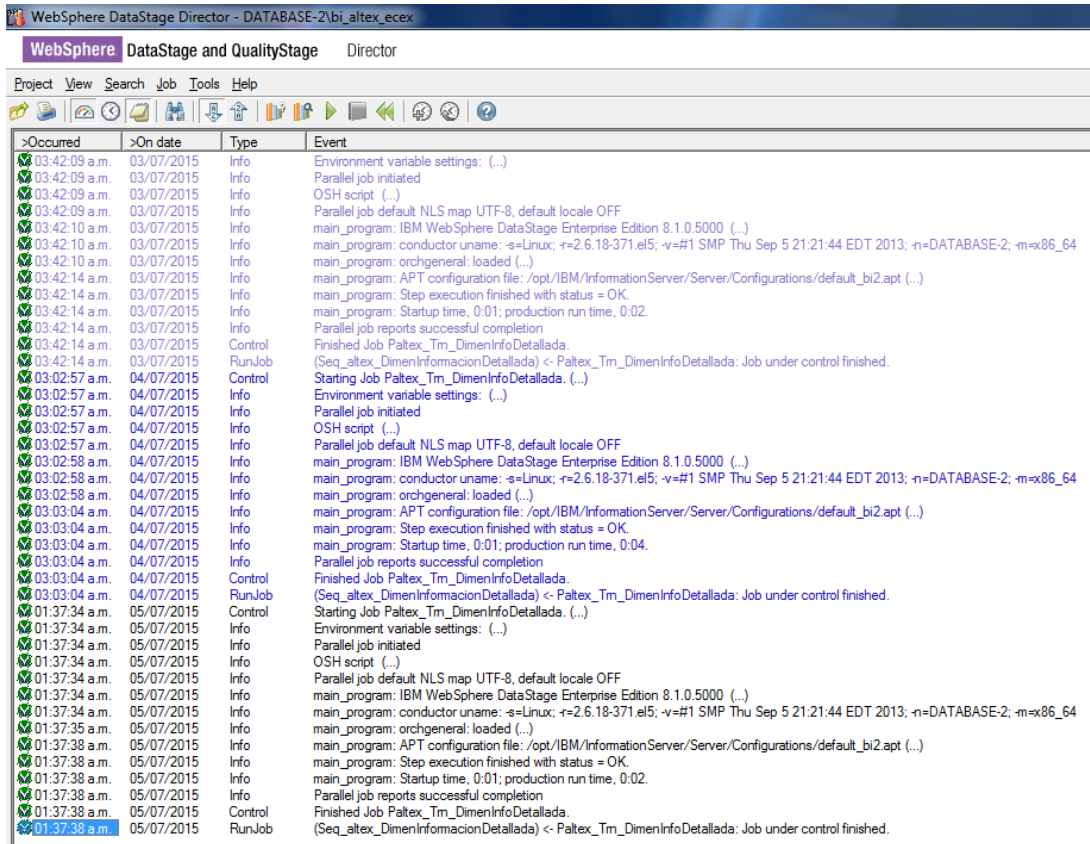


Figura 1.7 Ejemplo de ejecución de un job

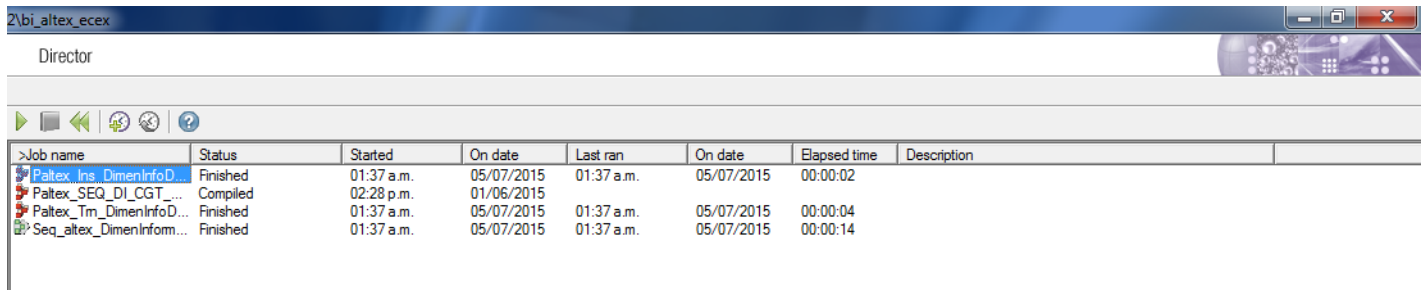


Figura 1.8 Ejemplo de director de un job

Existen diversos tipos de *jobs* en *Datastage* las que utilizamos fueron *Job Paralelo* (Paralell Job) y *Job Secuencial* (Sequence Job).

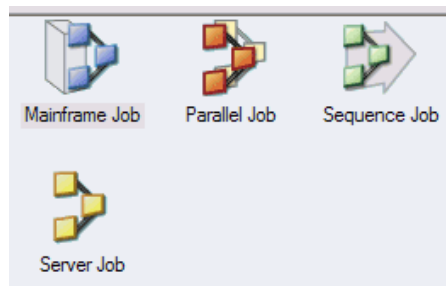


Figura 1.9 Tipos de Job

El *job* paralelo es el que se utilizó en todo tipo de extracciones, transformaciones y carga de datos. La ventaja es que su funcionamiento como su nombre lo dice, es en paralelo por lo que la velocidad de procesamiento de datos aumenta.

El *job* de secuencia se utilizó para generar *jobs* generales de extracción, transformación, carga de datos y una secuencia maestra. Nos permiten agrupar en un solo *job* los distintos *jobs* paralelos que se tenían para una misma unidad de negocio, lo que nos permitía ejecutar solo ese *job* de extracción o transformación o carga. La secuencia puede ser tanto en cadena (Figura 1.10), como en paralelo.



Figura 1.10 Ejemplo de secuencia en cadena

Un *job* de secuencia maestra es aquel que invocaba un *job* de secuencias de extracción, el *job* de secuencias de transformación y el *job* de secuencias de carga. Esto permite ejecutar solo ese *job* y el resultado es que se ejecutaban todos los procesos correspondientes a cada unidad de negocio, además de que es posible calendarizar para que su ejecución sea automática.

3.6 Unidad de negocio 1

3.6.1 ODS Unidad de negocio 1

3.6.1.1 Extracciones Unidad de negocio 1

En las figura 2.0, se tiene un solo *job* de extracción, el cual es reutilizable. Contiene 2 stages uno que es el que consulta a la base de datos de oracle y el otro stage es un dataset de salida en donde se almacena la información extraída.

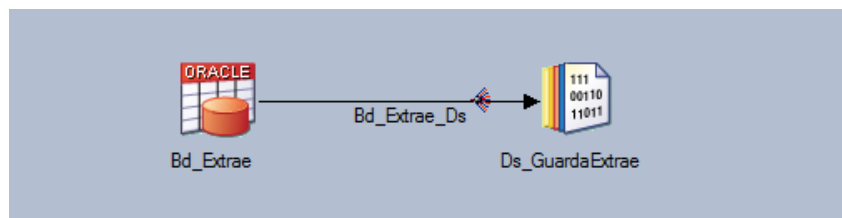


Figura 2.0 Job de extracción de la Unidad 1

3.6.1.2 Transformaciones Unidad de negocio 1

Dependiendo de la unidad de negocio son las transformaciones que se hacen, en algunas se generan columnas, en otras se toman valores máximos, en otras solamente se filtra, en algunas se hacen asociaciones de tablas, ó se ordenan los datos entre otras transformaciones existentes.

En la figura 2.1 se tiene como insumo un *dataset* de la dirección de la solicitud, la cual se filtra por el tipo de persona y tomando solo los identificadores únicos de la dirección de la solicitud, por último se agregan columnas para identificar que usuario hizo el desarrollo y en qué fecha se cargó.

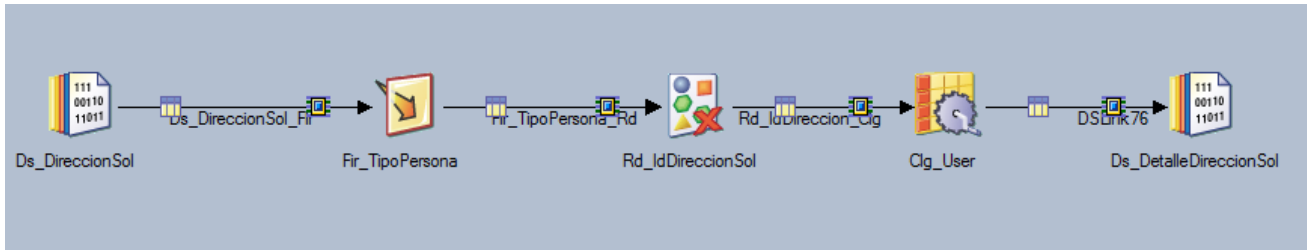


Figura 2.1 Job de transformación de la dirección de la solicitud de la Unidad 1

Para la figura 2.2 se asocian los datos de la unidad 1 con las personas asociadas a la solicitud, en seguida se toman solo cierto tipo de personas y se agregan columnas para identificar que usuario hizo el desarrollo y en qué fecha se cargó

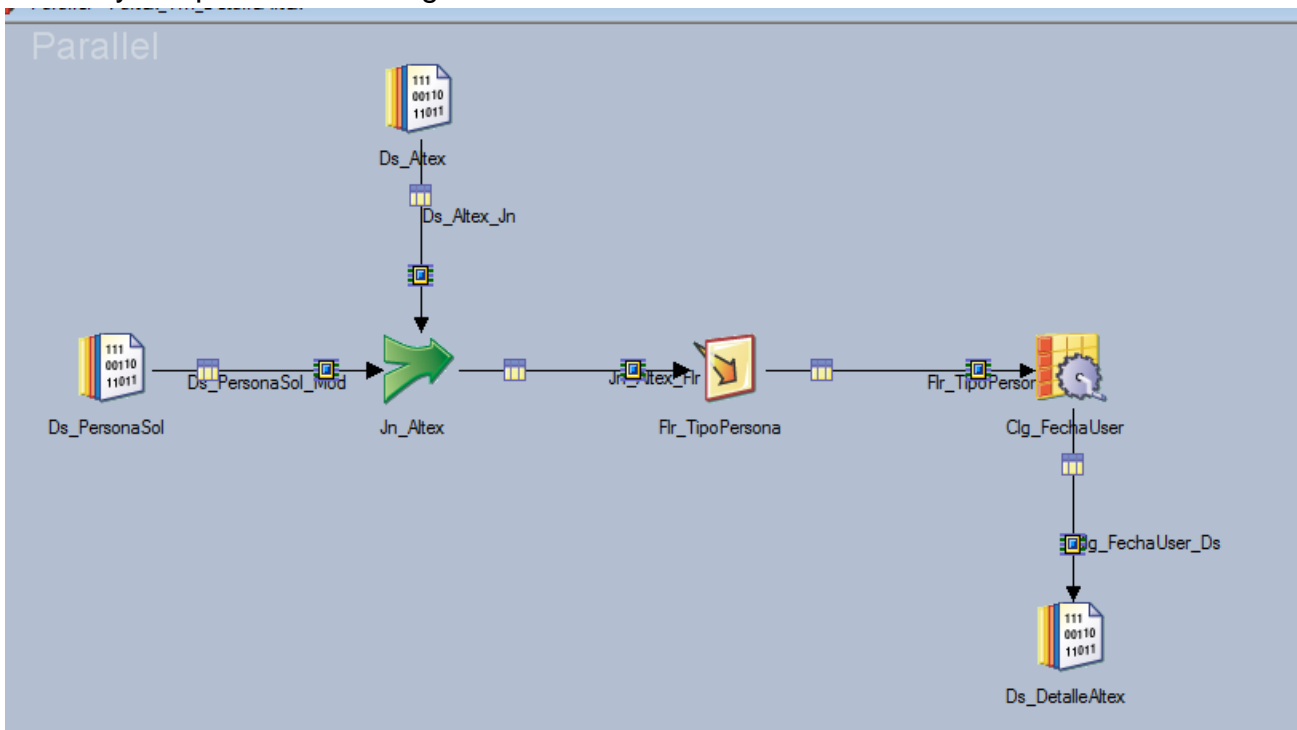


Figura 2.2 Job de transformación del trámite de la Unidad 1

El *job* de la figura 2.3 está asociado a las firmas de los solicitantes las cuales se filtran por clave de rol y tomando solo identificadores de firmas únicas, su resultado se guarda en un archivo tipo *dataset*.

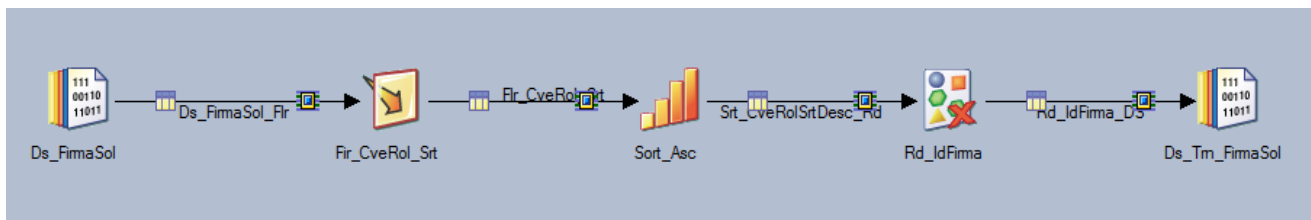


Figura 2.3 Job de transformación de la firma de la solicitud de la Unidad 1

En la figura 2.4 se tienen los trámites de la unidad 1, a los cuales se les agrega la resolución y la firma de la solicitud a partir del número de folio de trámite y del identificador de la solicitud.

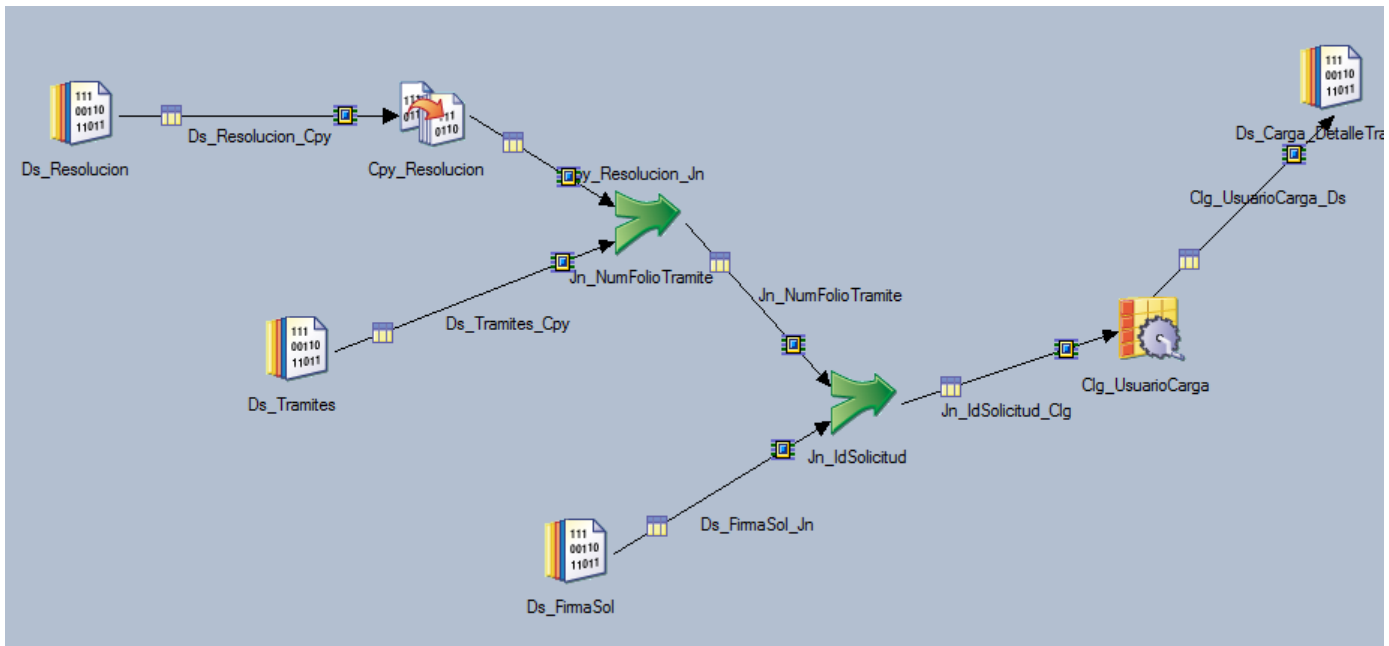


Figura 2.4 Job de transformación del detalle del trámite de la Unidad 1

En el job de la figura 2.5 es el proceso principal, ya que es el proceso que conforma los trámites de la unidad de negocio 1.

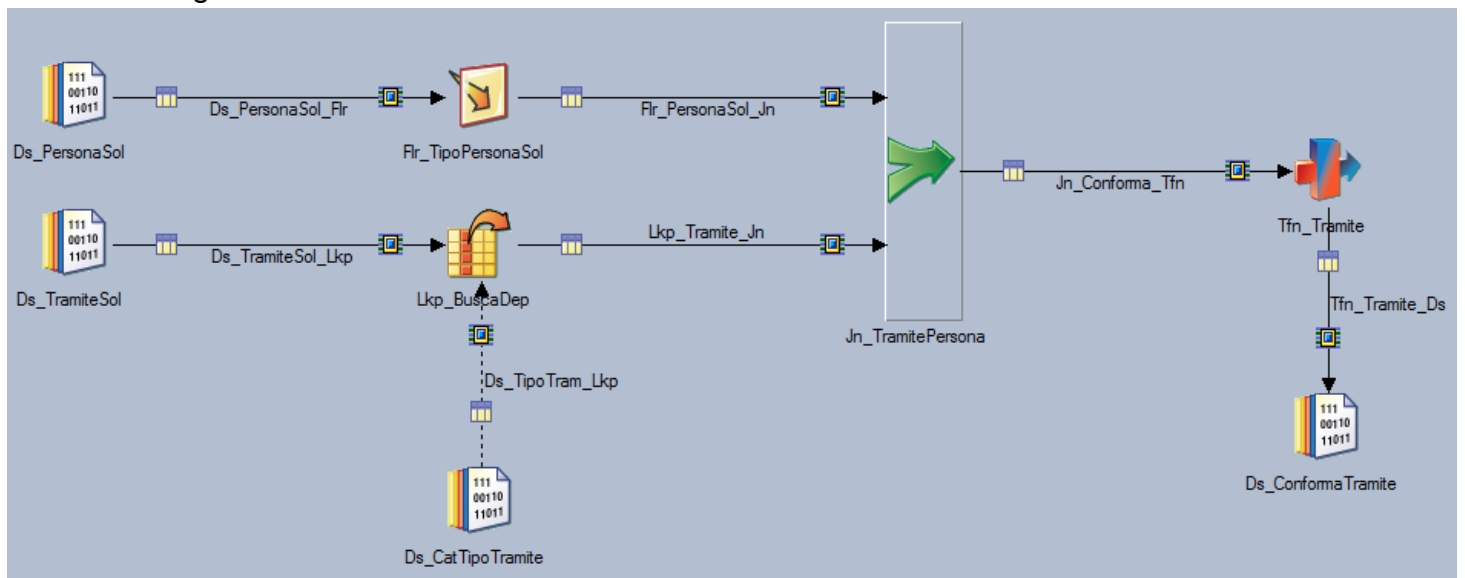


Figura 2.5 Job de transformación del trámite de la Unidad 1

3.6.1.3 Cargas Unidad de negocio 1

Una vez teniendo el *dataset* con todas las transformaciones se realiza la carga a la tabla correspondiente, la cual tiene un archivo *dataset* de rechazo que indica los registros que por cualquier causa no pudieron ser insertados en la base de datos. (Figura 2.6, Figura 2.7, Figura 2.8, Figura 2.9)

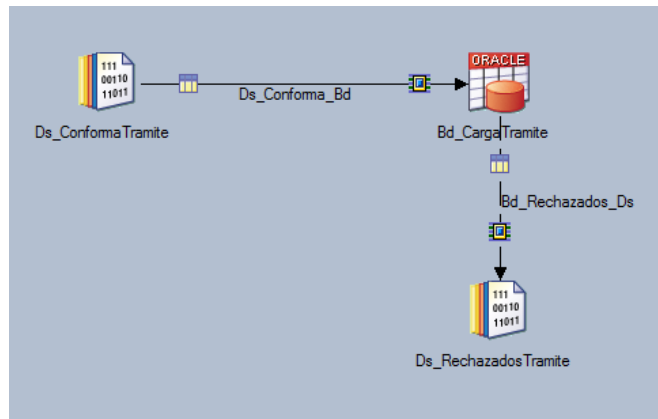


Figura 2.6 Job de carga del trámite de la Unidad1

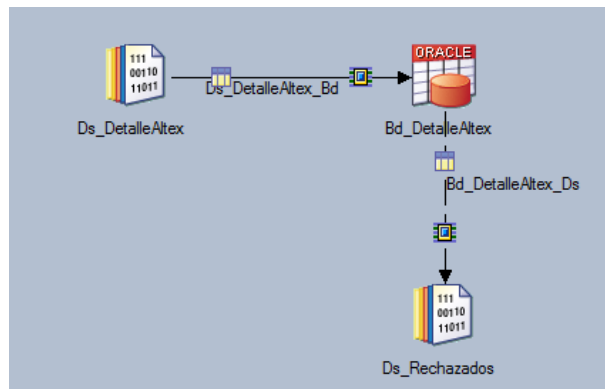


Figura 2.7 Job de carga del detalle de la Unidad 1

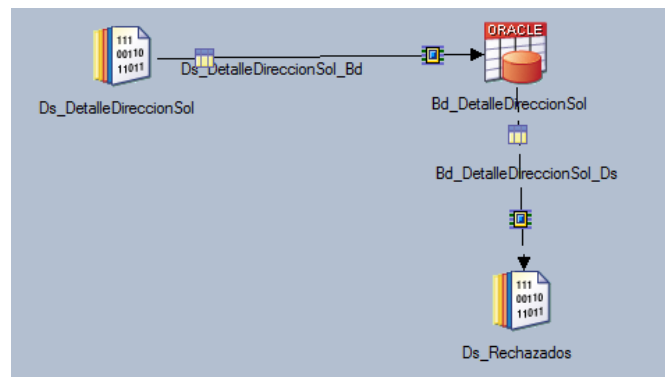


Figura 2.8 Job de carga del detalle de la dirección de la solicitud de la Unidad 1

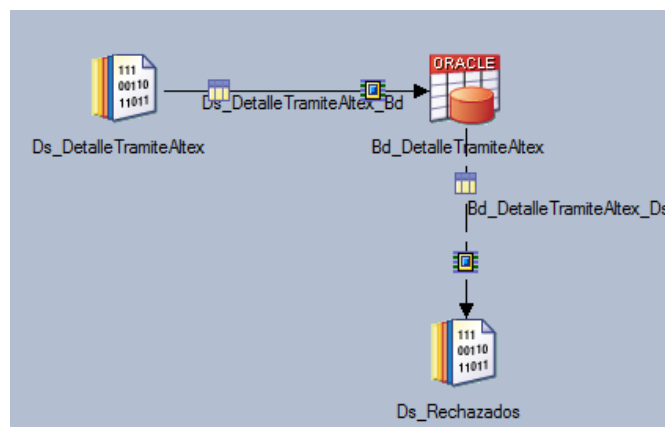


Figura 2.9 Job de carga del detalle del trámite de la unidad 1

En el caso de la figura 2.10, es la carga de los trámites, y se tiene un *stage* intermedio para renombrar las columnas con el mismo nombre que su destino, también es un *job* de carga.

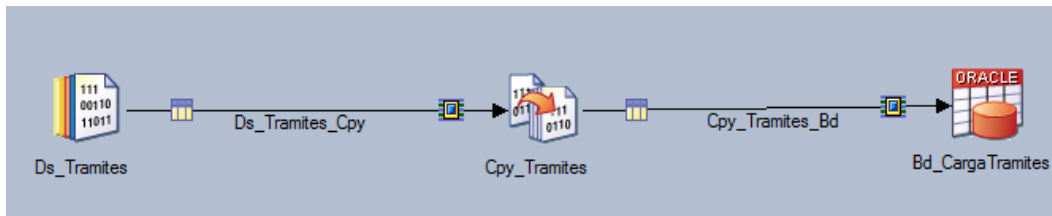


Figura 2.10 Job de carga de los trámites de la Unidad 1

3.6.1.4 Secuencias ODS Unidad de negocio 1

Las secuencias indican el orden en que los *jobs* se ejecutarán teniendo así, un orden específico para cada unidad de negocio. La secuencia temporal es la que manda a llamar la ejecución de los trámites correspondientes a cada unidad de negocio, la secuencia de extracción toma solo los *jobs* de extracción de la unidad de negocio, la secuencia de transformación solo los *jobs* de transformación, la secuencia de carga solo los *jobs* de carga y la secuencia maestra manda a llamar a las secuencias temporal, de extracción, transformación y carga de cada unidad de negocio.

En la figura 2.11 se tiene un ejemplo de una secuencia temporal, esto se nota al tener tanto un *job* de extracción como un *job* de carga dentro de la misma secuencia, además de que toda secuencia temporal, es aquella que conforma a los trámites de cada unidad.

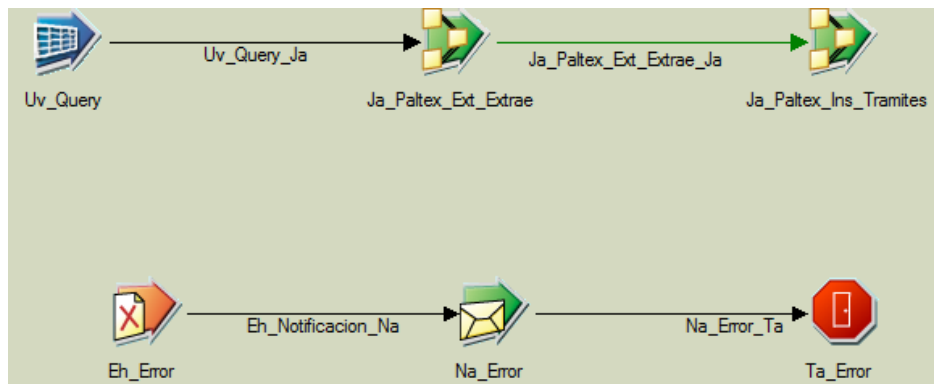


Figura 2.11 Secuencia que conforma los trámites de la Unidad 1

La figura 2.12 es una secuencia de extracción, se compone de los stage de creación de variables, el job de extracción reutilizable, el loop de inicio y el loop de fin de proceso.

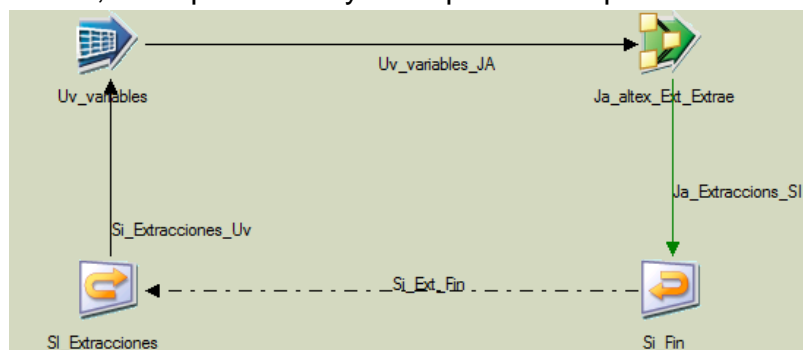


Figura 2.12 Secuencia de extracción de la Unidad 1

En la figura 2.13 se tiene una secuencia de transformación, es un claro ejemplo de ejecución en paralelo junto con una ejecución en cadena, cada job corresponde a las transformaciones realizadas para el ODS de la unidad 1.

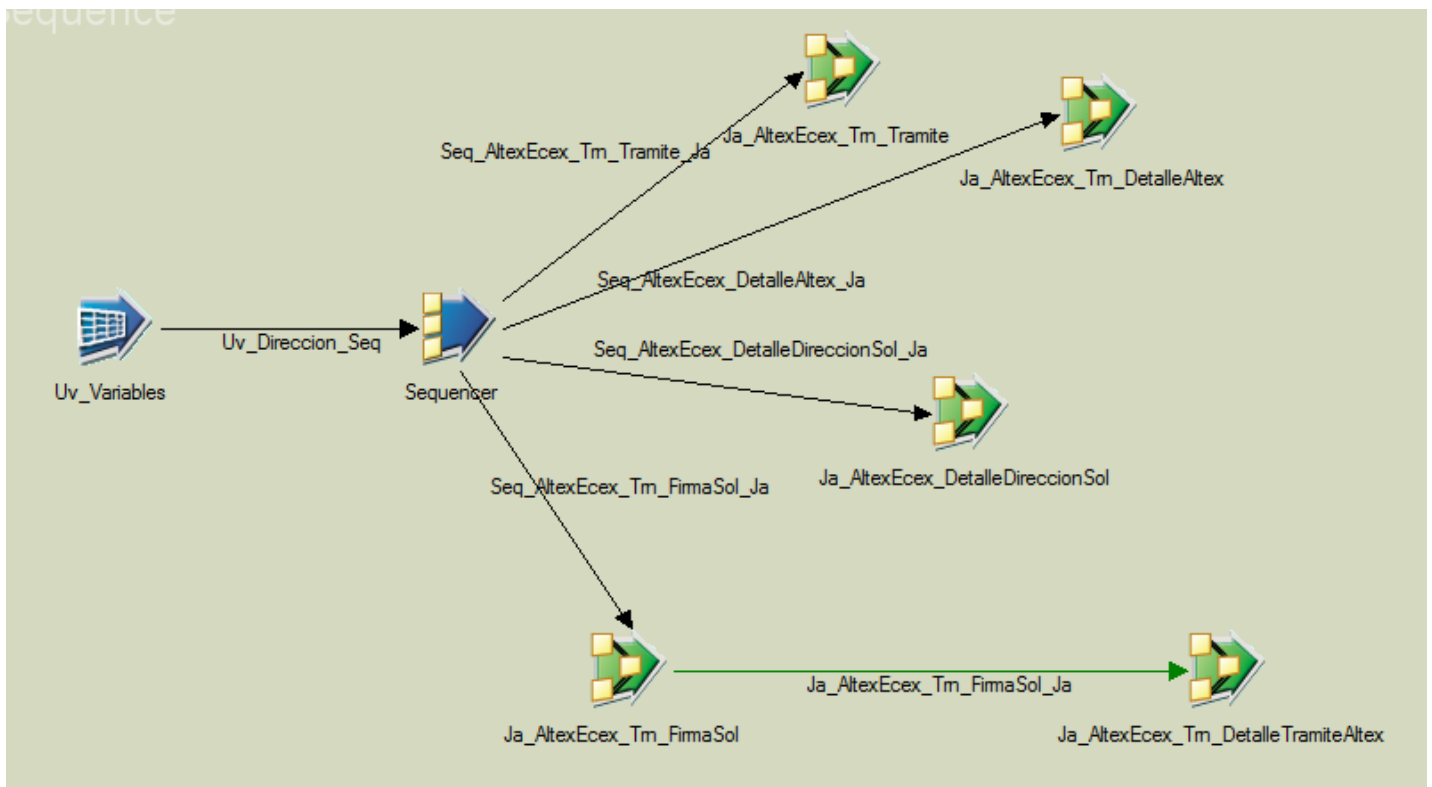


Figura 2.13 Secuencia de transformación de la Unidad 1

En la figura 2.14 se tiene la secuencia de carga, en la cual solo se manda a llamar cada job de carga creado, se pudo tambien haber creado en paralelo.

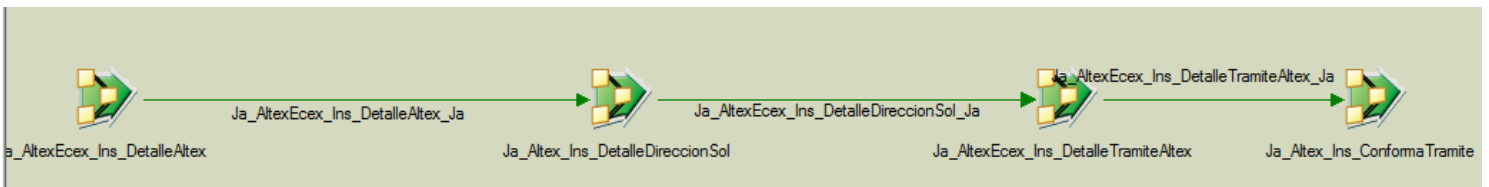


Figura 2.14 Secuencia de carga de la Unidad 1

La secuencia maestra mostrada en la figura 2.15 manda a llamar cada una de las secuencias antes mencionadas, con lo que, su ejecución implica crear cada uno de los jobs desarrollados en el ODS para la unidad 1.

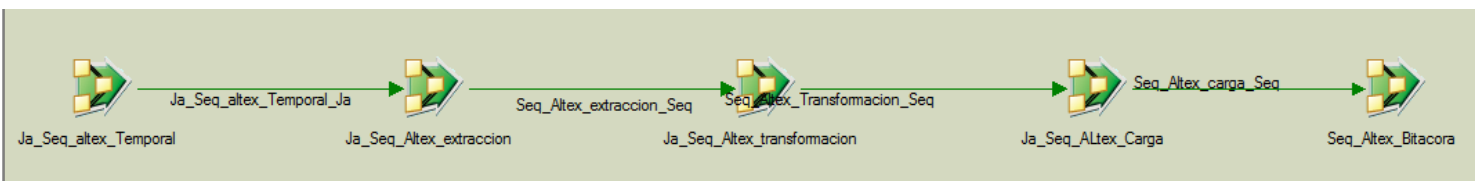


Figura 2.15 Secuencia maestra de la Unidad 1

3.6.2 DataMart Unidad de negocio 1

3.6.2.1 Transformaciones Unidad de negocio 1

Para cada una de las dimensiones, primero es necesario crear su llave subrogada, la cual es la llave primaria de la dimensión y que sirve para identificar como único a cada registro de la dimensión. (Figura 2.16, Figura 2.17, Figura 2.18, Figura 2.19)

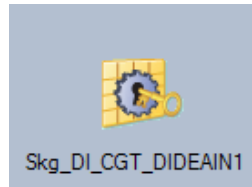


Figura 2.16 Job de creación de llave subrogada de la dimensión detallada de la información

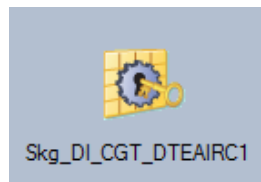


Figura 2.17 Job de creación de llave subrogada de la dimensión de los trámites de la unidad 1

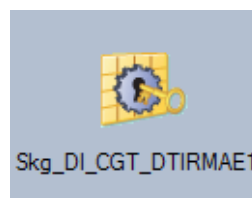


Figura 2.18 Job de creación de llave subrogada de la dimensión del trámite



Figura 2.19 Job de creación de llave subrogada de la tabla de hechos del trámite

Una vez creadas las llaves subrogadas, se procede a crear la transformación correspondiente de cada dimensión.

En la figura 2.20 se tiene la transformación de la dimensión de la información detallada, la cual toma los trámites de la unidad 1 agregando los detalles de la dirección de la solicitud y diferentes campos de cada catálogo necesario. Enseguida se genera la llave subrogada. La transformación se guarda en un archivo tipo *dataset*.

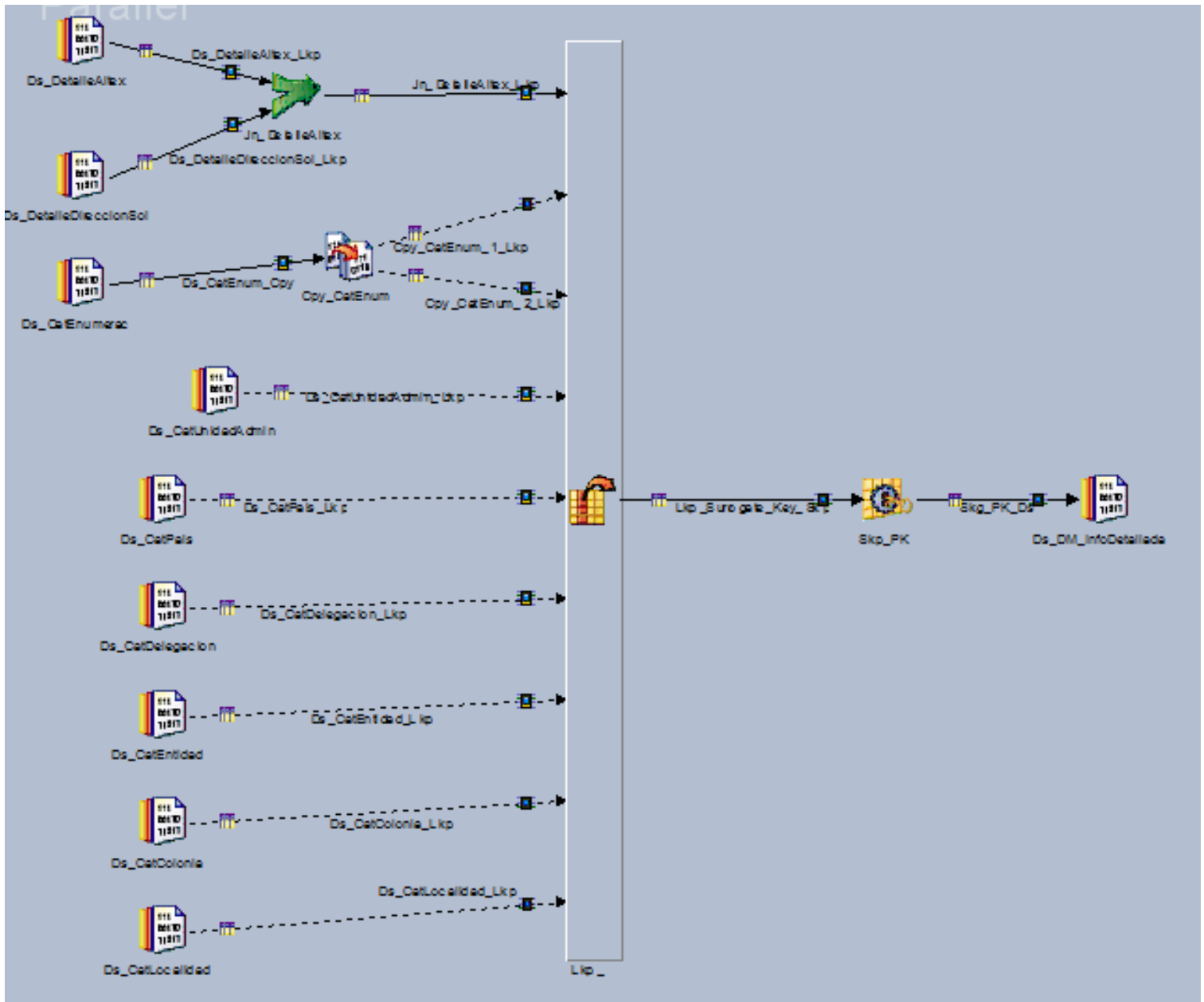


Figura 2.20 Job de transformación de la dimensión de la información detallada

La figura 2.21 corresponde a la transformación de los trámites de la unidad 1 en donde se agregan unidades administrativas, enumeraciones para la dimensión, al final se agrega la llave subrogada que se creó antes. El resultado se guarda en un archivo tipo *dataset*.

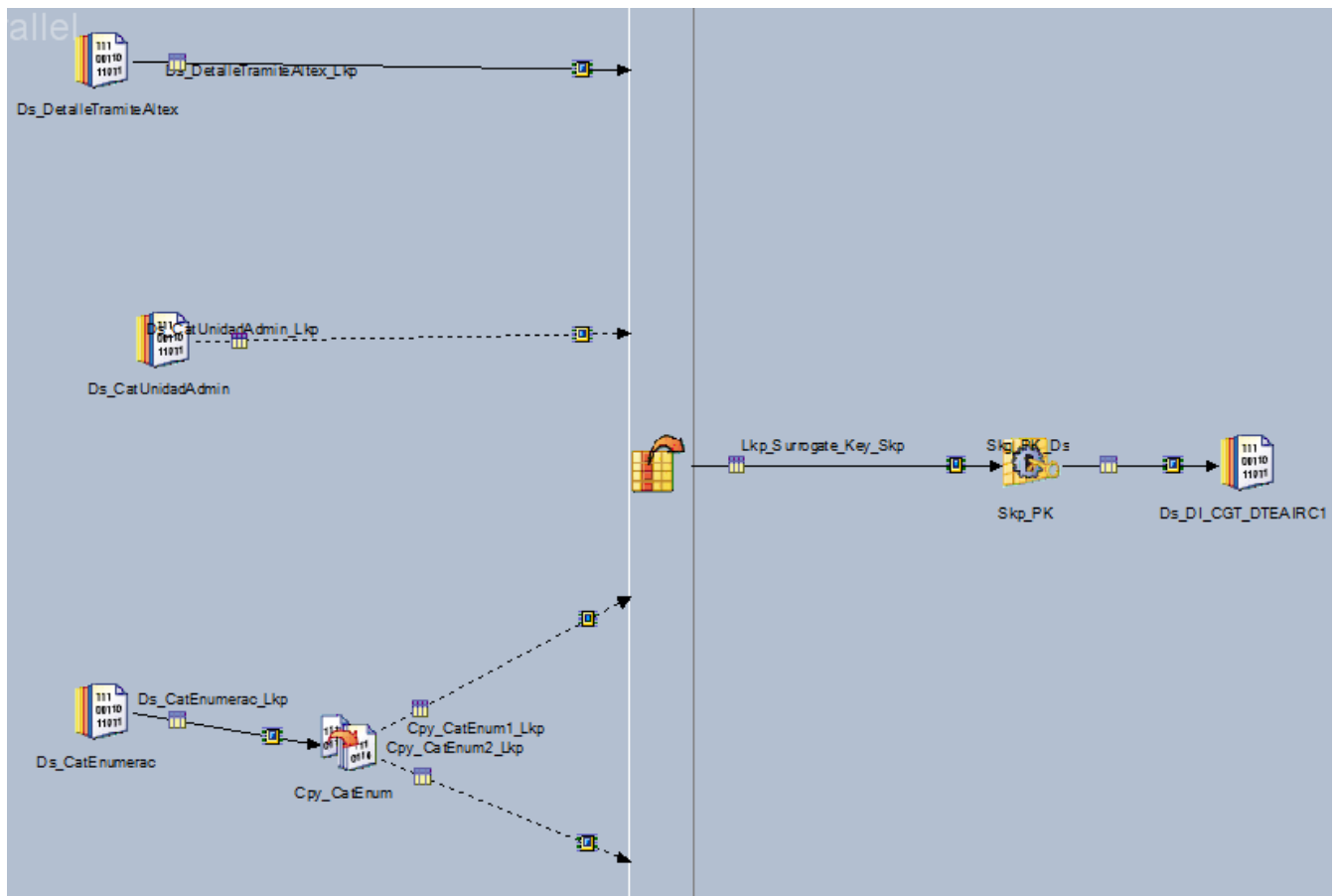


Figura 2.21 Job de transformación de los trámites de la unidad 1

Para la dimensión del trámite, se agregaron muchos datos de diferentes catálogos, como localidad, colonia, entidad, municipio, país, tipo de trámite para complementar la información del trámite.

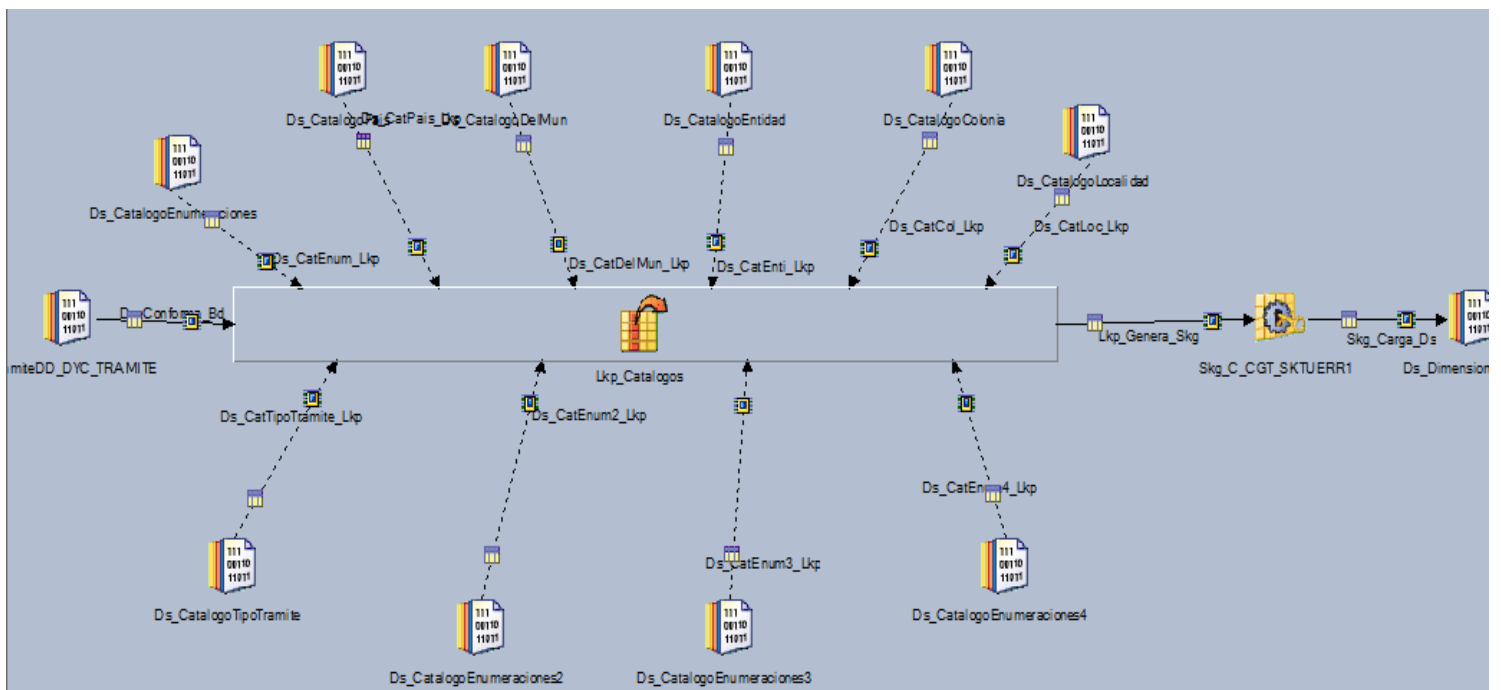


Figura 2.22 Job de transformación de la dimensión del trámite

Una vez teniendo todas las dimensiones creadas, se procede a crear la tabla de hechos del trámite de la unidad1, asociando cada dimensión con los trámites, generando también columnas de carga de usuario y fecha con su respectiva llave subrogada.

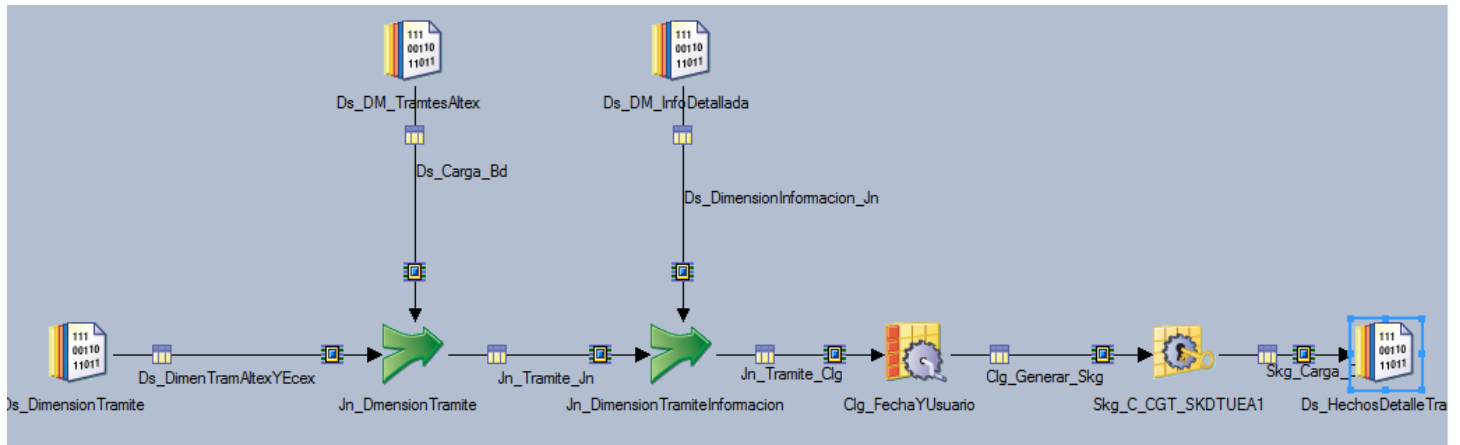


Figura 2.23 Job de transformación de los hechos del trámite

3.6.2.2 Cargas Unidad de negocio 1

Cada dimensión de la unidad de negocio de negocio se carga a la tabla correspondiente. Se tiene un archivo tipo *dataset* como archivo de rechazo, el cual se utiliza en caso de que por alguna razón no pueda ser insertado algún registro. (Figura 2.24, Figuras 2.25, Figuras 2.26, Figuras 2.27).

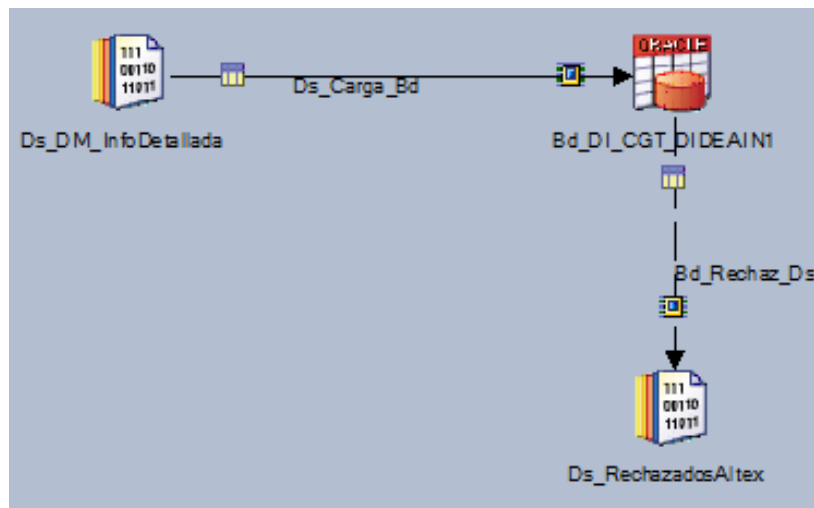


Figura 2.24 Job de carga de la dimensión de la información detallada

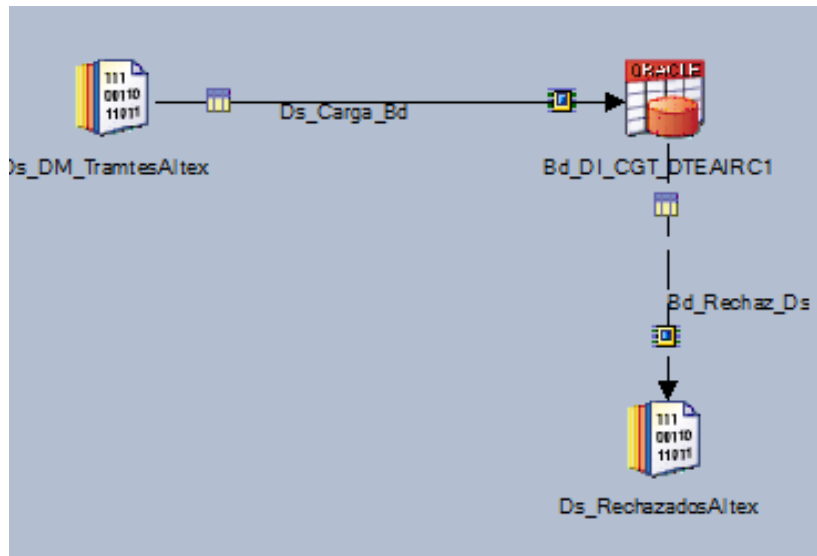


Figura 2.25 Job de carga de la dimensión de los trámites de la unidad 1

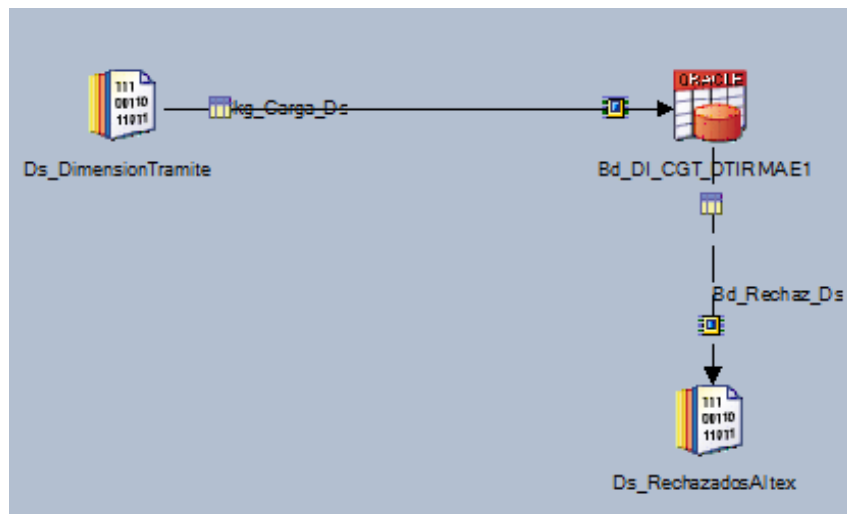


Figura 2.26 Job de carga de la dimensión del trámite

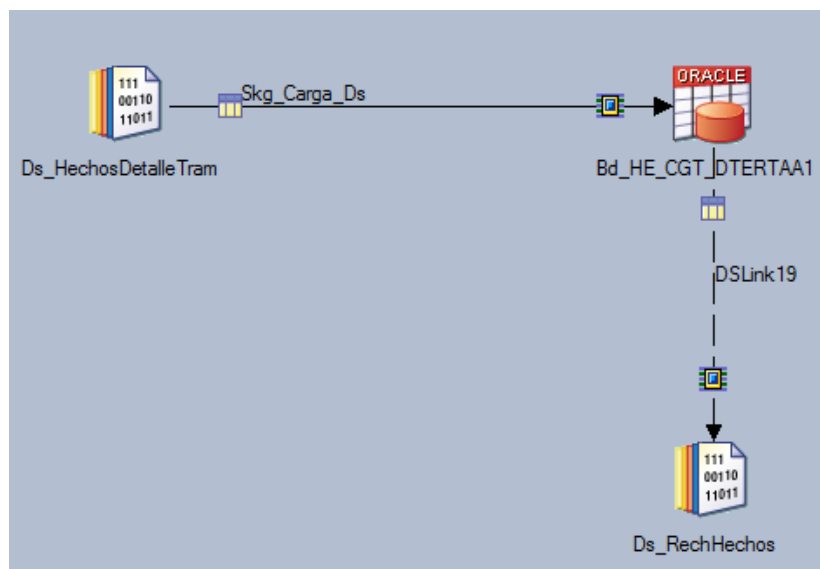


Figura 2.27 Job de carga de los hechos del trámite

3.6.2.3 Secuencias *Datamart* Unidad de negocio 1

En el caso de las secuencias del *datamart*, primero se hace una secuencia que engloba a la extracción de catálogos (Figura 2.32), después una secuencia que engloba a un job de transformación y un job de carga de cada dimensión (Figura 2.28, Figura 2.29, Figura 2.30) y cada tabla de hechos creada (Figura 2.31). Así mismo también se tiene una secuencia maestra, la cual manda a llamar a cada una de las secuencias de las dimensiones. Figura 2.33

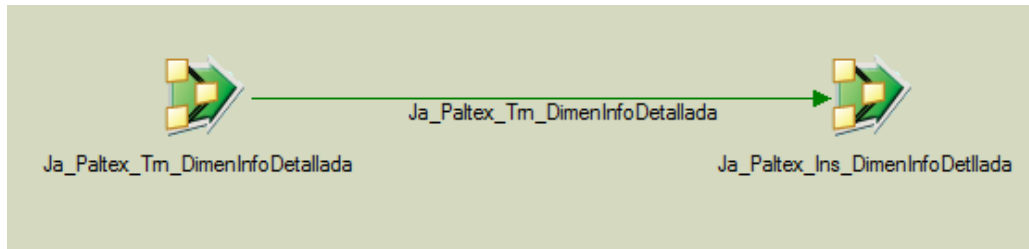


Figura 2.28 Secuencia de la dimensión de la información detallada

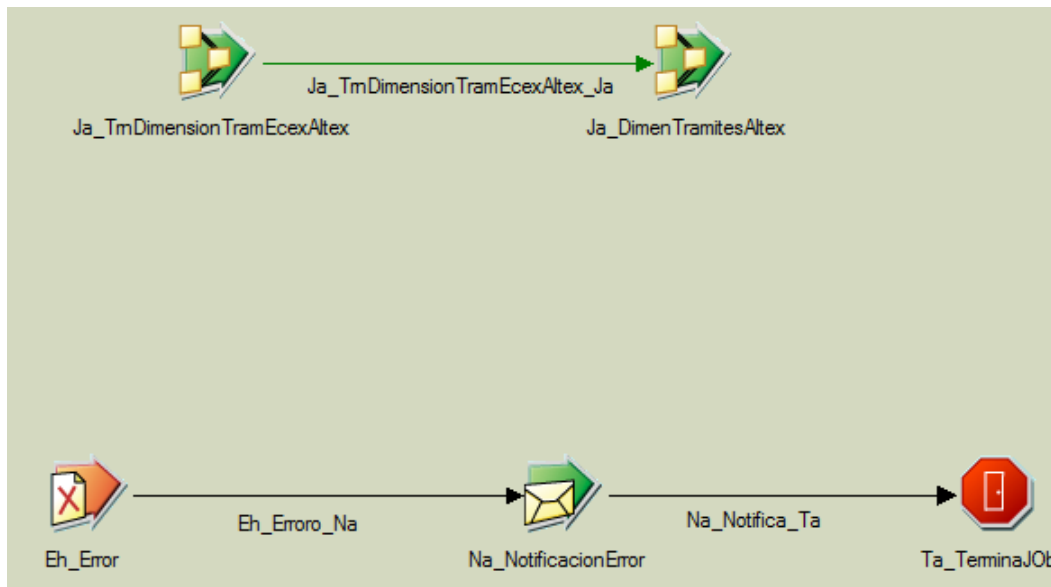


Figura 2.29 Secuencia de la dimensión de los trámites de la unidad 1

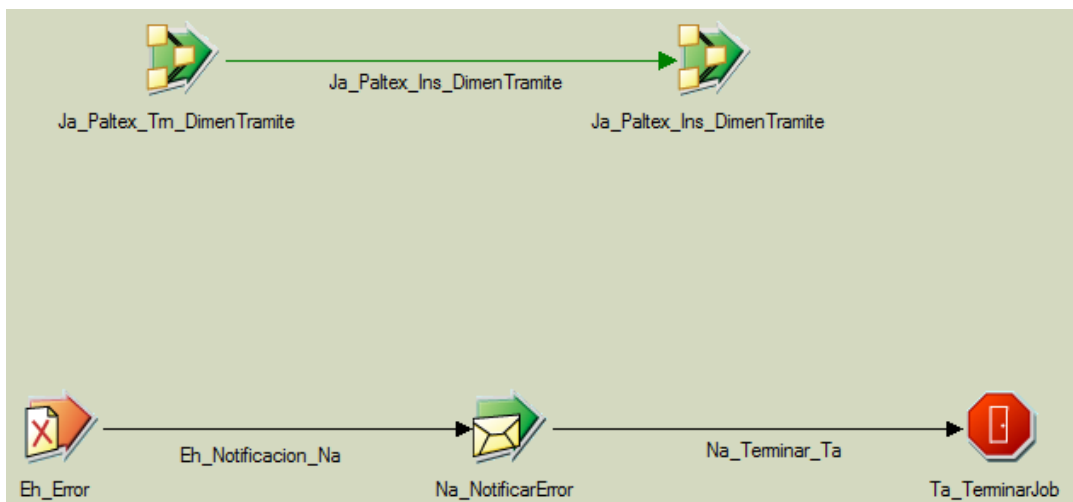


Figura 2.30 Secuencia de la dimensión del trámite

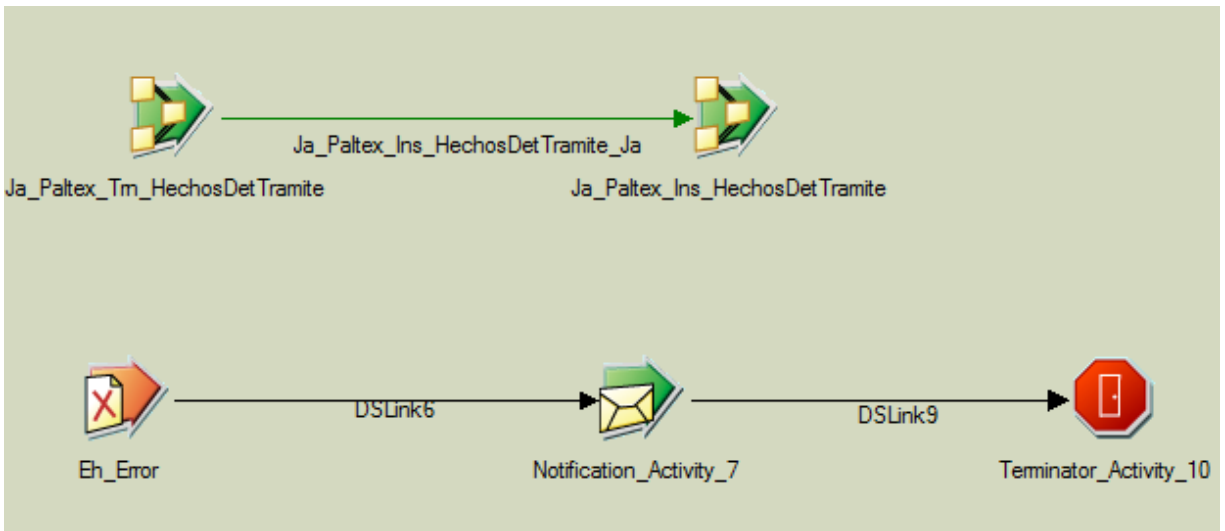


Figura 2.31 Secuencia de los hechos del trámite

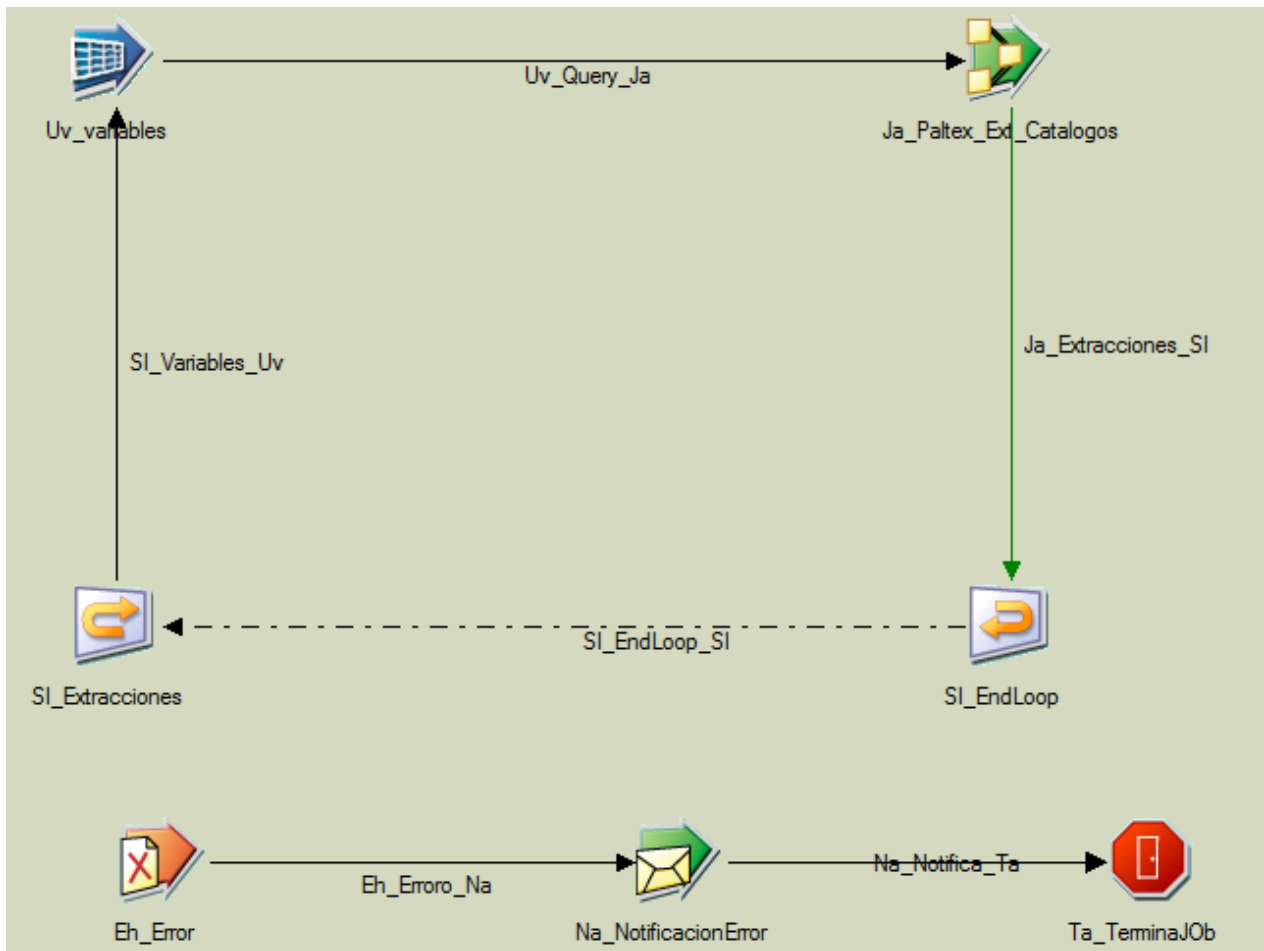


Figura 2.32 Secuencia de extracción de catálogos

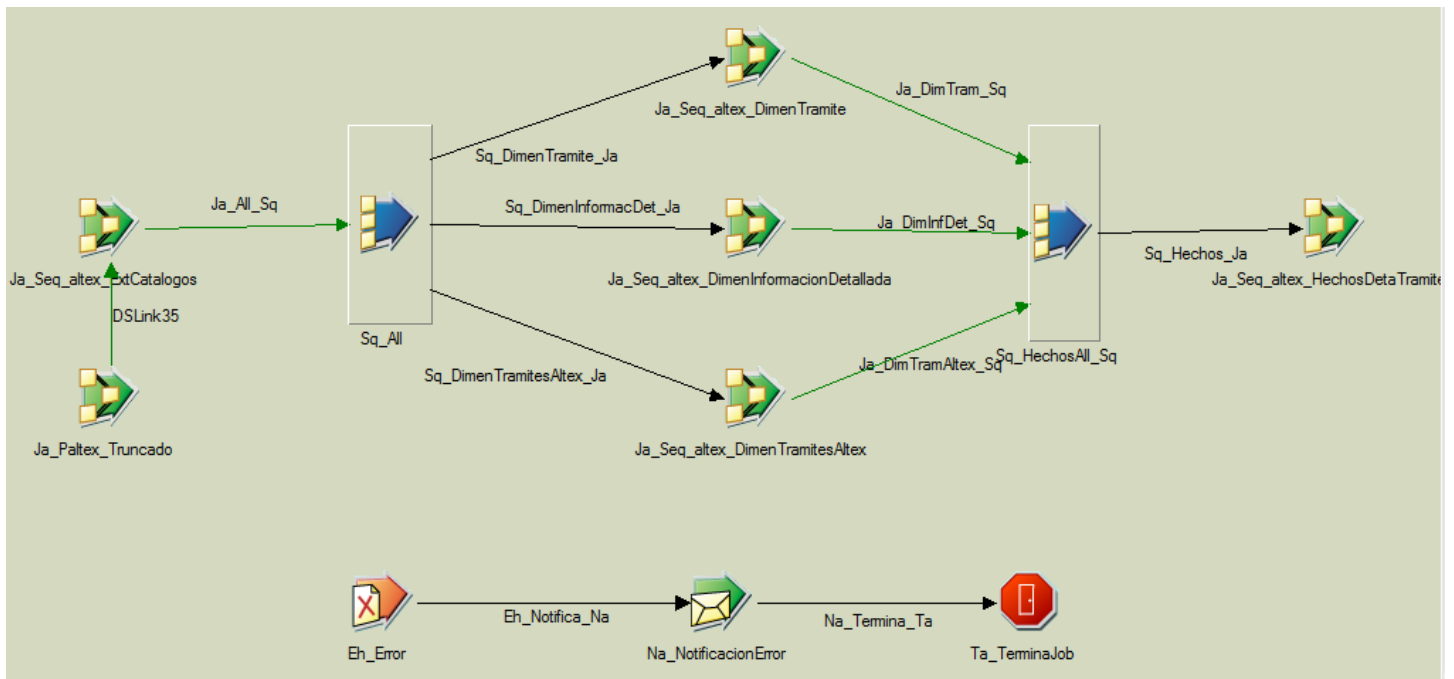


Figura 2.33 Secuencia maestra del datamart de la unidad 1

3.7 Unidad de negocio 2

3.7.1 ODS Unidad de negocio 2

3.7.1.1 Extracciones Unidad de negocio 2

En la extracción se genera un *job* reutilizable. Se compone de 2 *stages*, el primero con la conexión a base de datos de Oracle y el segundo es un archivo *dataset* en donde se guarda la extracción.

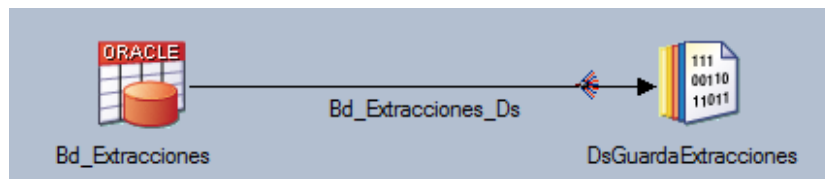


Figura 2.34 Job de extracciones de la unidad 2

3.7.1.2 Transformaciones Unidad de negocio 2

En la Figura 2.35 se renombran las columnas de origen para el destino, de manera que estandarizamos la información. En la Figura 2.36 para la transformación del job, solo se agregan las columnas de carga de usuario y fecha de carga. Para los procesos de la Figura 2.38 y Figura 2.39, podríamos decir que son procesos similares a la Figura 2.36.

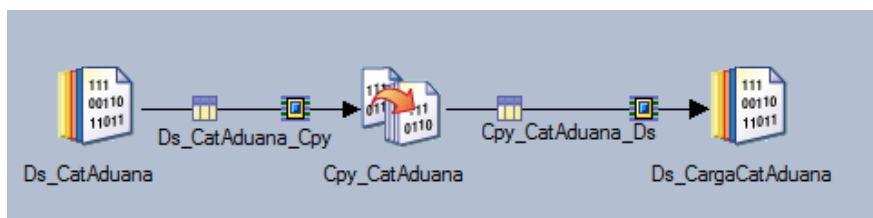


Figura 2.35 Job de transformación del catálogo de aduanas

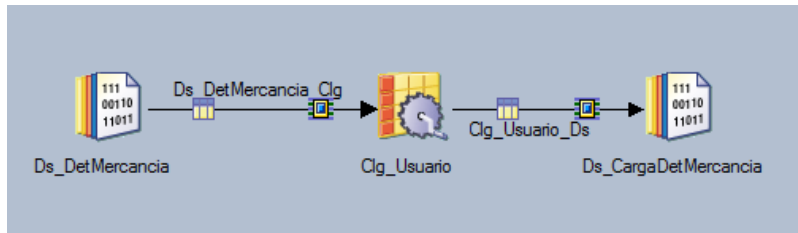


Figura 2.36 Job de transformación del detalle de la mercancía

Para el job de la Figura 2.37, se agrega información a la mercancía con el catálogo de isótopos y se generan nuevamente columnas de usuario y fecha de carga.

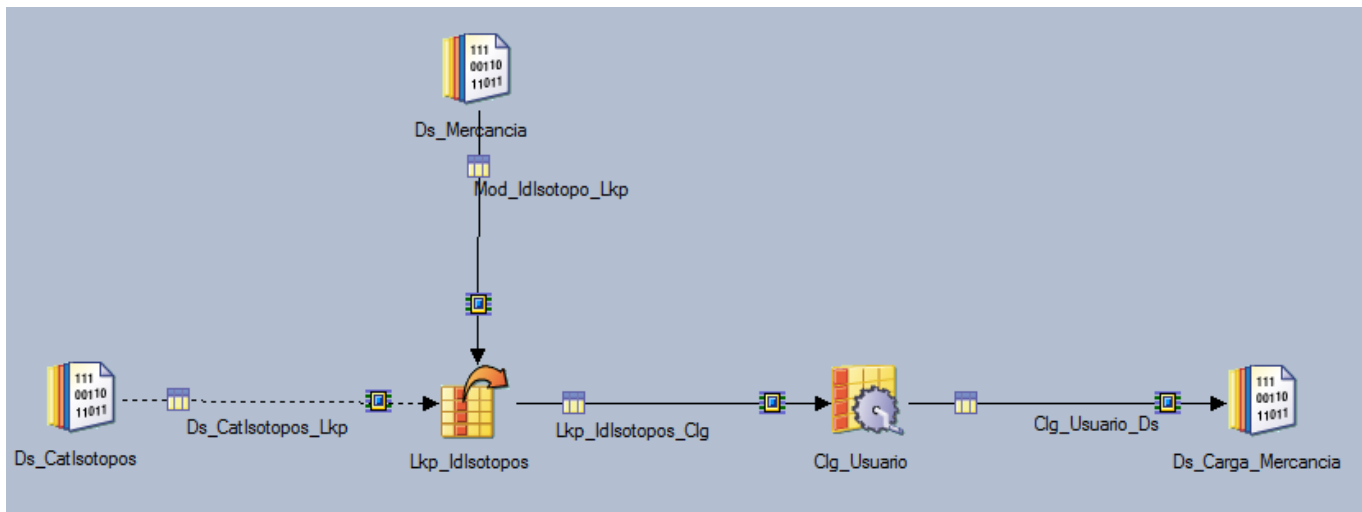


Figura 2.37 Job de transformación de la mercancía de la unidad 2

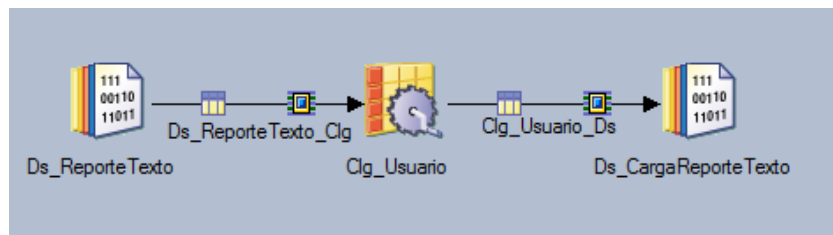


Figura 2.38 Job de transformación de los reportes de texto de la unidad 2

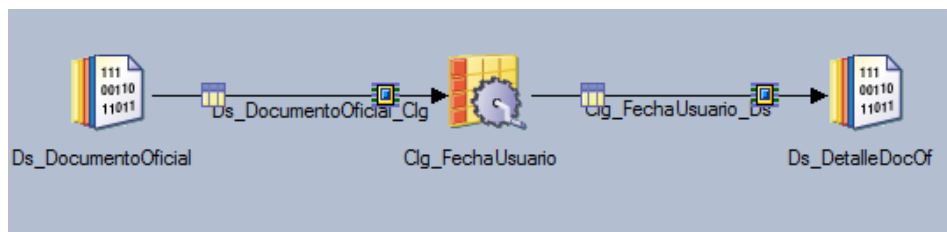


Figura 2.39 Job de transformación de los documentos oficiales de la unidad 2

3.7.1.3 Cargas Unidad de negocio 2

Para hacer una carga a una tabla, se tiene de inicio el archivo *dataset* generado en la transformación, el cual va hacia el stage de base de datos. Además se tiene un *dataset* de rechazo que indica los registros que por cualquier causa no pudieron ser insertados en la base de datos. (Figura 2.40, Figura 2.41, Figura 2.42, Figura 2.43, Figura 2.44)

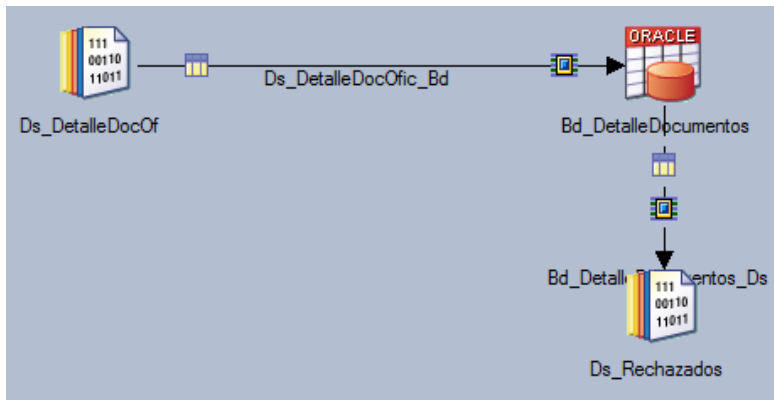


Figura 2.40 Job de carga de los documentos oficiales de la unidad 2

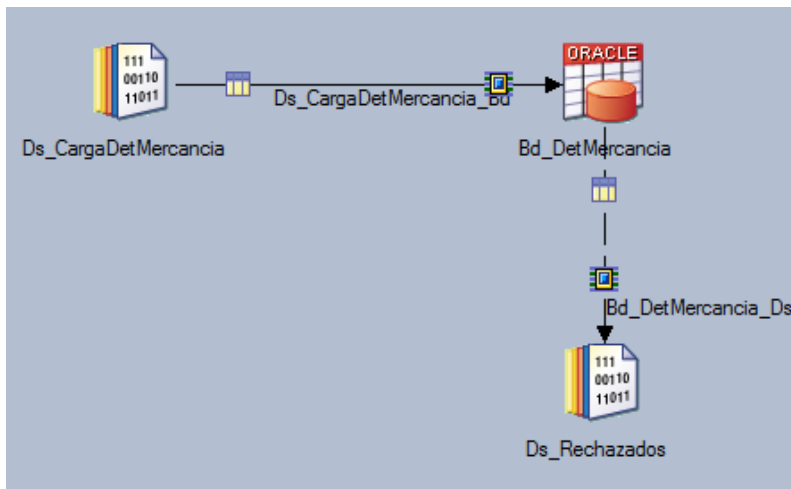


Figura 2.41 Job de carga del detalle de la mercancía de la unidad 2

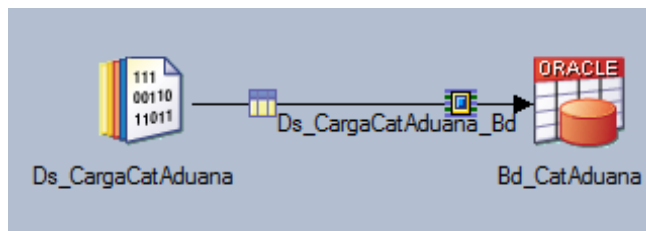


Figura 2.42 Job de carga del catálogo de la aduana

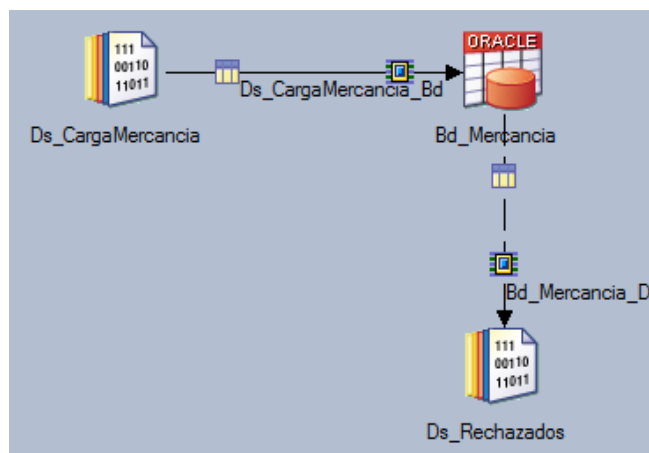


Figura 2.43 Job de carga de la mercancía de la unidad 2

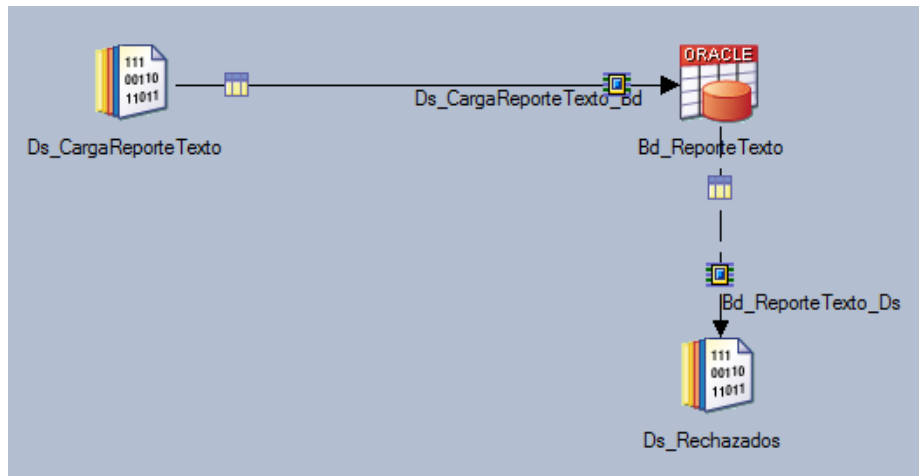


Figura 2.44 Job de carga de los reportes de texto de la unidad 2

3.7.1.4 Secuencia ODS Unidad de negocio 2

En este caso de secuencia del ODS de la unidad 2, solo se muestra la secuencia maestra la cual manda a llamar a la secuencia temporal, de extracción, de transformación y carga, ya que para esta unidad solo aporté el desarrollo de una de las secuencias más los desarrollos previos mostrados en esta misma unidad de negocio 2.

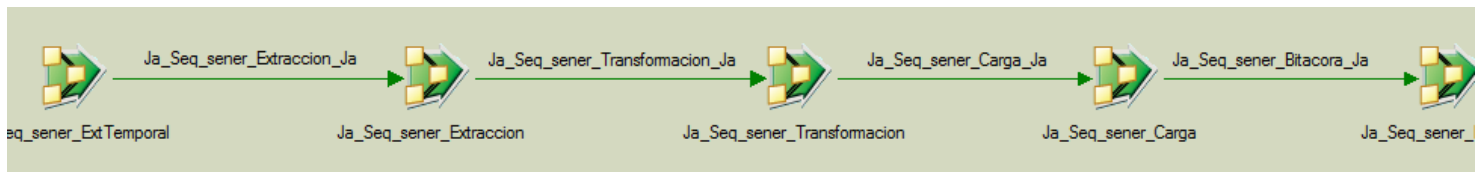


Figura 2.45 Secuencia maestra de la Unidad 2

3.7.2 DataMart Unidad de negocio 2

3.7.2.1 Transformaciones Unidad de negocio 2

Al igual que en la unidad 1, para cada una de las dimensiones, primero es necesario crear su llave subrogada. (Figura 2.46, Figura 2.48) posteriormente se crean las transformaciones relacionadas con cada dimensión (Figura 2.47, Figura 2.49) y su respectiva carga (Figura 2.50 Figura 2.51). Por último se crea la transformación de la tabla de hechos con las dimensiones previamente cargadas y se hace su respectiva carga.

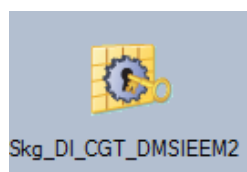


Figura 2.46 Job de creación de la llave subrogada de la dimensión de la mercancía de la unidad 2

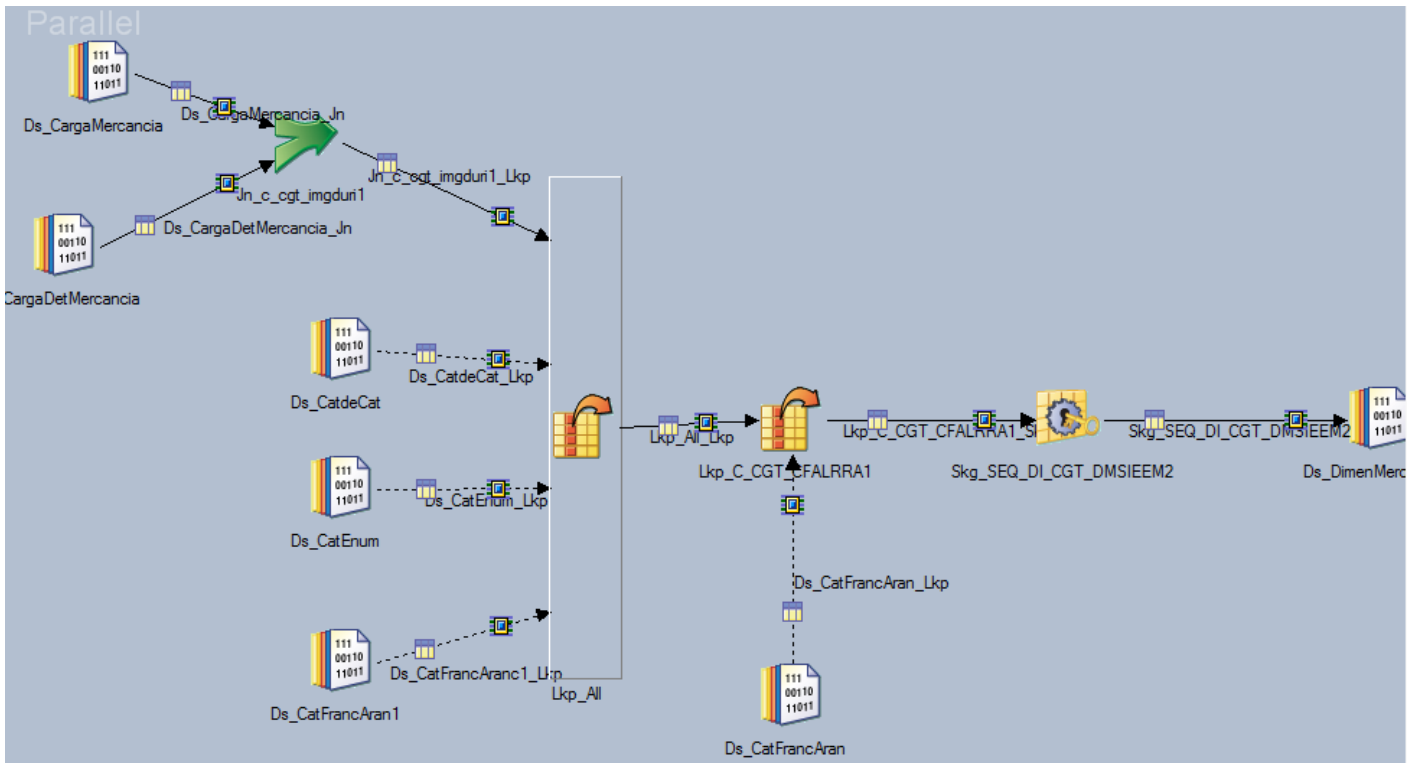


Figura 2.47 Job de transformación de la dimensión de la mercancía de la unidad 2



Figura 2.48 Job de creación de la llave subrogada de la dimensión del reporte de texto libre de la unidad 2

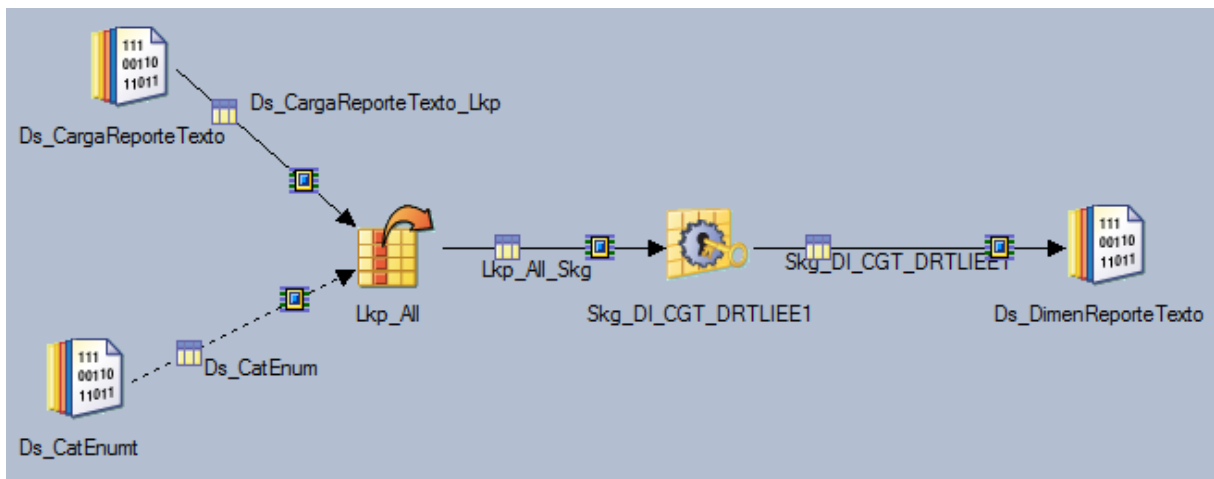


Figura 2.49 Job de transformación de la dimensión del reporte de texto libre de la unidad 2

3.7.2.2 Cargas Unidad de negocio 2

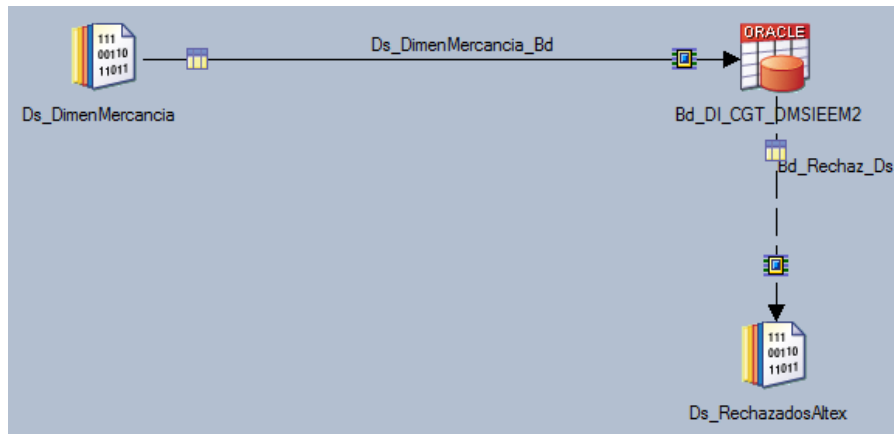


Figura 2.50 Job de carga de la dimensión de la mercancía de la unidad 2

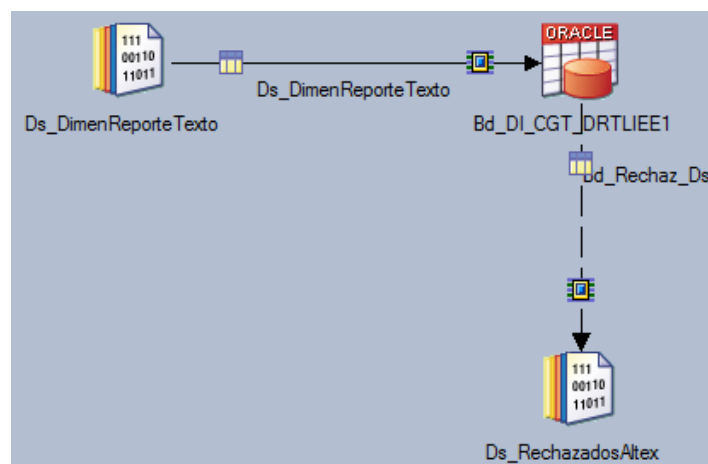


Figura 2.51 Job de carga de la dimensión del reporte de texto de la unidad 2

3.7.2.3 Secuencias *Datamart* Unidad de negocio 2

En el caso de un datamart, un job de secuencia se crea por dimensión, de manera que una secuencia contempla la transformación y la carga de cada dimensión (Figura 2.52) Después se crea la secuencia de la tabla de hechos, que contenía tanto su transformación como carga. Finalmente se creaba la secuencia maestra por *datamart* (Figura 2.53)

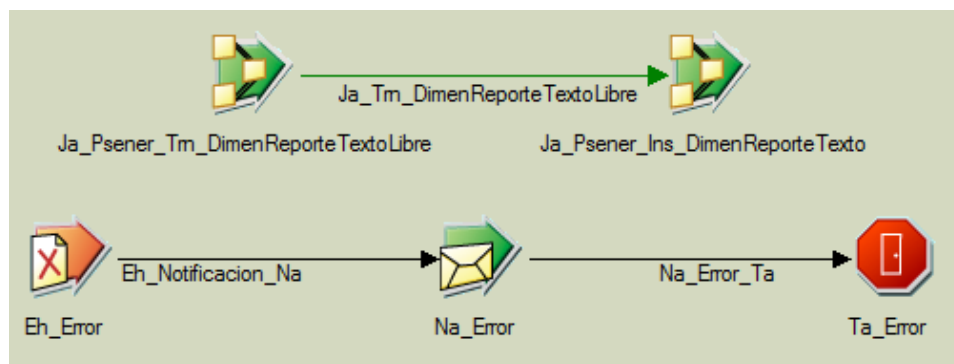


Figura 2.52 Secuencia de la dimensión del reporte de texto de la unidad 2

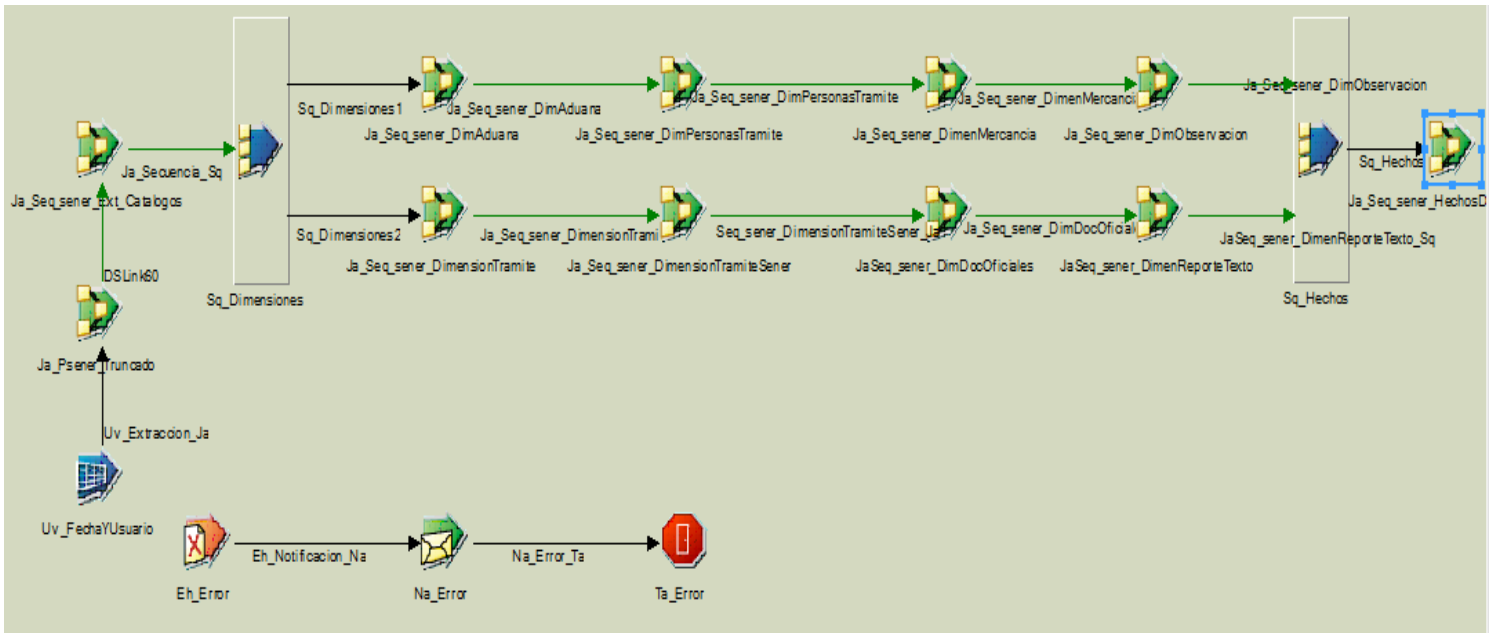


Figura 2.53 Secuencia maestra del datamart de la unidad 2

3.8 Unidad de negocio 3

3.8.1 ODS Unidad de negocio 3

3.8.1.1 Extracciones Unidad de negocio 3

Para las extracciones, se utiliza un job de extracción, el cual es reutilizable. En la Figura 2.54 se muestra la extracción de ODS para la unidad 3.

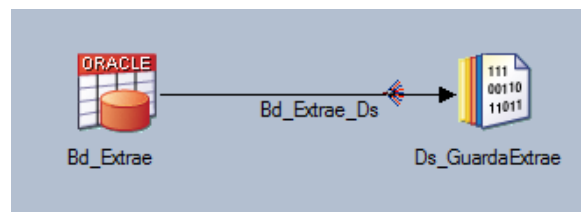


Figura 2.54 Job de extracción de la unidad 3

3.8.1.2 Transformaciones Unidad de negocio 3

La Figura 2.55 es un *job* de transformación donde solo se le agrega información a la declaración de la solicitud y se indica quien y en qué fecha se hizo el desarrollo.

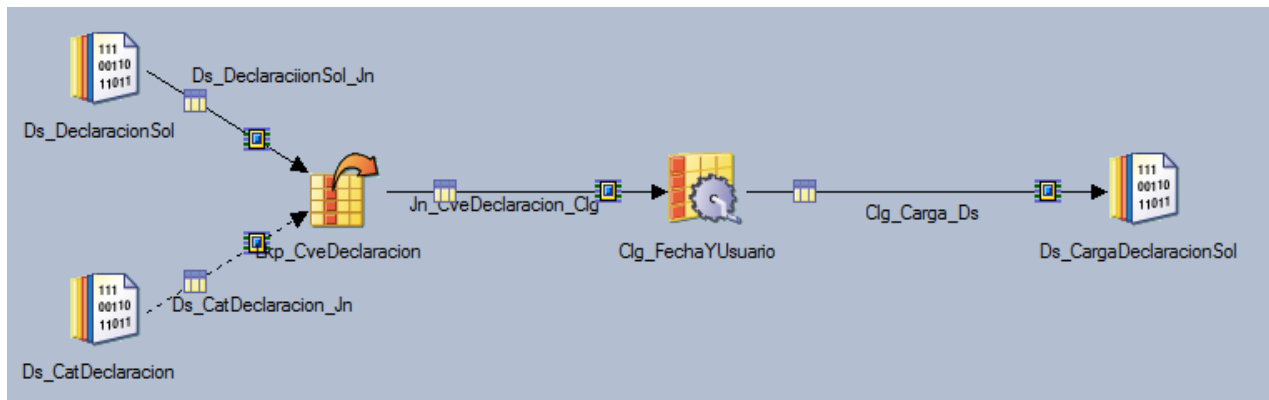


Figura 2.55 Job de transformación de la declaración de la solicitud de la unidad 3

Para la figura 2.56 la cual tiene campos tipo *clob*, se hizo un *substring* de los datos para que datastage lo pudiera transformar, ya que tiene un límite y no puede procesar los datos tipo *clob* en un mismo campo.

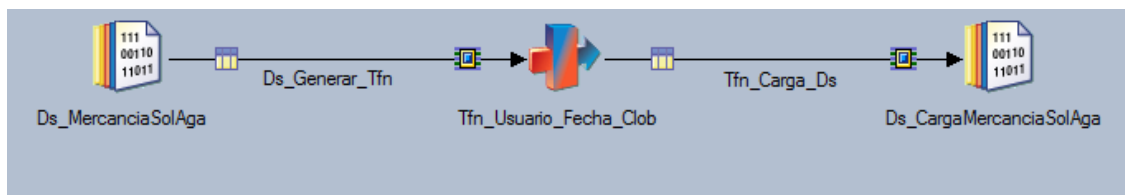


Figura 2.56 Job de transformación del detalle de la mercancía de la solicitud de la unidad 3

En la Figura 2.57 se transforman las observaciones del trámite, agregando las personas asociadas a la observación a partir de su clave de usuario. En seguida solo se agregan columnas para saber cuándo y quien hizo el desarrollo.

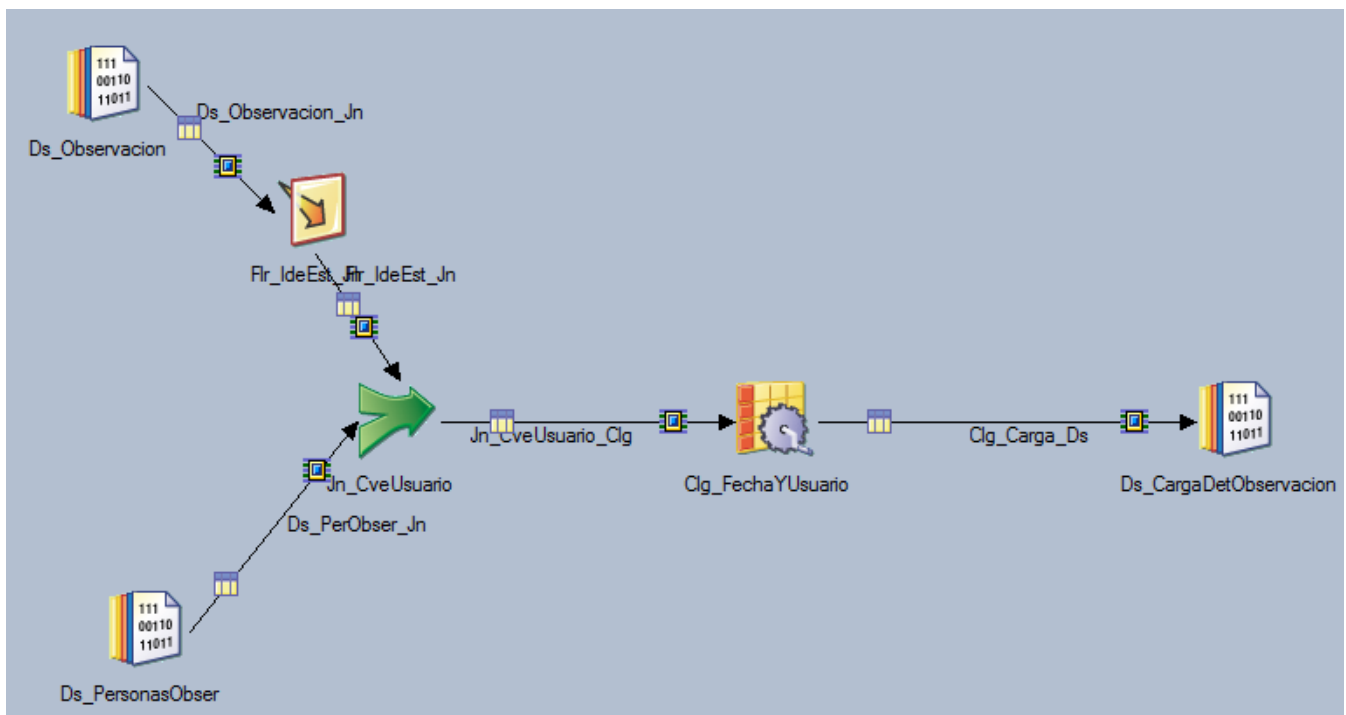


Figura 2.57 Job de transformación del detalle de la observación de la unidad 3

Para la Figura 2.58, el *job* de transformación es para que solo se modificaran los campos en caso que vinieran nulos, se ponía 0.

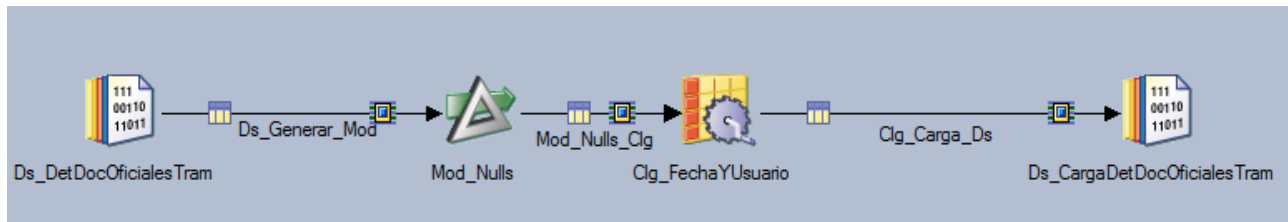


Figura 2.58 Job de transformación del detalle de los documentos oficiales del trámite de la unidad 3

En la Figura 2.59 para el requerimiento del trámite, solo se agregaron las personas asociadas al requerimiento, a partir del identificador del requerimiento. Además se hizo un *substring* de los campos tipo *clob*.

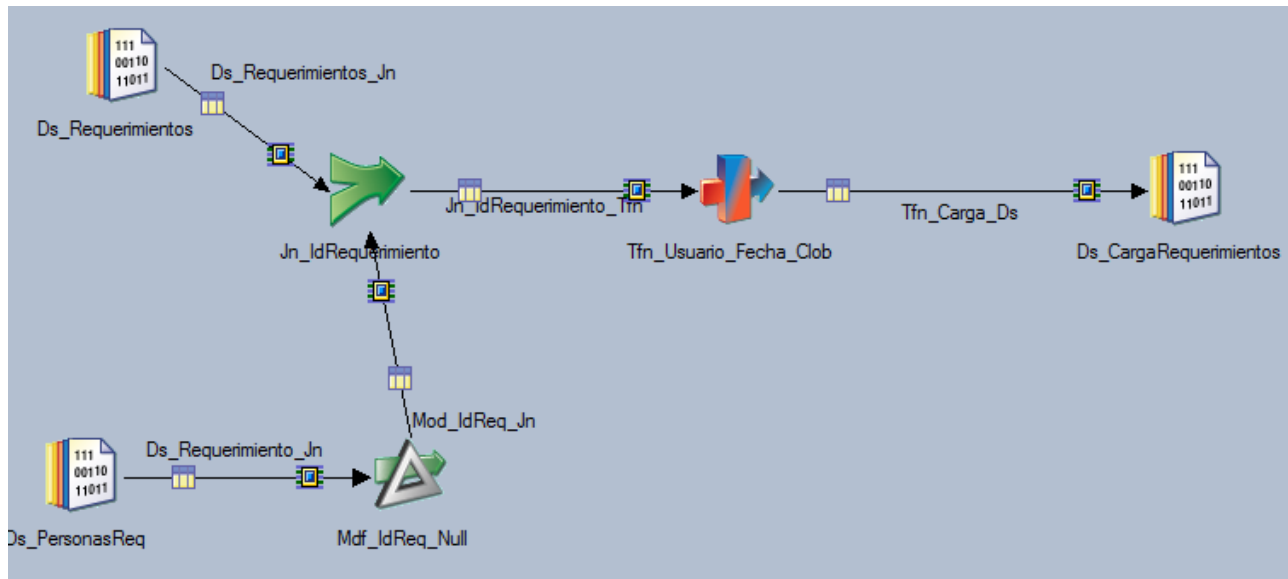


Figura 2.59 Job de transformación del detalle de los requerimientos de la unidad 3

En la Figura 2.60 se agregaron los datos necesarios para el detalle del atributo de la solicitud de la mercancía. Al final se agregaron los campos para identificar que usuario y cuando se hizo el desarrollo.

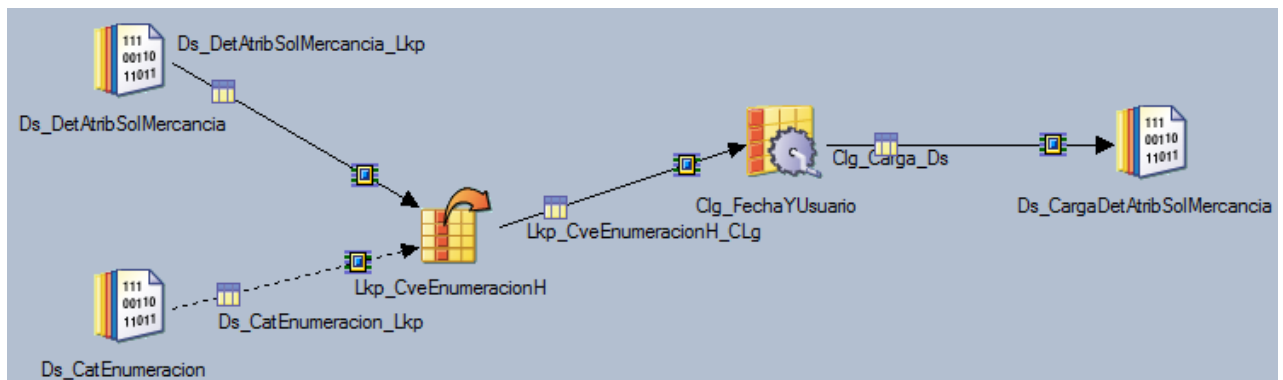


Figura 2.60 Job de transformación del detalle del atributo de la solicitud de la mercancía

Para la Figura 2.61, el *job* de transformación es para que solo se modificaran los campos en caso que vinieran nulos, se ponía 0.

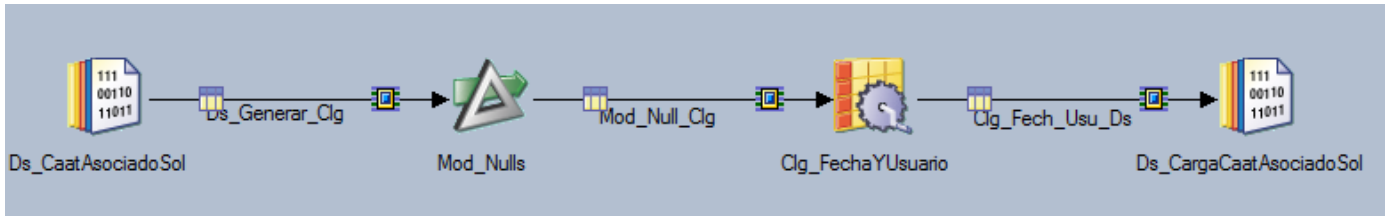


Figura 2.61 Job de transformación del detalle del CaatAsociado a la Solicitud

Para la figura 2.62 la cual tiene campos tipo *clob*, se hizo un *substring* de los datos para que datastage lo pudiera transformar, ya que tiene un límite y no puede procesar los datos tipo *clob* en un mismo campo.



Figura 2.62 Job de transformación de la mercancía asociada a la solicitud de la unidad 3

3.8.1.3 Cargas Unidad de negocio 3

Para cada carga a una tabla, se tiene de inicio el archivo *dataset* generado en la transformación, el cual va hacia el stage de base de datos, además, se agrega un dataset de rechazo que indica los registros que por cualquier causa no pudieron ser insertados en la base de datos. Las cargas al ODS de la unidad 3 se muestran en las imágenes: Figura 2.63, Figura 2.64, Figura 2.65, Figura 2.66, Figura 2.67, Figura 2.68, Figura 2.69, Figura 2.70.

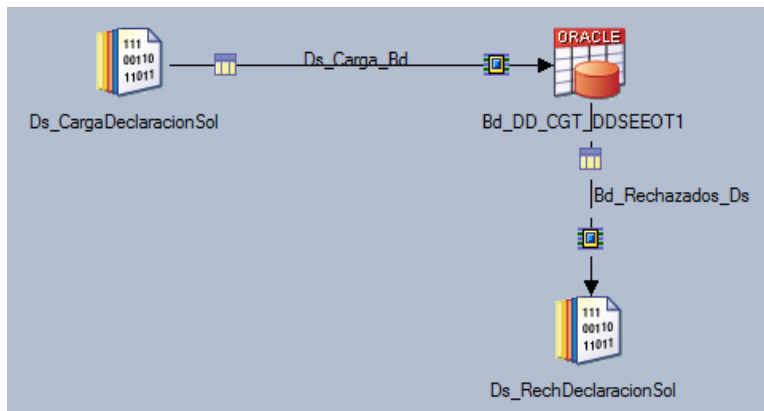


Figura 2.63 Job de carga de la declaración de la solicitud de la unidad 3

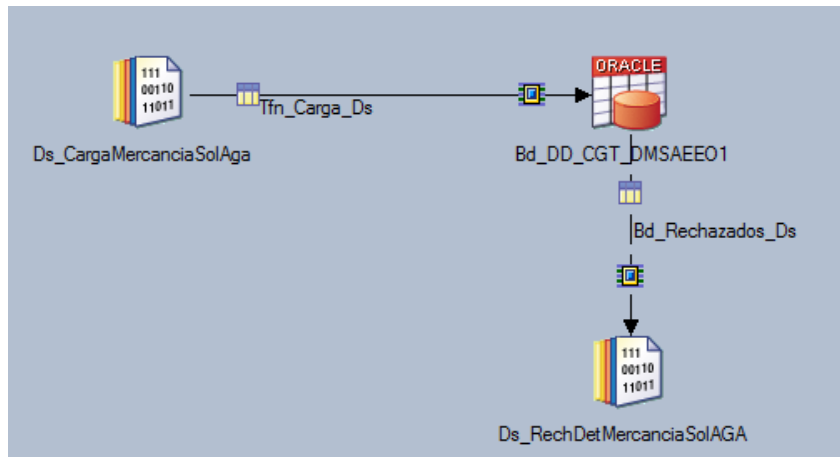


Figura 2.64 Job de carga del detalle de la mercancía de la solicitud de la unidad 3

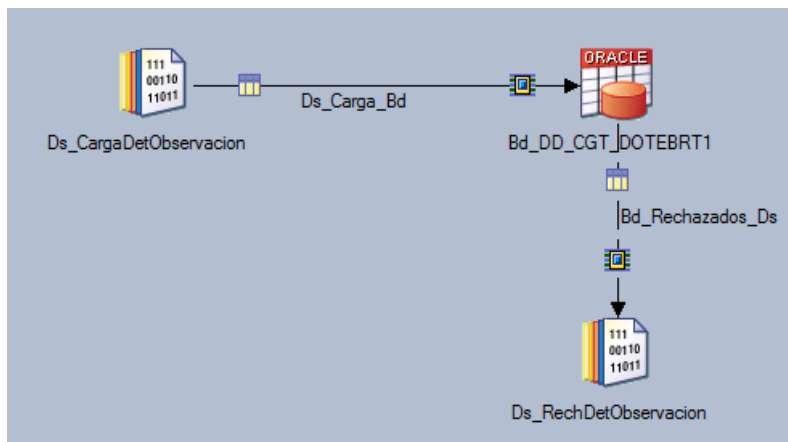


Figura 2.65 Job de carga del detalle de la observación de la unidad 3

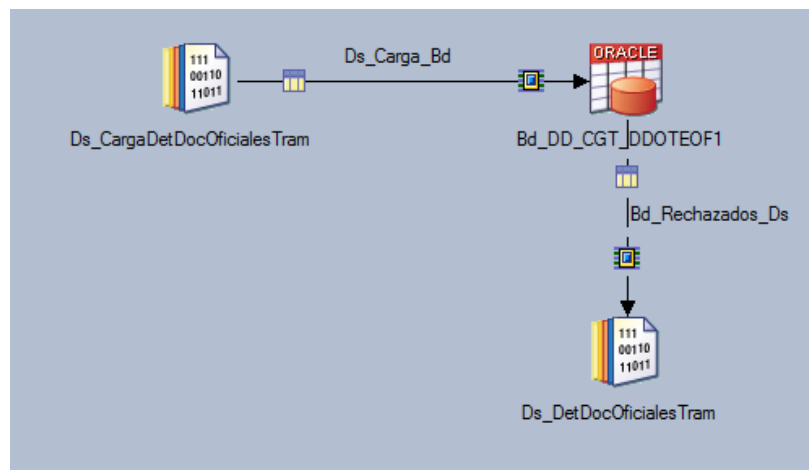


Figura 2.66 Job de carga del detalle de los documentos oficiales de la unidad 3

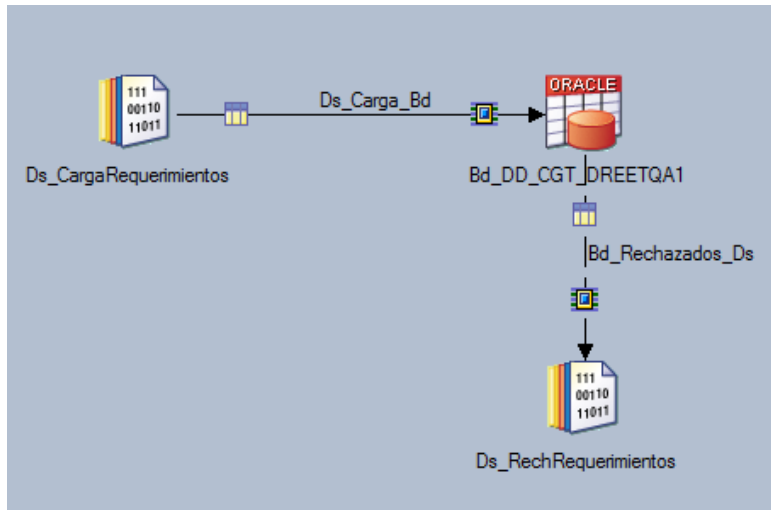


Figura 2.67 Job de carga del detalle de los requerimientos de la unidad 3

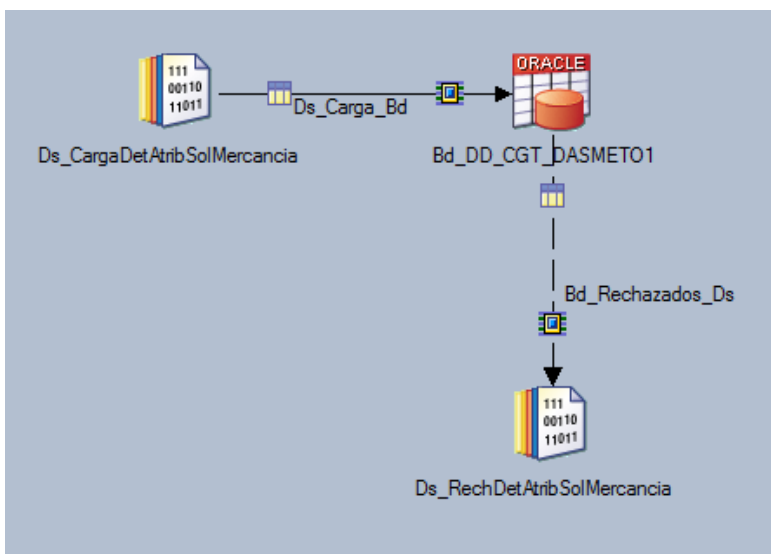


Figura 2.68 Job de carga del detalle de los atributos de la solicitud de la mercancia

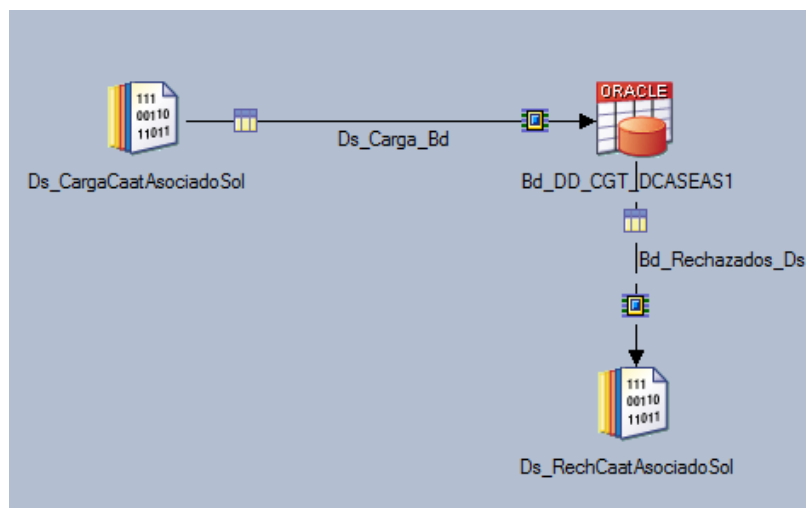


Figura 2.69 Job de carga del caat asociado a la solicitud de la unidad 3

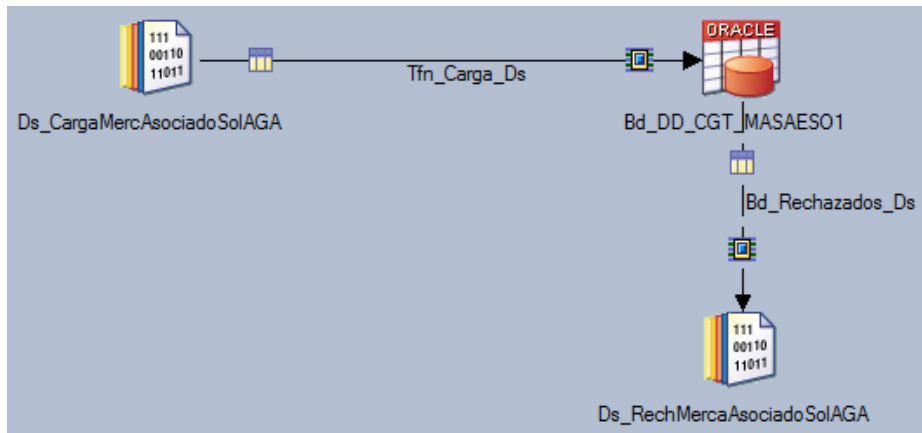


Figura 2.70 Job de carga de la mercancía asociada a la solicitud de la unidad 3

3.8.1.4 Secuencia ODS Unidad de negocio 3

En cada unidad se generaron las secuencias temporal, extracción, transformación y carga de datos, así como la secuencia maestra que contiene las 4 secuencias antes mencionadas. La secuencia maestra del ODS de la unidad 3 es mostrada en la Figura 2.71

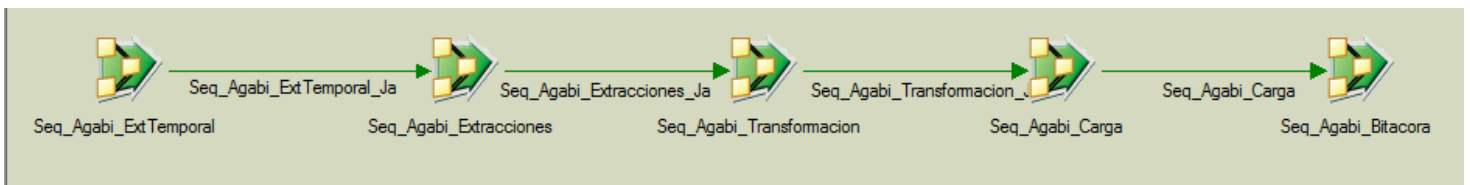


Figura 2.71 Secuencia maestra del ODS de la unidad 3

3.8.2 DataMart Unidad de negocio 3

3.8.2.1 Transformaciones Unidad de negocio 3

Al igual que en la unidad 1 y unidad 2, en la unidad 3 para cada una de las dimensiones, primero es necesario crear su llave subrogada. (Figura 2.72, Figura 2.74, Figura 2.76, Figura 2.78, Figura 2.80, Figura 2.82) posteriormente se crean las transformaciones relacionadas con cada dimensión (Figura 2.73, Figura 2.75, Figura 2.77, Figura 2.79, Figura 2.81). Por último se crea la transformación de la tabla de hechos agregando las dimensiones previamente cargadas (Figura 2.83).

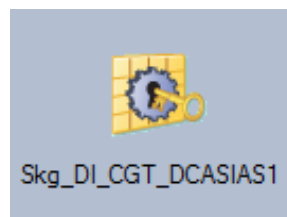


Figura 2.72 Job de creación de llave subrogada de la dimensión de CAAT de la unidad 3

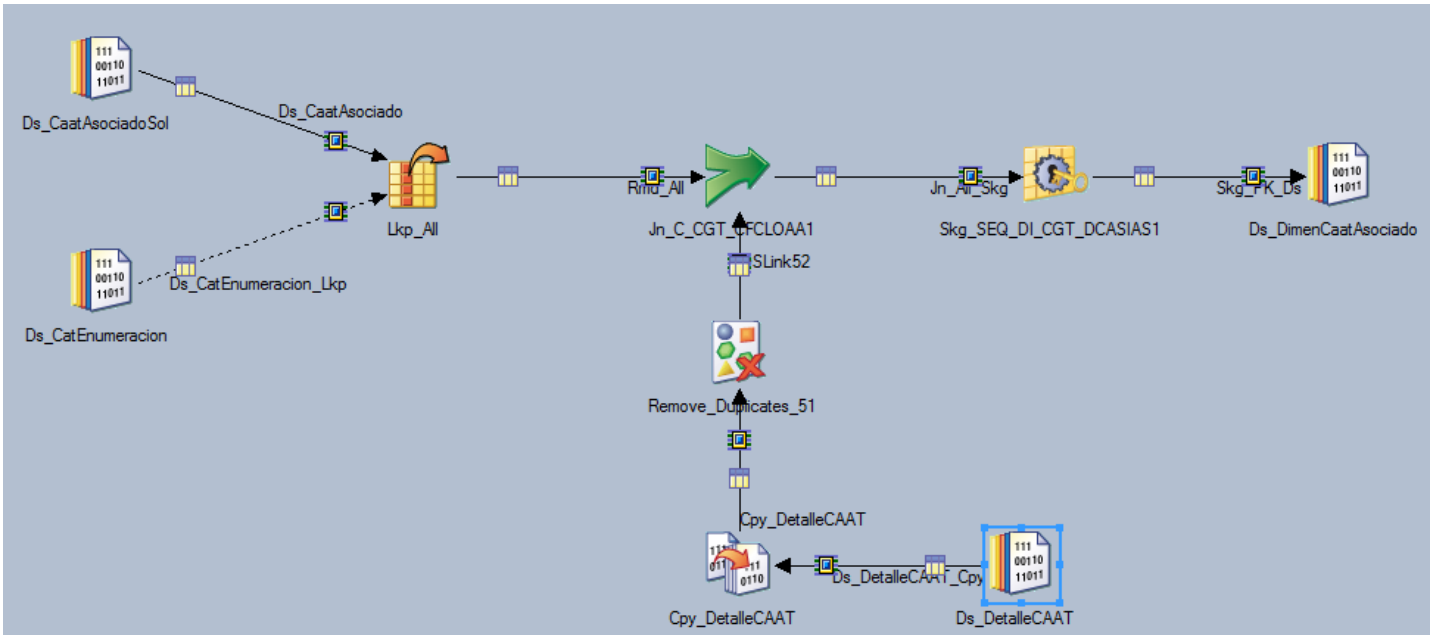


Figura 2.73 Job de transformación de la dimensión de CAAT de la unidad 3

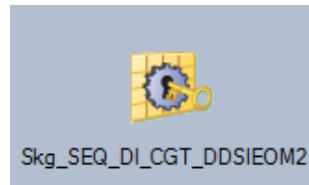


Figura 2.74 Job de creación de llave subrogada de la dimensión de la declaración de la solicitud de la unidad 3

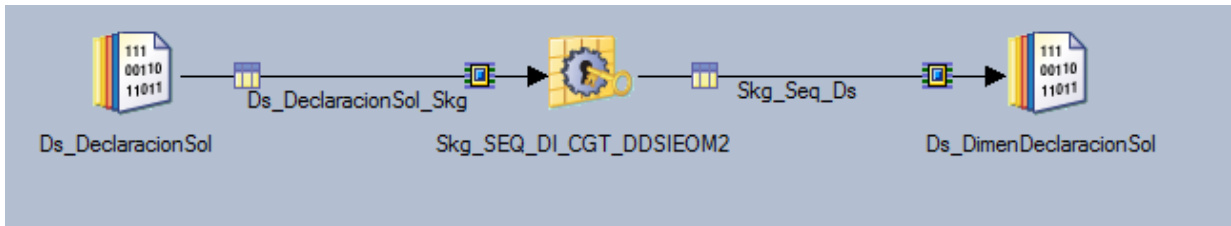


Figura 2.75 Job de transformación de la dimensión de la declaración de la solicitud de la unidad 3



Figura 2.76 Job de creación de llave subrogada de la dimensión de los documentos oficiales de la unidad3

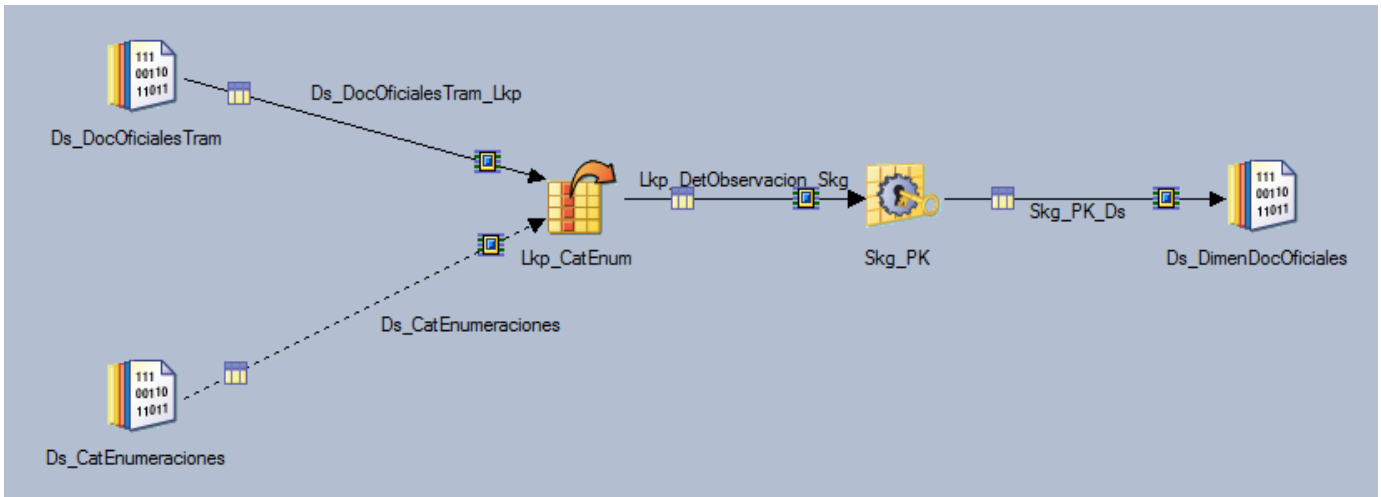


Figura 2.77 Job de transformación de la dimensión de los documentos oficiales de la unidad 3

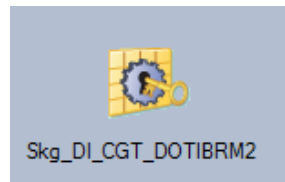


Figura 2.78 Job de creación de llave subrogada de la dimensión de las observaciones del trámite de la unidad 3

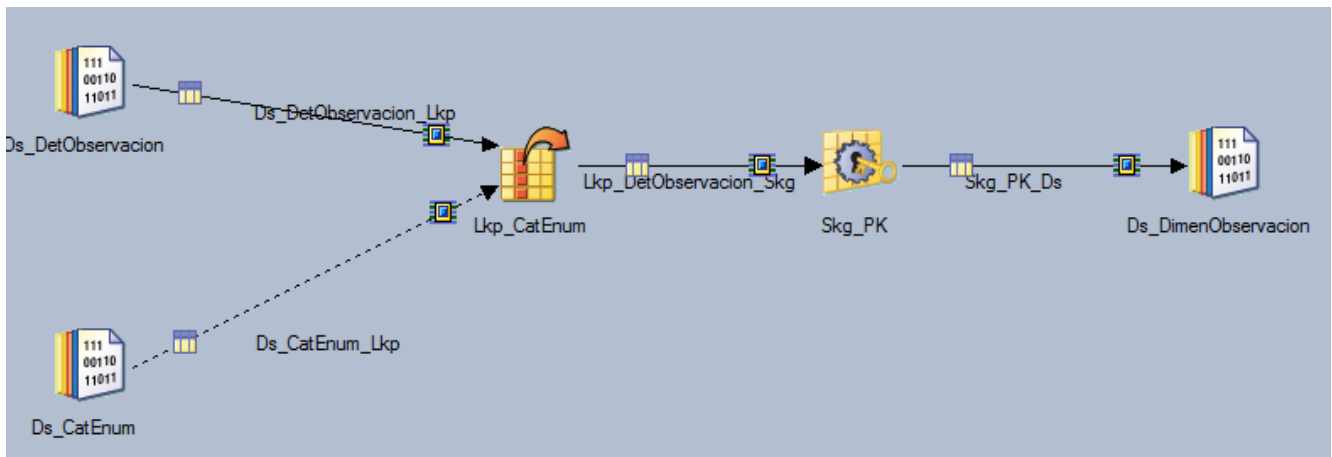


Figura 2.79 Job de transformación de la dimensión de las observaciones del trámite



Figura 2.80 Job de creación de la llave subrogada de la dimensión de los requerimientos de la unidad 3

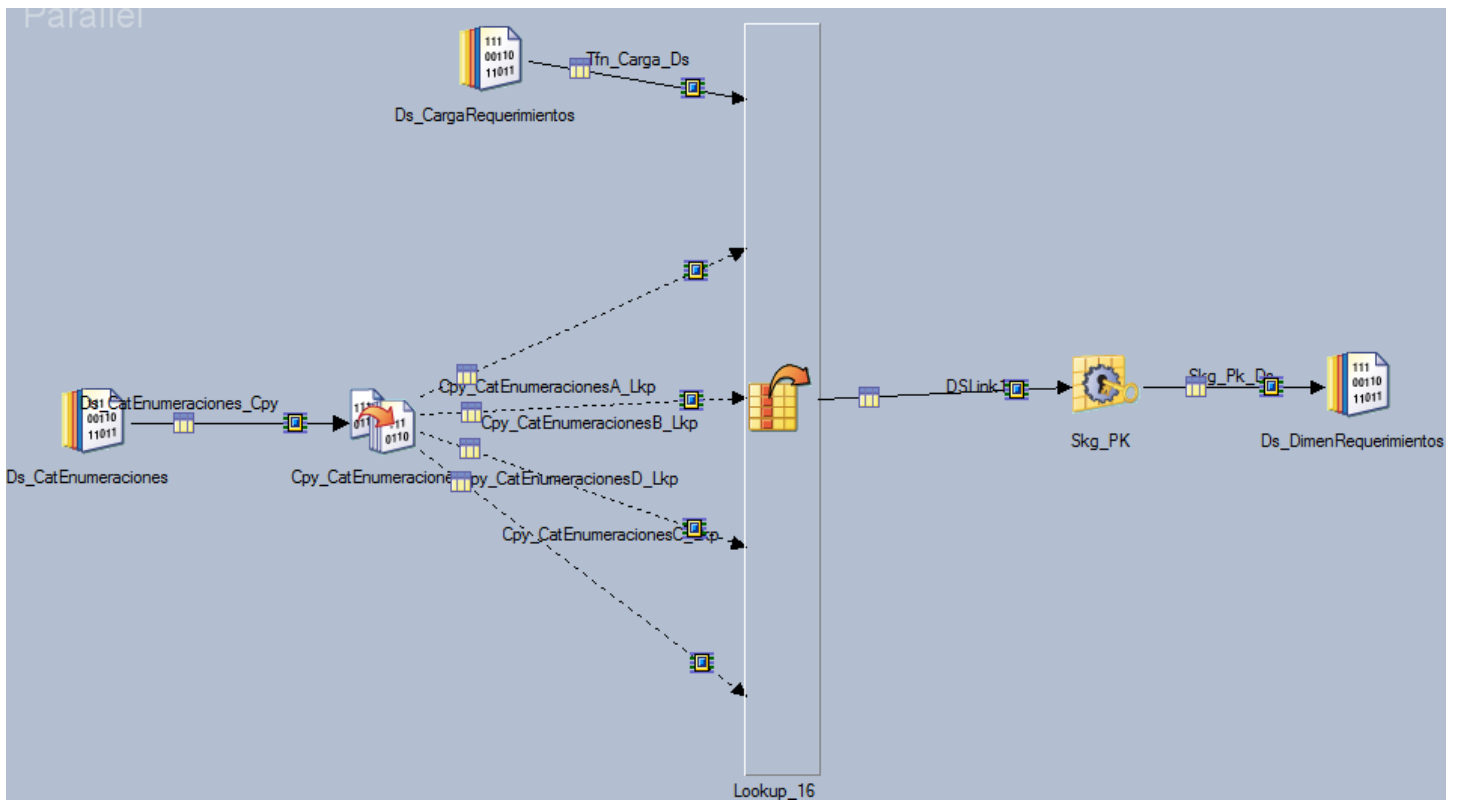


Figura 2.81 Job de transformación de la dimensión de los requerimientos de la unidad 3

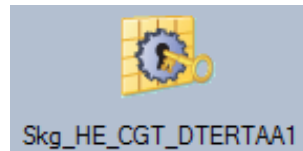


Figura 2.82 Job de creación de llave subrogada de los hechos del detalle del trámite de la unidad 3

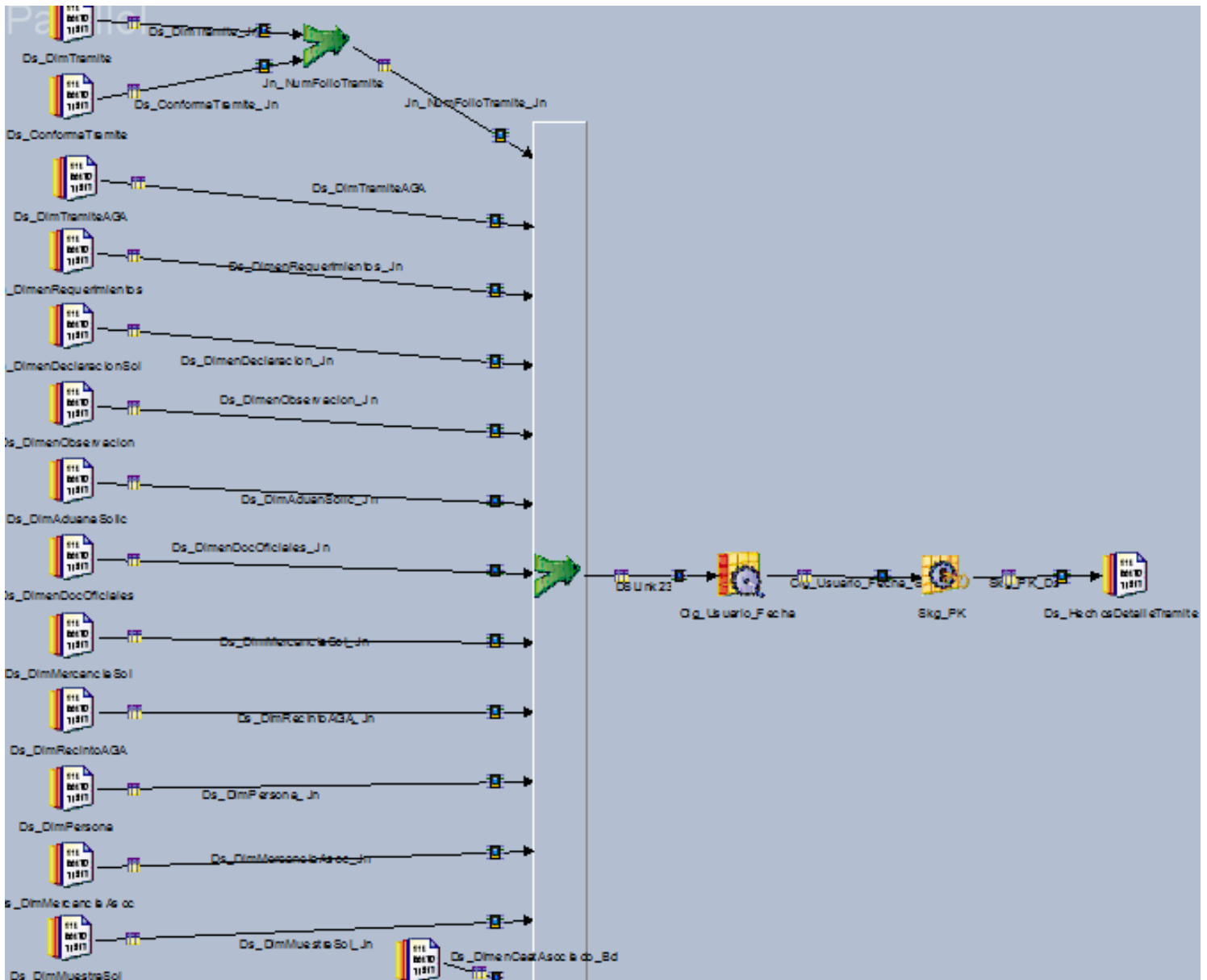


Figura 2.83 Job de transformación de los hechos del detalle del trámite de la unidad 3

3.8.2.2 Cargas Unidad de negocio 3

Para cada carga a una tabla, se tiene de inicio el archivo *dataset* generado en la transformación, el cual va hacia el stage de base de datos, además, se agrega un dataset de rechazo que indica los registros que por cualquier causa no pudieron ser insertados en la base de datos. Las cargas al *datamart* de la unidad 3 se muestran en las imágenes: Figura 2.84, Figura 2.85, Figura 2.86, Figura 2.87, Figura 2.88, Figura 2.89.

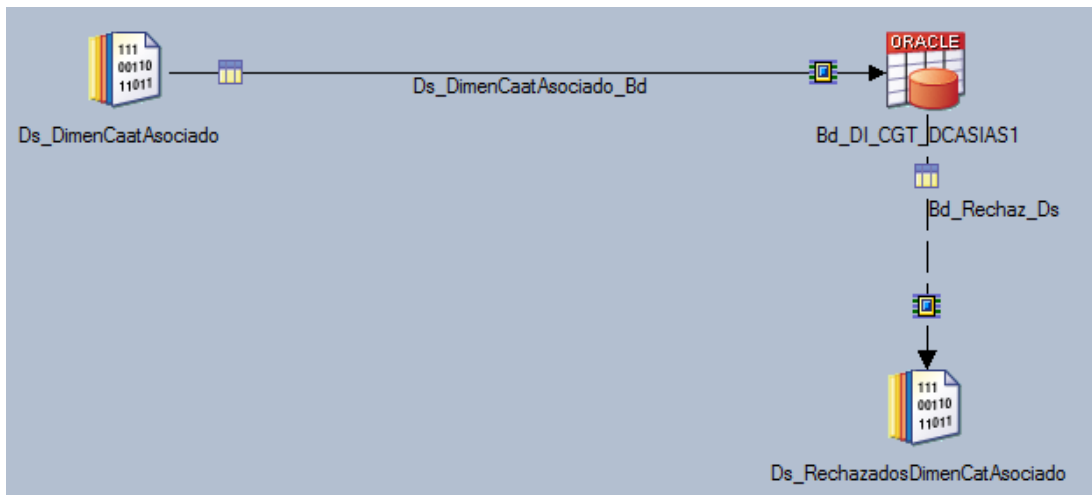


Figura 2.84 Job de carga de la dimensión del CAAT

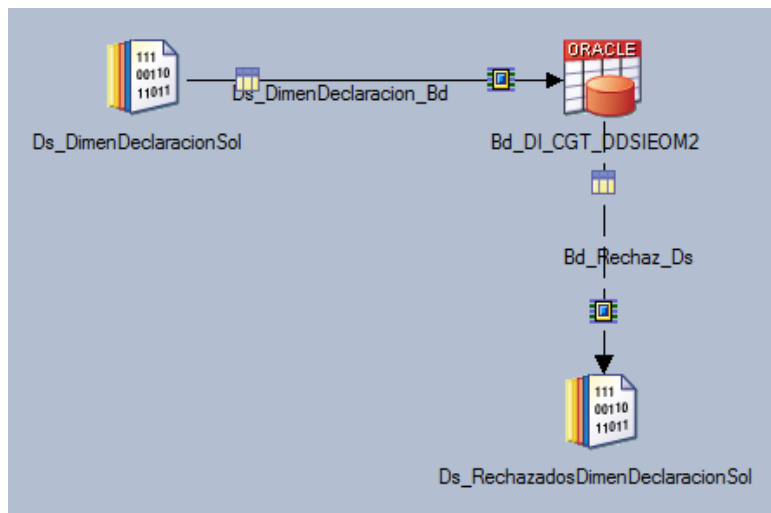


Figura 2.85 Job de carga de la dimensión de la declaración de la solicitud de la unidad 3

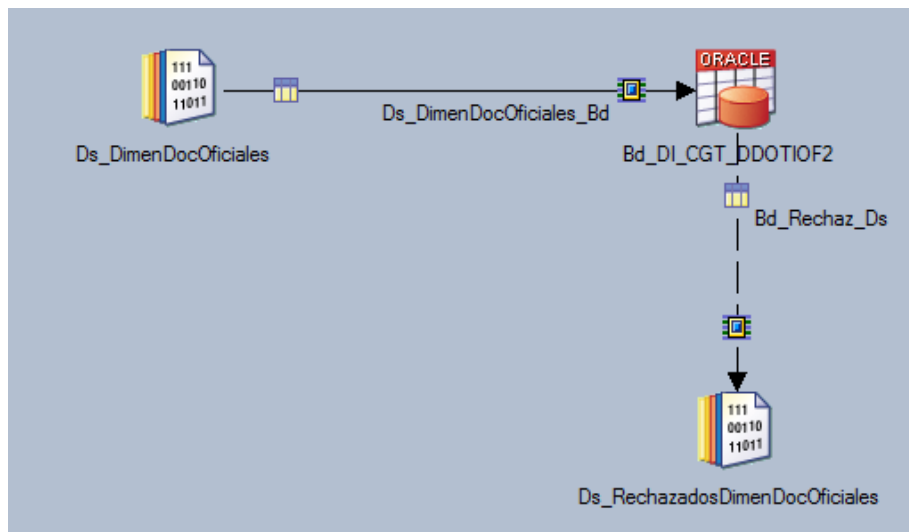


Figura 2.86 Job de carga de la dimensión de los documentos oficiales de la unidad 3

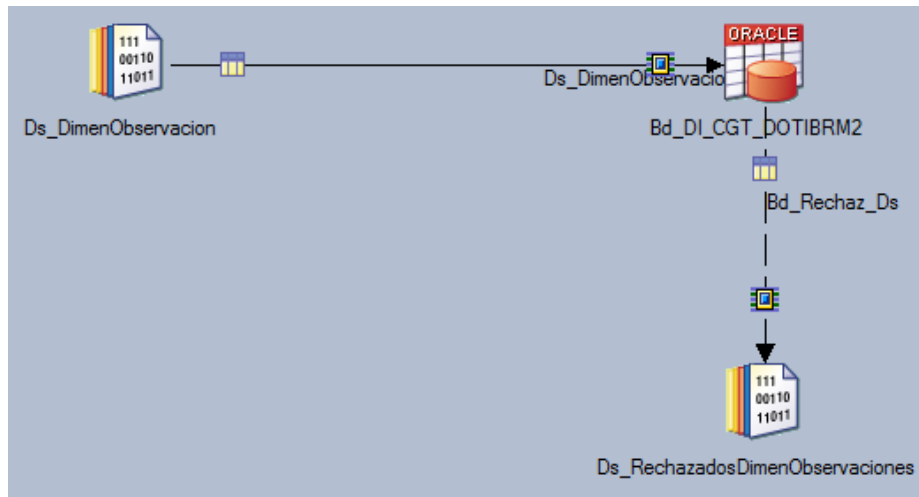


Figura 2.87 Job de carga de la dimensión de las observaciones del trámite de la unidad 3

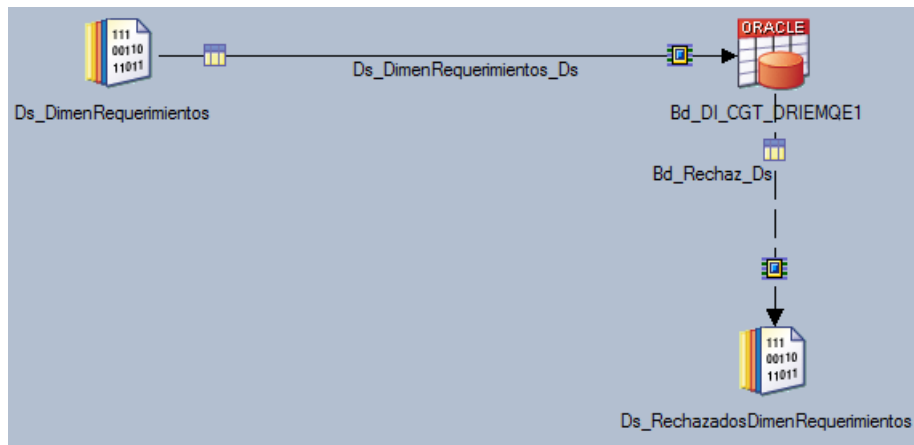


Figura 2.88 Job de carga de la dimensión del requerimiento de la unidad 3

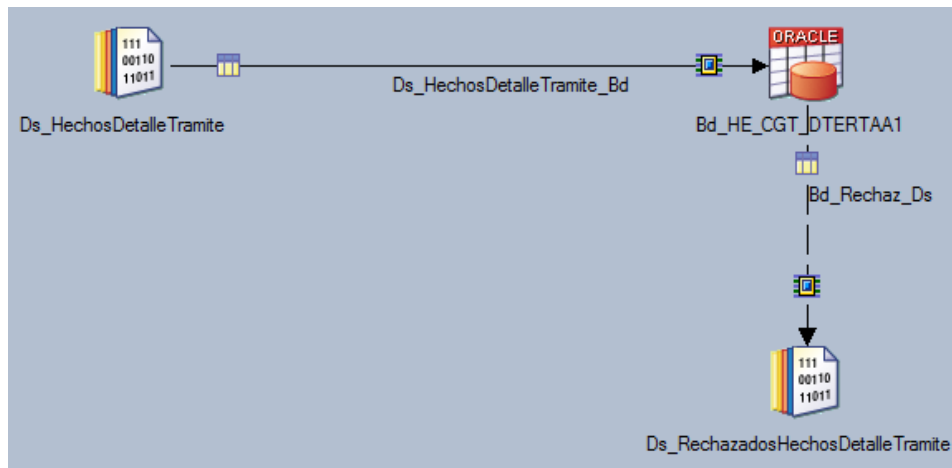


Figura 2.89 Job de carga de los hechos del detalle del trámite de la unidad 3

3.8.2.3 Secuencias Datamart Unidad de negocio 3

En el *datamart* de la unidad 3 se generaron las secuencias de transformación y carga de datos de cada dimensión (Figura 2.90, Figura 2.91, Figura 2.92, Figura 2.93 y Figura 2.94) y de cada tabla de

hechos (Figura 2.95). Además también se genera la secuencia maestra que contiene las secuencias antes mencionadas. La secuencia maestra del *datamart* de la unidad 3 es mostrada en la Figura 2.96

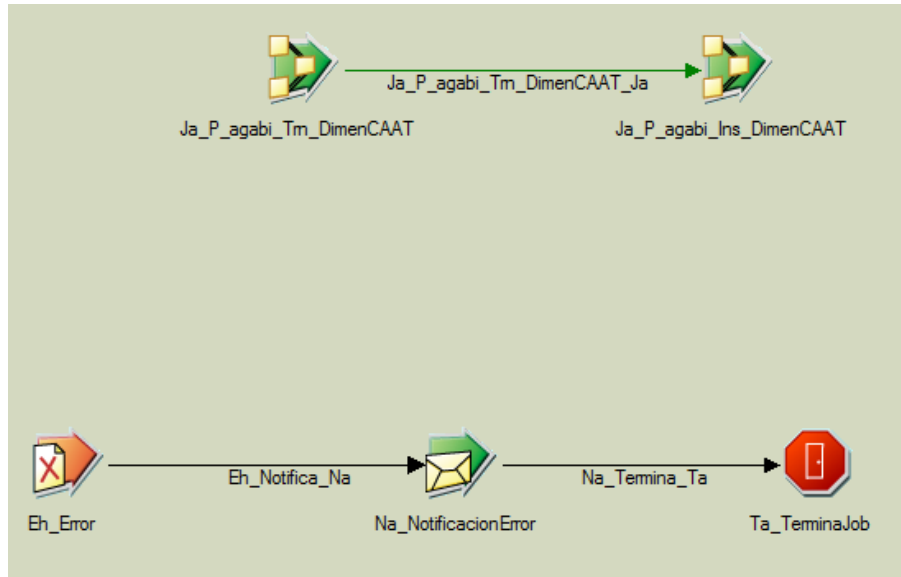


Figura 2.90 Secuencia de la dimensión del CAAT de la unidad 3

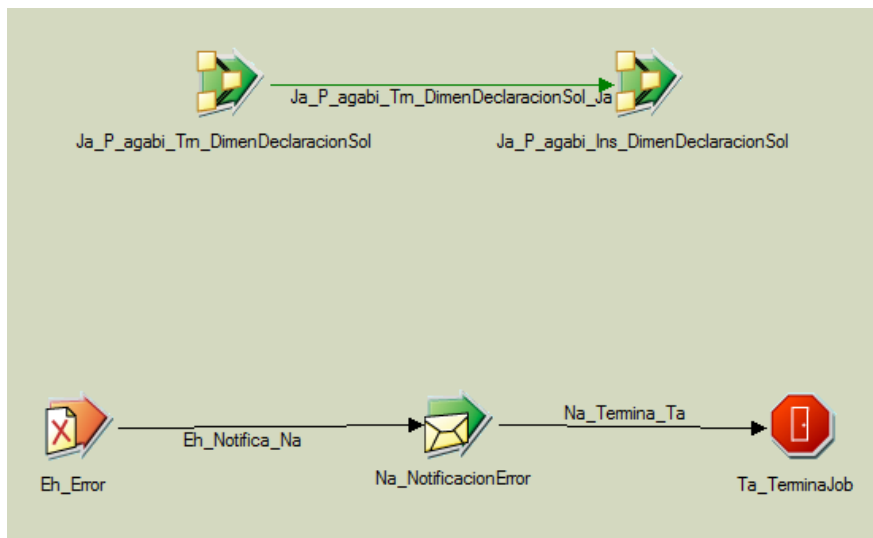


Figura 2.91 Secuencia de la dimensión de la declaración de la solicitud de la unidad 3

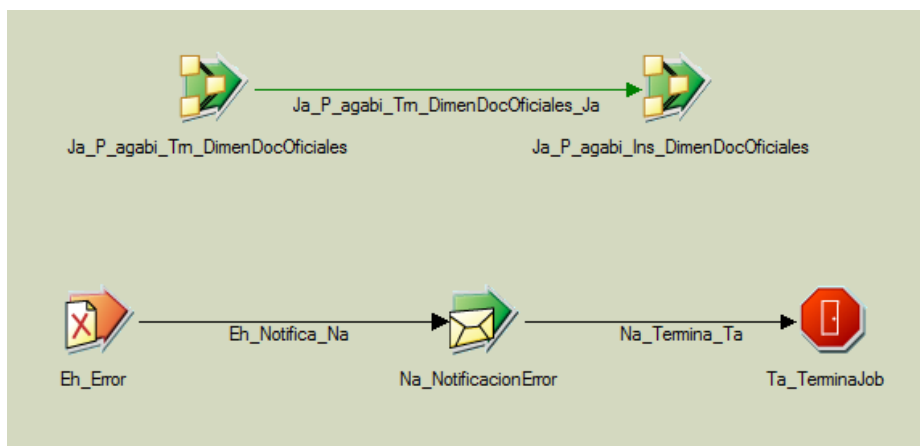


Figura 2.92 Secuencia de la dimensión de los documentos oficiales de la unidad 3

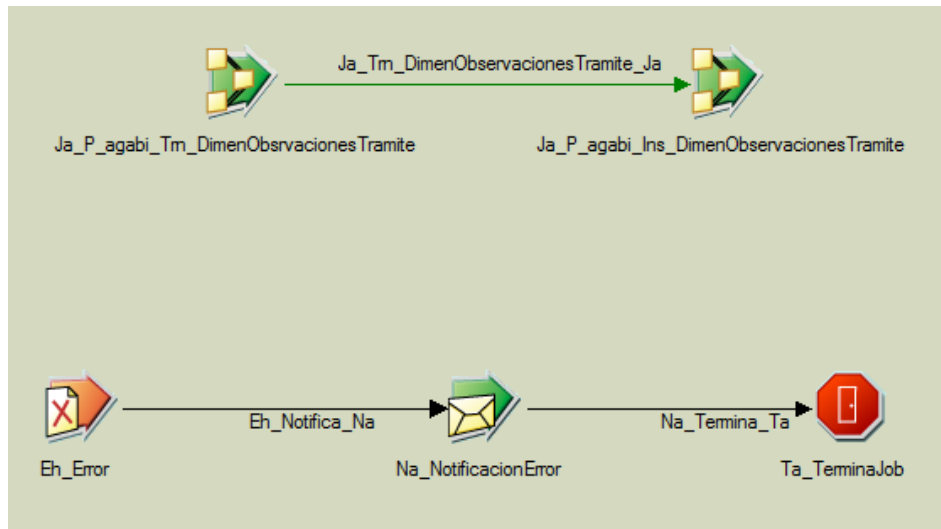


Figura 2.93 Secuencia de la dimensión de las observaciones del trámite de la unidad 3

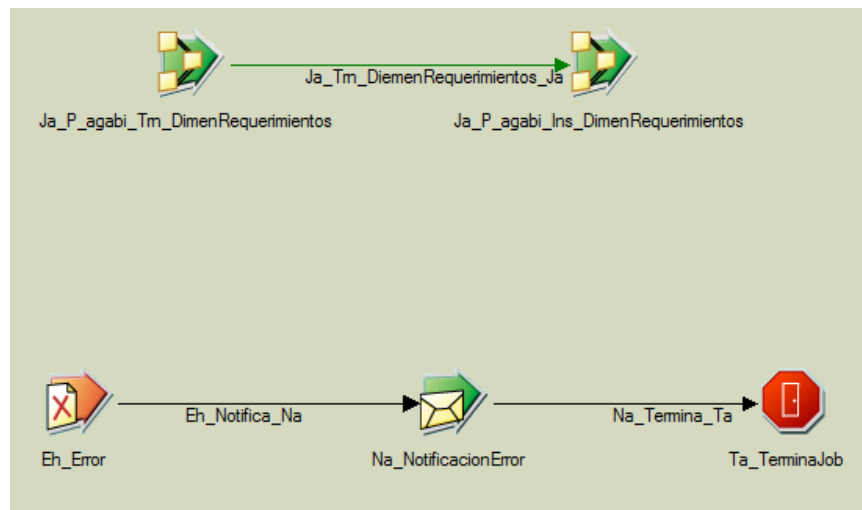


Figura 2.94 Secuencia de la dimensión del requerimiento de la unidad 3

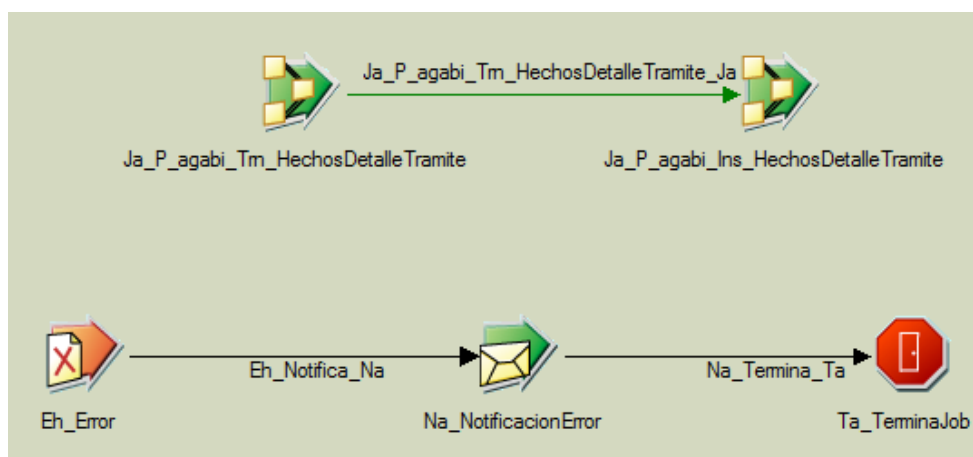


Figura 2.95 Secuencia de los hechos del trámite

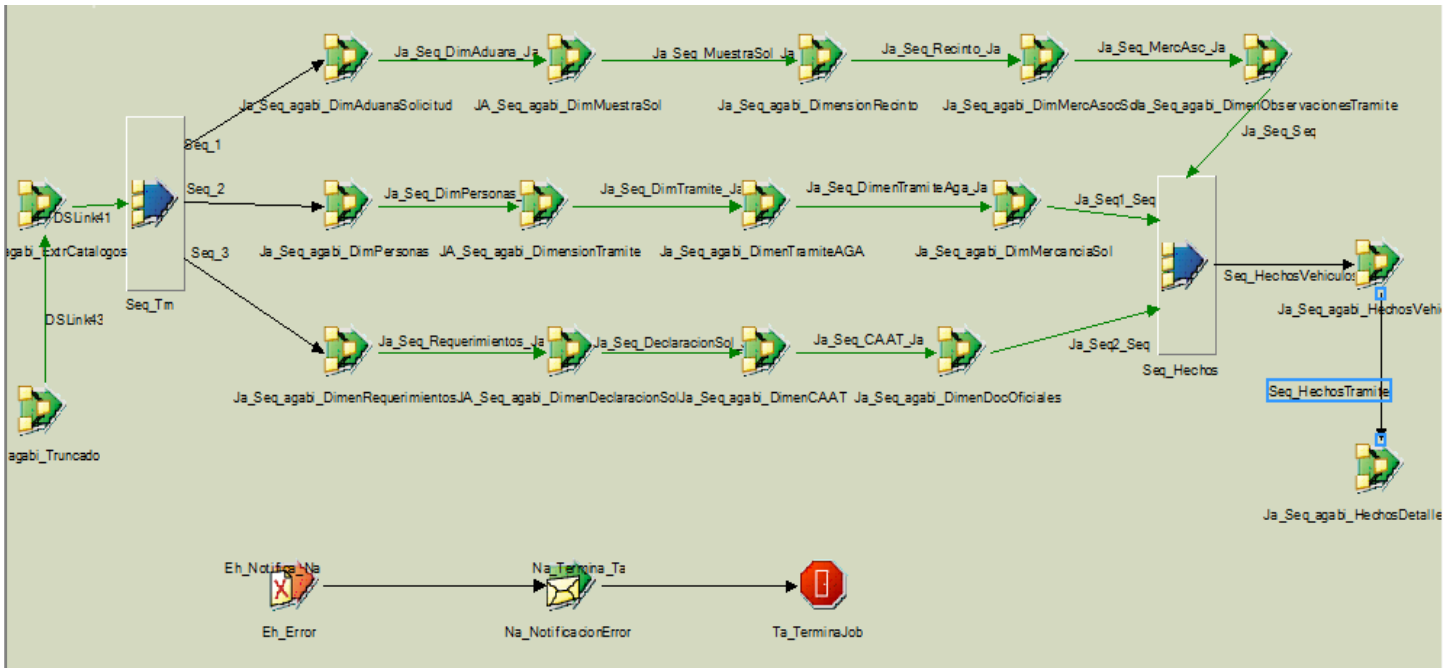


Figura 2.96 Secuencia maestra del datamart de la unidad 3

3.9 Testing

Para cada desarrollo realizado (con excepción de las tablas de hechos de los *datamarts* de cada unidad de negocio), se hace una consulta a la base de datos origen (base de datos de VUCEM) y la información y número de registros se compara con la información contenida y generada en el ETL primeramente cargada en el ODS, el cual es el destino filtrando la información con los trámites asociados a cada unidad de negocio.

Posteriormente en cada desarrollo de cada *datamart* de cada unidad de negocio, se hace la conciliación en número e información en los registros, solo que tomando como origen el ODS y como destino los *datamarts* correspondientes con los trámites asociados a cada unidad de negocio y previamente cargados con el ETL desarrollado.

En la figura 2.97 se pudo observar un claro ejemplo del tipo de prueba en cuanto a número entre el ODS y una de las dimensiones que se cargó para uno de los *datamarts*, y en este caso, existe un campo el cual no coincide en cantidad, por lo que se podría decir que existe un error en el desarrollo o en dado caso podría ser que faltó algún filtro al momento de hacer la validación.

The image displays two screenshots of a SQL query editor showing query results. The top screenshot shows a query with an inner join and a result table with 328 rows. The bottom screenshot shows a query with a filter and a result table with 318 rows.

Top Screenshot Query:

```

count (unique F_CGT_FFTEIRC1 ) F_CGT_FFTEIRC1 ,
count (unique F_CGT_FETESRC1 ) F_CGT_FETESRC1
from (select distinct
DD_DYC_TRAMITE1 . C_CGT_NFTUFRM1 ,
DD_DYC_TRAMITE1 . C_CGT_ISDOELN1 ,
DD_DYC_TRAMITE1 . F_DGT_FITENRC1 ,
DD_DYC_TRAMITE1 . F_CGT_FFTEIRC1 ,
DD_DYC_TRAMITE1 . F_CGT_FETESRC1
from DD_DYC_TRAMITE1
inner join DD_CGT_DTAERMT1 on DD_CGT_DTAERMT1.C_CGT_NFTUFRM1 = DD_DYC_TRAMITE1.C_CGT_NFTUFRM1); ---amecafe

```

Top Screenshot Result Table:

COUNT(*)	C_CGT_NFTUFRM1	C_CGT_ISDOELN1	F_DGT_FITENRC1	F_CGT_FFTEIRC1	F_CGT_FETESRC1
1	19835	19835	19835	17030	328

Bottom Screenshot Query:

```

Select Count (*),
count (unique C_CGT_NFTUFRM1 ) C_CGT_NFTUFRM1 ,
count (unique C_CGT_ISDOELN1 ) C_CGT_ISDOELN1 ,
count (unique F_CGT_FITENRC1 ) F_CGT_FITENRC1 ,
count (unique F_CGT_FFTEIRC1 ) F_CGT_FFTEIRC1 ,
count (unique F_CGT_FETESRC1 ) F_CGT_FETESRC1
from HE_CGT_DTERTAAL
where C_CGT_NFTUFRM1 like '110290%'
;

```

Bottom Screenshot Result Table:

COUNT(*)	C_CGT_NFTUFRM1	C_CGT_ISDOELN1	F_CGT_FITENRC1	F_CGT_FFTEIRC1	F_CGT_FETESRC1
1	19835	19835	19835	17030	318

2.97 Ejemplo de testing ODS vs Datamart

4.0 Documentación de los procesos

En los desarrollos realizados, se realizó la documentación correspondiente tanto de ODS como de *datamarts* de los cuales sólo adjunto algunos de la Unidad de negocio 1 del ODS en la sección de Anexos de éste documento.

Toda documentación incluye los primeros 5 puntos que a continuación explico. El punto 6 no es necesario, debido a que es probable que existan procesos que sean solo de tipo únicos.

- 1) Descripción – el cual es una breve síntesis de lo que trata el proceso
- 2) Entrada – indica el nombre y tipo de los insumos que se requieren para que el proceso sea satisfactorio, se incluye además información adicional relevante sobre cada una.
- 3) Salida – indica el nombre y tipo de lo que se obtiene al terminar el proceso, además se indica su regla de actualización.

- 4) Diagrama del proceso - es un diagrama de flujo operativo del proceso, indicado de manera general, donde se observan las entradas y salidas indicadas en los puntos anteriores.
- 5) Descripción detallada del proceso – como su nombre lo indica es la descripción detallada del proceso, en donde se especifica más a detalle del tipo de join que se hace, de donde se tomó la información para poder hacer el desarrollo, además de la salida que se obtenga.
- 6) Procesos similares – indica los procesos que se desarrollaron de manera similar al proceso especificado, se tiene el mismo número de entradas y salidas, su cambio radica en tener distintas entradas y salidas en cuanto a proceso, así mismo se incluye información adicional para su mejor entendimiento.

Un ejemplo es el proceso de extraer dirección de las personas solicitantes (anexo página 60) el cual se compone de los 6 puntos antes mencionados. Otro ejemplo el proceso extraer el catálogo tipo trámite(anexo página 61), el cual solo contiene los primeros 5 puntos. Ambos son correctos.

CAPITULO IV

Análisis de Resultados

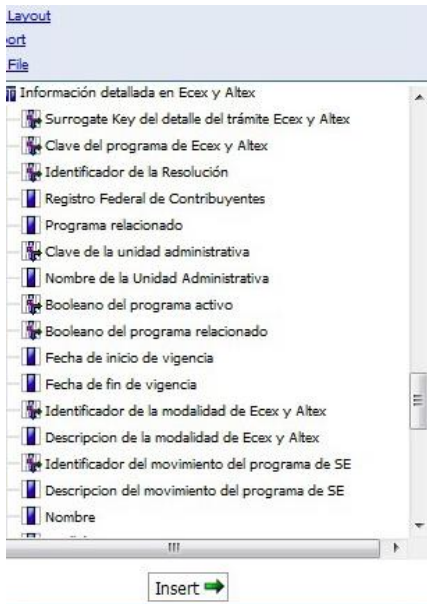
Con base en el desarrollo realizado por cada miembro del equipo, y al tener validaciones concluidas y satisfactorias después de cada desarrollo, se hicieron pruebas con el cliente. En un inicio se solicitaron más trámites que se tuvieron que agregar en cada dependencia solicitada. Posteriormente se hicieron otro tipo de pruebas de funcionalidad, de cantidad y calidad de la información, de diseño e implementación, inclusive se verificaba con el equipo de *testing* los números obtenidos, concluyendo con la aceptación del producto desarrollado y con los resultados obtenidos, de manera que el cliente puede acceder a Cognos para generar sus reportes o visualizar las consultas planeadas generadas.

La aportación que tuve en el proyecto consistió en el desarrollo de 3 unidades: Unidad 1, Unidad 2 y Unidad 3. Para cada una de ellas se realizó el análisis de requerimientos y a partir de ello comenzar el desarrollo de cada unidad.

La construcción de cada unidad contempla:

1. Crear extracciones hacia el ODS
2. Crear transformaciones para el ODS
3. Hacer la carga de cada unidad al ODS
4. Crear llave subrogada de cada dimensión
5. Crear transformaciones de cada dimensión
6. Hacer la carga de las dimensiones
7. Crear transformaciones de la tabla de hechos.
8. Hacer la carga de la tabla de hechos.
9. Hacer los reportes que se presentaban para cada usuario, ó generar una consulta con lo construido de cada unidad para obtener lo que el cliente necesitaba ver. (consultas no planeadas).

Un ejemplo es la figura 2.98 en la que se puede apreciar el resultado de una consulta no planeada de la Unidad 1, en la cual se puede apreciar diversos campos que se querían visualizar del datamart de la información detallada de la Unidad 1, fue implementado en Cognos Query Studio que es la herramienta que se utilizó para ver los reportes que se deseaban generar, y con la información estructurada en la forma que el usuario lo requiere.



Title				
Número del Folio del Trámite	Identificador de la Solicitud	Identificador del Tipo de Trámite y su Modalidad	Descripcion de la modalidad de Eceex y Altex	Nombre de la Unidad Administrativa
0209900200119992033000030	14420830	9902	ALTEX Exportador Directo	OAXACA
0209900200119992644000128	14420930	9902	ALTEX Exportador Indirecto	CIUDAD OBREGON
0209900200119990509000189	14420991	9902	ALTEX Exportador Directo	SALTILLO
0209900200119992848000229	14421031	9902	ALTEX Exportador Directo	MATAMOROS
0209900200119991729000275	14421077	9902	ALTEX Exportador Directo	CUERNAVACA
020990020011999118000277	14421079	9902	ALTEX Exportador Directo	CELAYA
020990020011999118000318	14421120	9902	ALTEX Exportador Directo	CELAYA
0209900200119999901000373	14421175	9902	ALTEX Exportador Directo	METROPOLITANA
0209900200119991628000421	14421223	9902	ALTEX Exportador Directo	MORELIA
0209900200119992235000452	14421254	9902	ALTEX Exportador Directo	QUERETARO
0209900200119991017000496	14421298	9902	ALTEX Exportador Directo	GOMEZ PALACIO
0209900200119991017000499	14421301	9902	ALTEX Exportador Directo	GOMEZ PALACIO
0209900200119991424000500	14421302	9902	ALTEX Exportador Directo	GUADALAJARA
0209900200119993159000537	14421339	9902	ALTEX Exportador Directo	MERIDA
020990020011999254000632	14421434	9902	ALTEX Exportador Directo	CULIACAN
0209900200119992134000633	14421435	9902	ALTEX Exportador Directo	PUEBLA
0209900200119992134000637	14421439	9902	ALTEX Exportador Directo	PUEBLA
0209900200119990204000695	14421497	9902	ALTEX Exportador Directo	TIJUANA
0209900200119991931000797	14421599	9902	ALTEX Exportador Directo	MONTERREY

Figura 2.98 Ejemplo de reporte de consulta no planeada

Todo lo presentado en el capítulo III fue parte de los desarrollos que me tocó realizar. Para el diseño previo a la implementación, el proyecto ya contaba con una metodología, la cual se tenía que respetar. Un ejemplo son las extracciones y las cargas de datos ya que prácticamente se hacía de la misma forma para cada proceso, lo que cambiaba es:

1. Lo que se deseaba extraer(no eran los mismos trámites asociados a cada unidad ni los mismos datos requeridos)
2. Para la carga de catálogos no se tenía un *dataset* de rechazo, para cualquier otro proceso era necesario.
3. El asociar los trámites de cada unidad para tener una única fuente de información "inicial".

Las transformaciones de cada proceso podían ser de distintas formas en algunas se generan columnas, en otras se toman valores máximos, en otras solamente se filtra, en algunas se hacen asociaciones de tablas, ó se ordenan los datos entre otras transformaciones existentes, a pesar de la gran diversidad de transformaciones existentes, se podía dar el caso de tener procesos similares, por ejemplo la transformación de catálogos.

Las ventajas que pude encontrar al diseñar e implementar con la herramienta de ETL, es que permite interactuar y desarrollar de manera más simple y gráfica, lo que ayuda a simplificar la programación y automatización de los procesos. Otra ventaja es la compatibilidad con diferentes bases de datos como por ejemplo Oracle, DB2, Informix, Teradata etc. Al ser una herramienta permite compilar cada desarrollo, lo que facilita la detección de errores sintácticos. Además permite la parametrización de datos, lo que te ayuda si en algún momento deseas cambiar algún dato para realizar una diferente ejecución.

Una de las desventajas es que es necesario conocer gran parte de la herramienta, principalmente por cuestiones de tiempo de procesamiento, a veces utilizar una herramienta puede alentar el

procesamiento de los datos por la cantidad de transformaciones que se tiene o simplemente porque el desarrollo que se implementó no es el óptimo.

Otra desventaja es el costo de la licencia, sin embargo, a mi parecer es fundamental saber lo básico de programación en SQL, ya que si no se tienen claro esos conceptos, es muy probable que el manejo y manipulación de la herramienta se dificulte.

Poder hacer BI fue una experiencia muy grata pero creo que debería de haber mejores herramientas open source, ya que la mayoría de las herramientas utilizadas en la vida real son herramientas comerciales, lo cual en cierta parte te llega a limitar para poder seguir desarrollando y experimentando, en mi caso, fuera del ámbito laboral.

CONCLUSIONES

La inteligencia de negocios está pensado para agilizar y sistematizar uno o varios procesos de negocio de cada empresa o negocio.

Para poder llevarlo a cabo se necesita que las personas involucradas en plantear este tipo de soluciones se 'infoxiquen' del proceso que se lleva o llevará a cabo, de manera que al plantear la solución, sea la óptima y sencilla en cuanto a uso y solución que se le dará al cliente.

Una solución es la óptima cuando el proceso que se genera toma la menor cantidad de tiempo al menor costo pero sin afectar los resultados que el cliente solicita.

El análisis y diseño de la solución es la parte más importante del proyecto, ya que si hay un mal análisis, es muy probable que no se llegue a completar la construcción y por lo mismo la ejecución sería inexistente.

Una vez teniendo el análisis correcto y de acuerdo a lo que el cliente solicita se podrá proceder a la construcción, sin embargo es importante tomar en cuenta que cada cambio que llegue a haber, afectará considerablemente los tiempos de desarrollo y mientras más cambios haya mayor retraso se tendrá en la entrega del proyecto.

Teniendo en cuenta lo anterior, para el desarrollo del proyecto de BI de VUCEM fue necesario un rediseño de BI, todo ello se logró en un tiempo aproximado de un año y con aproximadamente 20 personas involucradas en ello.

Con esto se logró que, para el negocio que contemplaba 18 unidades de negocio, se unificaran en 1 solo cliente, sin perder de vista el desarrollo específico de cada unidad. Cada desarrollo se realizó con base a estándares y lineamientos contemplados para el proyecto, con el menor procesamiento posible.

Primeramente se creó y verificó en cuanto a número y calidad de la información, la extracción, transformación y carga de datos de la base de datos transaccional hacia el ODS. Una vez aprobado, se hizo el mismo proceso anterior pero teniendo como insumo el ODS y como destino los *datamarts* correspondientes de cada unidad de negocio.

Al ser aprobado cada dimensión del *datamart* se crean las tablas de hechos, las cuales son las que nos brindan la información relevante para el usuario.

Finalmente se procede a la presentación de los datos al cliente en donde se tienen consultas planeadas de las cuales se obtiene información 'automatizada' que el usuario quiere saber, siendo la consulta muy específica de acuerdo al negocio, y también se tienen consultas no planeadas, las cuales son las que el cliente va a poder generar vía web, en donde obtendrá la información que el desee en el momento que así lo solicite, evitando así la generación de tickets y teniendo completo acceso a su

respectiva unidad de negocio ó en dado caso de ser el cliente principal, sólo él tendría acceso a consultar la información de todas las unidades de negocio.

En lo personal me pareció un proyecto bien organizado, con objetivos y tiempos cumplidos, además de tener pocos defectos en lo desarrollado, el cual fue reflejo de un buen líder y un grandioso equipo de trabajo.

Por lo anterior, tuve un gran y amplio aprendizaje desde lo personal y profesional aprendiendo a convivir, complementar y trabajar en equipo en un ambiente real, además de ser más responsable debido a que una vez terminado el desarrollo, se pasaba a producción en el cual se trataba de evitar todo tipo de error en la ejecución. Aprendí mayormente el lenguaje de las bases de datos (SQL), analizando cada consulta que realizaba así como el resultado de la misma y teniendo en cuenta qué es lo que se quería obtener.

Además aprendí a utilizar la herramienta para el desarrollo de ETL, permitiéndonos desarrollar los procesos de manera más gráfica lo que nos ayudaba a tener la menor cantidad de errores pero sin perder de vista que el proceso se podía realizar de manera manual y sin dependencia total de la herramienta.

Por último quiero agregar que se garantiza el cumplimiento de los requerimientos establecidos, teniendo así un proyecto con un desarrollo exitoso, contemplando los trámites de cada unidad de negocio.

ANEXOS

PROCESO – [EXTRAER DIRECCIÓN DE LAS PERSONAS SOLICITANTES]

Descripción	Extraer la información relacionada a las direcciones de las personas solicitantes de Unidad 1.
--------------------	--

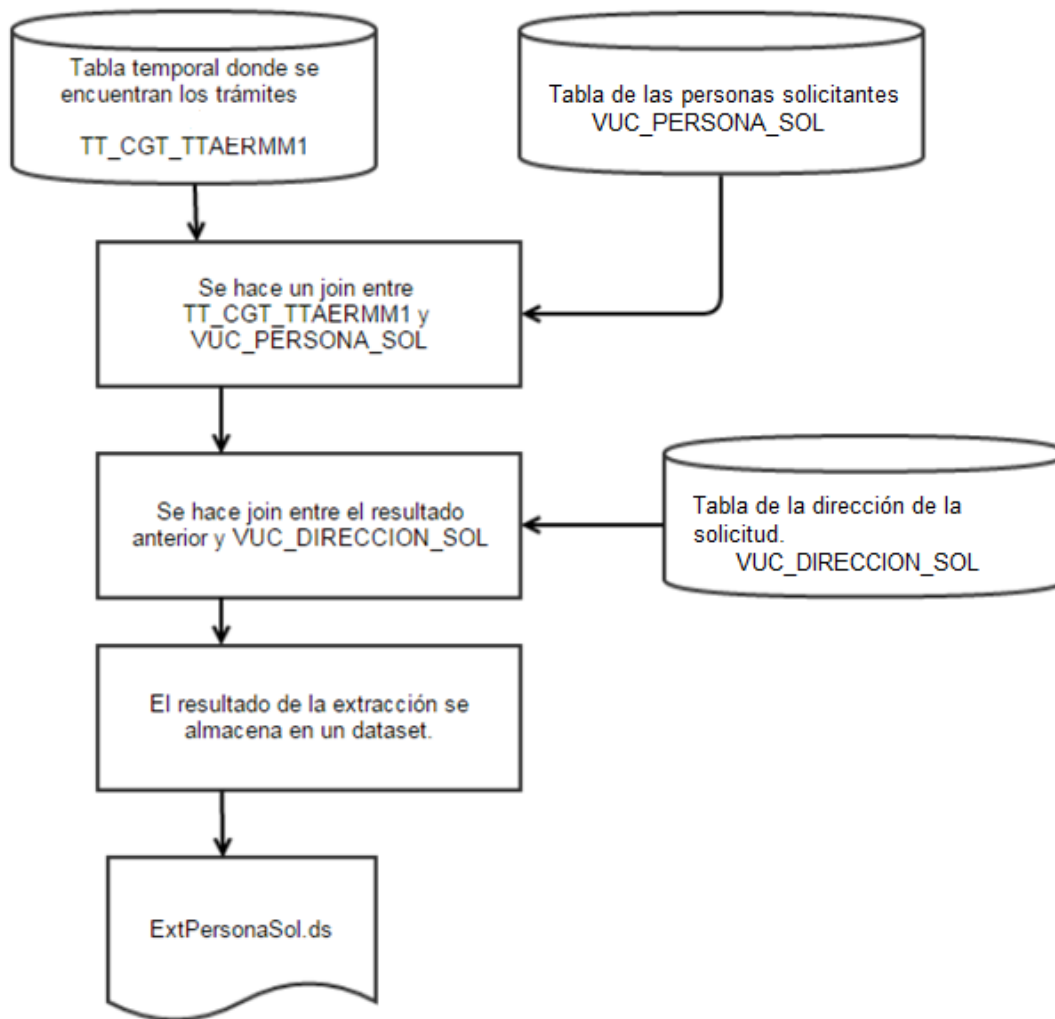
ENTRADA

Tipo	Nombre	Información Adicional
Tabla temporal	TT_CGT_TTAERMM1	Tabla temporal de VUCEM de los trámites correspondientes a los detalles de Unidad 1.
Tabla	VUC_PERSONA_SOL	Tabla de VUCEM que contiene los datos de las personas solicitantes de Unidad 1.
Tabla	VUC_DIRECCION_SOL	Tabla de VUCEM que contiene los datos de la dirección de la solicitud de Unidad 1.

SALIDA

Tipo	Nombre	Regla de Actualización
Archivo	Dataset: ExtPersonaSol.ds archivo donde se almacena la información de la extracción de persona solicitante.	Reescribir archivo

DIAGRAMA DEL PROCESO



DESCRIPCIÓN DETALLADA DEL PROCESO

1. Extraer los campos de la tabla personas solicitantes “VUC_PERSONA_SOL” los cuales se especifican en el documento, **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**. Se realiza un inner join entre la tabla temporal TT_CGT_TTAERMM1 y la tabla de “VUC_PERSONA_SOL” por medio del identificador de la solicitud “ID_SOLICITUD”.
2. Al resultado de ésta extracción, se realiza un left join con la tabla de la dirección de la solicitud “VUC_DIRECCION_SOL” por medio del identificador de la dirección de la solicitud “ID_DIRECCION_SOL”.
3. Por último, el resultado de la extracción se almacena en un dataset con el nombre “ExtPersonaSol.ds” en la ruta donde se guardan todos los archivos de extracción para Unidad 1.

PROCESOS SIMILARES

ID	Nombre del Proceso	Entrada	Salida	Información Adicional
1	Extraer detalle de Unidad 1	Tabla1: TT_CGT_TTTAEA E1 Tabla 2 : VUC_RESOLUCIO N Tabla 3: VUC_ECEX_ALTE X	Dataset: vucAltex.ds	1.- Hacer un inner join entre la tabla temporal “TT_CGT_TTTAEA1” y “VUC_RESOLUCION” por medio del número de folio del trámite “NUM_FOLO_TRAMITE”. 2.- Al resultado anterior, hacer un inner join con la tabla “VUC_ECEX_ALTEX” por medio del identificador de la resolución “ID_RESOLUCION”. 3.- El resultado de la extracción se almacena en un dataset con el nombre “vucAltex.ds”

PROCESO – [EXTRAER EL CATÁLOGO TIPO TRÁMITE]

Descripción	Extraer la información del catálogo de tipos de trámite relacionados a los trámites de Unidad 1.
--------------------	--

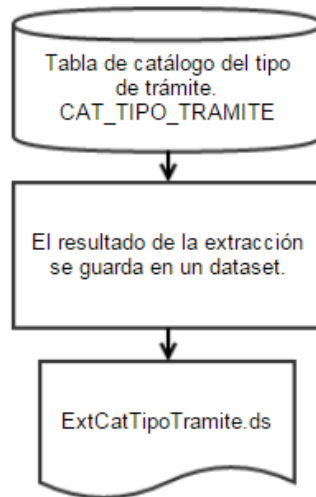
ENTRADA

Tipo	Nombre	Información Adicional
Tabla	CAT_TIPO_TRAMITE	Tabla de catálogo de los tipos de trámite

SALIDA

Tipo	Nombre	Regla de Actualización
Archivo	Dataset: ExtCatTipoTramite.ds, archivo donde se guarda la información de la extracción del catálogo de tipos de trámite.	Reescribir archivo

DIAGRAMA DEL PROCESO



DESCRIPCIÓN DETALLADA DEL PROCESO

1. Extraer todos los campos de la tabla de catálogo del tipo de trámite “CAT_TIPO_TRAMITE” el resultado de la extracción se almacena en un dataset con el nombre “ExtCatTipoTramite.ds” en la ruta donde se guardan todos los archivos de extracción para Unidad 1.

PROCESO – [EXTRAER LA DIRECCIÓN DE LA SOLICITUD]

Descripción	Extraer las direcciones de la solicitud asociadas a los trámites de Unidad 1.
--------------------	---

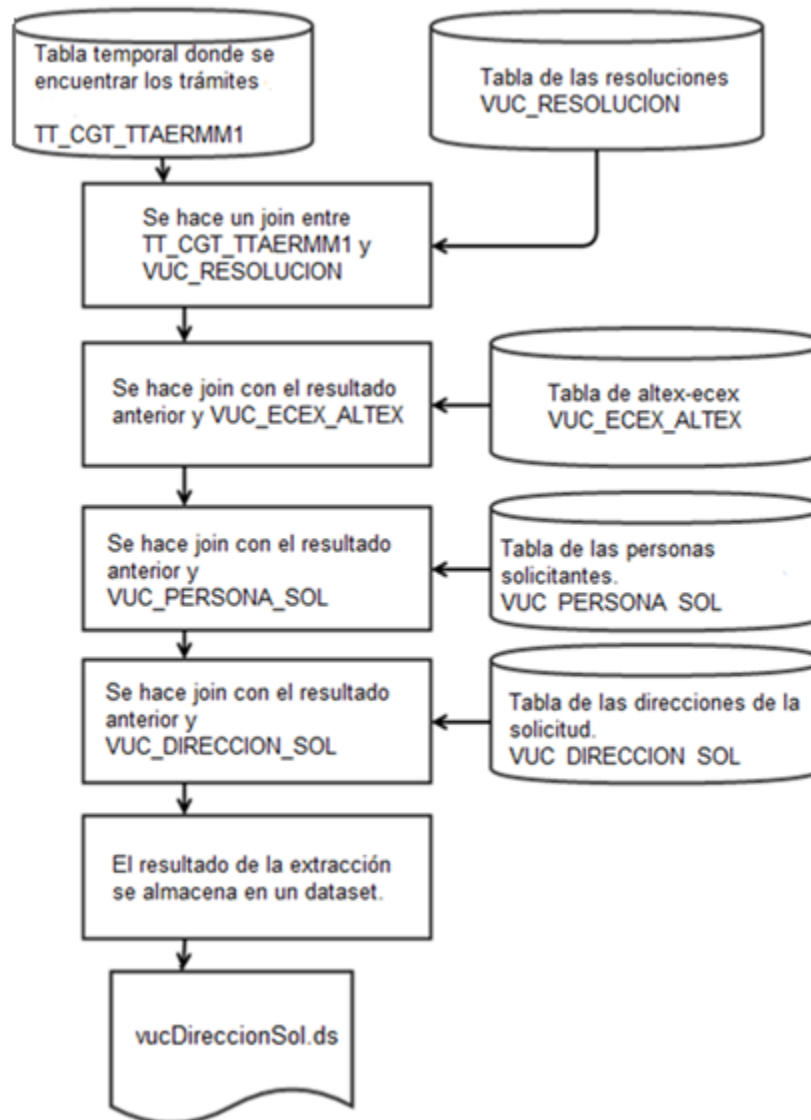
ENTRADA

Tipo	Nombre	Información Adicional
Tabla	TT_CGT_TTPAER1	Tabla temporal que contiene los trámites de Unidad 1.
Tabla	VUC_RESOLUCION	Tabla de las resoluciones de VUCEM de Unidad 1.
Tabla	VUC_ECEX_ALTEX	Tabla de detalle de Unidad 1.
Tabla	VUC_PERSONA_SOL	Tabla de las personas solicitantes de Unidad 1.
Tabla	VUC_DIRECCION_SOL	Tabla de las direcciones de la solicitud de VUCEM de Unidad 1.

SALIDA

Tipo	Nombre	Regla de Actualización
Archivo	vucDireccionSol.ds	Reescribir el archivo.

DIAGRAMA DEL PROCESO



DESCRIPCIÓN DETALLADA DEL PROCESO

1. Realizar un inner join entre la tabla temporal “TT_CGT_TTAERMM1” que contiene los trámites de Unidad 1, cuyo estado se encuentra en “Resuelto” o “Concluido”, y la tabla que contiene las resoluciones de VUCEM, “VUC_RESOLUCION”, por el número de folio del trámite “NUM_FOLIO_TRAMITE”, extrayendo los campos que se especifican en el documento **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**.
2. Al resultado de extracción se realiza un inner join con la tabla de VUCEM de Unidad 1 “VUC_ECEX_ALTEX”, por el identificador de la resolución “ID_RESOLUCION”, extraer de cada tabla los campos especificados en el documento **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**.
3. Con el resultado de la extracción anterior, se realiza un inner join con la tabla de personas solicitantes “VUC_PERSONA_SOL” por medio del identificador de la solicitud “ID_SOLICITUD”

extrayendo de cada tabla los campos especificados en el documento **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**.

4. A partir del resultado anterior se hace un inner join con la tabla de dirección de la solicitud "VUC_DIRECCION_SOL" por medio del identificador de la dirección de la solicitud "ID_DIRECCION_SOL"
5. Por último, el resultado de la extracción se guardará en el dataset "vucDireccionSol.ds", localizado en la ruta de extracción de archivos para Unidad 1.

PROCESO – [EXTRAER FIRMA DE LA SOLICITUD]

Descripción	Extraer la información de la firma de la solicitud relacionados a los trámites de Unidad 1.
--------------------	---

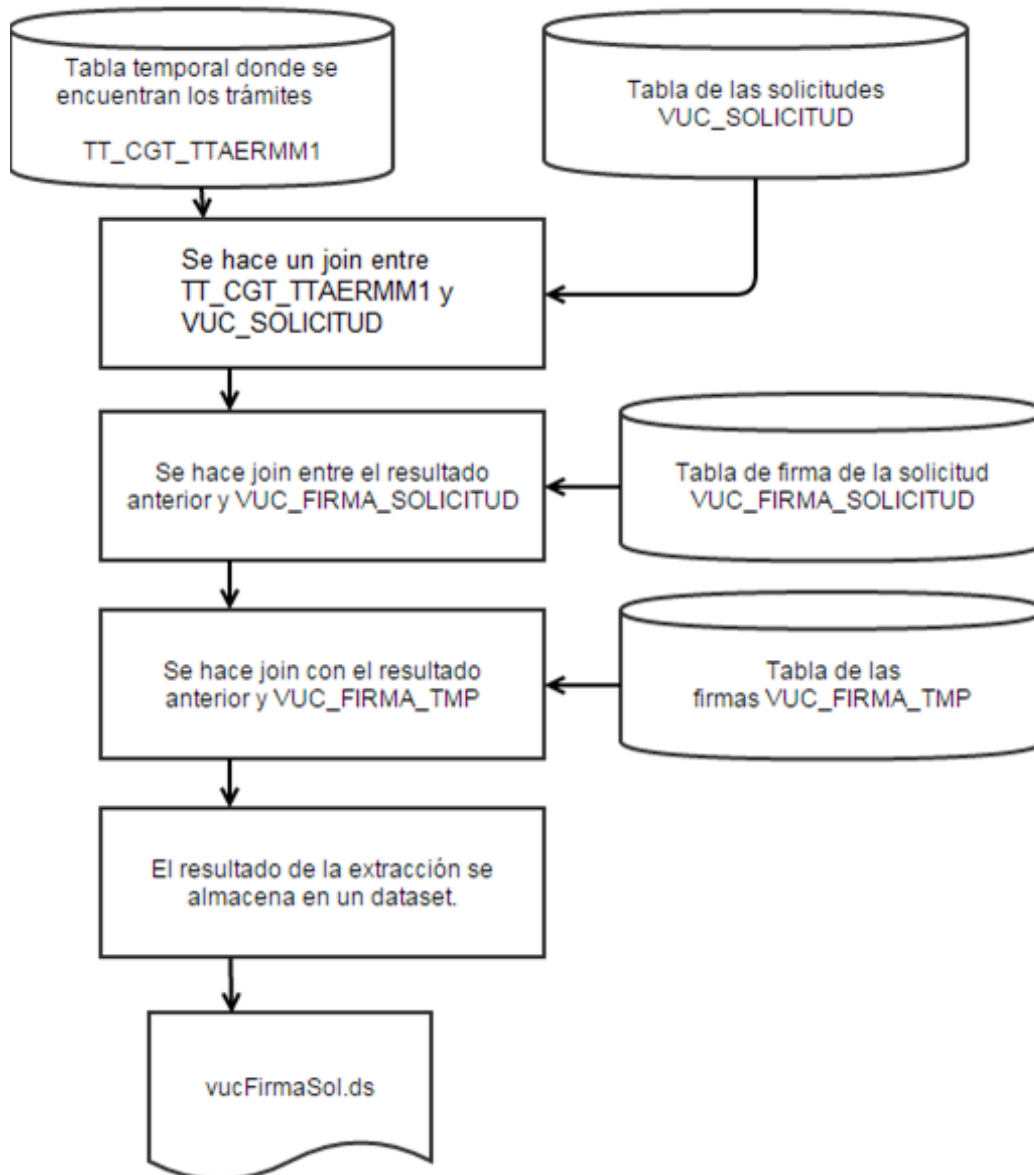
ENTRADA

Tipo	Nombre	Información Adicional
Tabla temporal	TT_CGT_TTTAEAE1	Tabla temporal de los tramites asociados a Unidad 1
Tabla	VUC_SOLICITUD	Tabla de la solicitudes de VUCEM
Tabla	VUC_FIRMA_SOLICITUD	Tabla de las firmas de la solicitudes de VUCEM
Tabla	VUC_FIRMA_TMP	Tabla de las firmas de VUCEM

SALIDA

Tipo	Nombre	Regla de Actualización
Archivo	Dataset: vucFirmaSol.ds, archivo donde se guarda la información de la extracción.	Reescribir archivo

DIAGRAMA DEL PROCESO



DESCRIPCIÓN DETALLADA DEL PROCESO

1. Realizar un inner join entre la tabla temporal “TT_CGT_TTAERMM1” que contiene los trámites de Unidad 1, cuyo estado se encuentra en “Resuelto” o “Concluido”, y la tabla que contiene las solicitudes de VUCEM, “VUC_SOLICITUD”, por el identificador de la solicitud “ID_SOLICITUD”, extrayendo los campos que se especifican en el documento **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**.
2. Al resultado de extracción se realiza un inner join con la tabla de firma de la solicitudes “VUC_FIRMA_SOLICITUD” mediante el identificador de la solicitud “ID_SOLICITUD”, extraer de cada tabla los campos especificados en el documento **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**.
3. Con el resultado de la extracción anterior, se realiza un left join con la tabla de firma “VUC_FIRMA_TMP” por medio del identificador de la firma “ID_FIRMA” extrayendo de cada

tabla los campos especificados en el documento **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**.

- Por último, el resultado de la extracción se guarda en el dataset “vucFirmaSol.ds”, localizado en la ruta de extracción de archivos para Unidad 1.

PROCESOS SIMILARES

ID	Nombre del Proceso	Entrada	Salida	Información Adicional
1	Extraer resolución de la solicitud	Tabla1: TT_CGT_TTTAEAE1 Tabla 2 : VUC_RESOLUCION Tabla 3: VUC_FIRMA_RESOLUCION Tabla 4: VUC_FIRMA_TMP	Dataset: vucResolucion.ds	1.- Hacer un inner join entre la tabla temporal “TT_CGT_TTTAEAE1” y “VUC_RESOLUCION” por medio del número de folio del trámite “NUM_FOLO_TRAMITE”. 2.- Al resultado anterior, hacer un left join con la tabla VUC_FIRMA_RESOLUCION por medio del identificador de la resolución “ID_RESOLUCION”. 3.-Hacer un left join con la tabla VUC_FIRMA_TMP por medio del identificador de la firma “ID_FIRMA”. 4.- El resultado de la extracción se guarda en el dataset “vucResolucion.ds”.

PROCESO – [EXTRAER LAS PERSONAS SOLICITANTES]

Descripción	Extraer las personas solicitantes asociadas a los trámites de Unidad 1.
--------------------	---

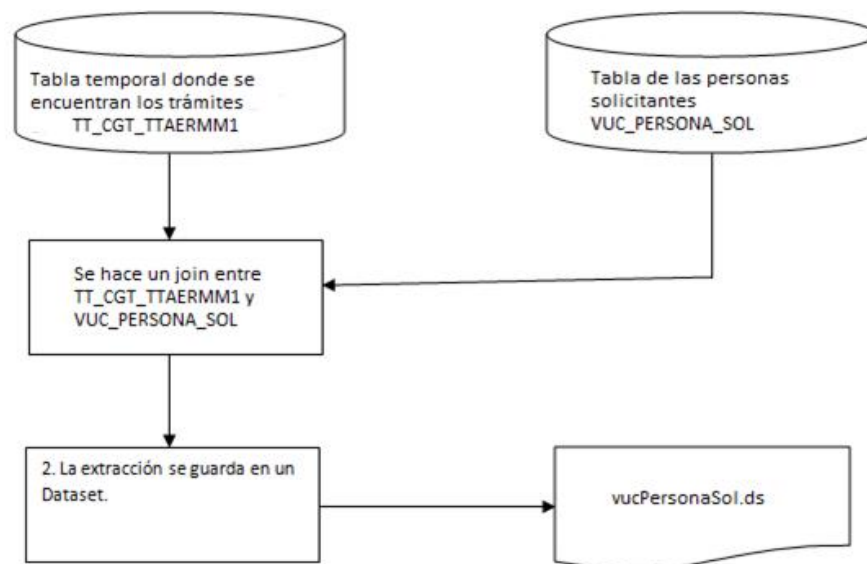
ENTRADA

Tipo	Nombre	Información Adicional
Tabla	TT_CGT_TTPAER1	Tabla temporal que contiene los trámites de Unidad 1.
Tabla	VUC_PERSONA_SOL	Tabla de las personas solicitantes de Unidad 1.

SALIDA

Tipo	Nombre	Regla de Actualización
Archivo	vucPersonaSol.ds	Reescribir el archivo.

DIAGRAMA DEL PROCESO



DESCRIPCIÓN DETALLADA DEL PROCESO

1.- Realizar un inner join entre la tabla temporal “TT_CGT_TTAERMM1” que contiene los trámites de Unidad 1, cuyo estado se encuentra en “Resuelto” o “Concluido”, y la tabla que contiene las personas solicitantes, “VUC_PERSONA_SOL”, por el identificador de la solicitud “ID_SOLICITUD”, extrayendo los campos que se especifican en el documento **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**.

2.- El resultado de la extracción se guardará en el dataset “vucPersonaSol.ds”, localizado en la ruta de extracción de archivos para Unidad 1.

PROCESO – [TRANSFORMAR DETALLE DE UNIDAD 1]

Descripción	Transformar los detalles de Unidad 1 relacionadas a los trámites de Unidad 1.
--------------------	---

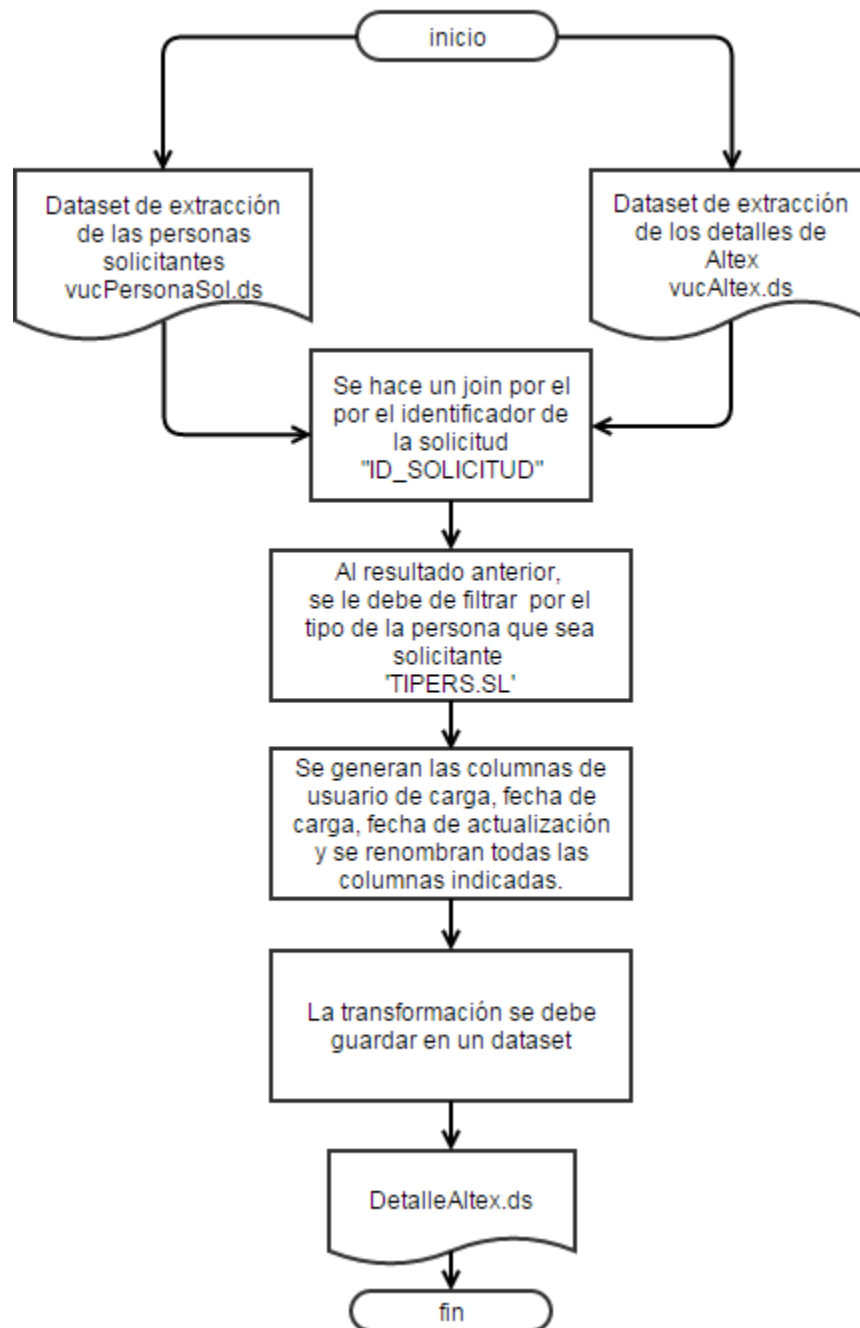
ENTRADA

Tipo	Nombre	Información Adicional
Archivo	Dataset: vucPersonaSol.ds	Dataset que contiene la información extraída de las personas solicitantes de Unidad 1.
Archivo	Dataset: vucAltex.ds	Dataset que contiene la información extraída de los detalles de Unidad 1.

SALIDA

Tipo	Nombre	Regla de Actualización
Archivo	Dataset: DetalleAltex.ds archivo donde almacena la información de la transformación lista para la carga.	Reescribir archivo

DIAGRAMA DEL PROCESO



DESCRIPCIÓN DETALLADA DEL PROCESO

1. A partir de los datasets que se obtuvieron en la extracción, "vucPersonaSol.ds" y "vucAltex.ds" que contienen los campos que se especifican en el documento, **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**, se realiza un left join por el campo identificador de la solicitud "ID_SOLICITUD".
2. Al resultado anterior, se le debe filtrar por el tipo de persona que sea solicitante "TIPERS.SL".

3. A continuación, se le agregan las columnas de usuario de carga, fecha de carga y fecha de actualización y se renombran las columnas indicadas y especificadas en el documento, **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**.
4. Por último, el resultado de la transformación se guardará en el dataset "DetalleAltex.ds", localizado en la ruta de carga de archivos para Unidad 1.

PROCESO – [TRANSFORMAR DETALLE DE LA DIRECCIONES DE LA SOLICITUD]

Descripción	Transformar las direcciones de la solicitud de Unidad 1
--------------------	---

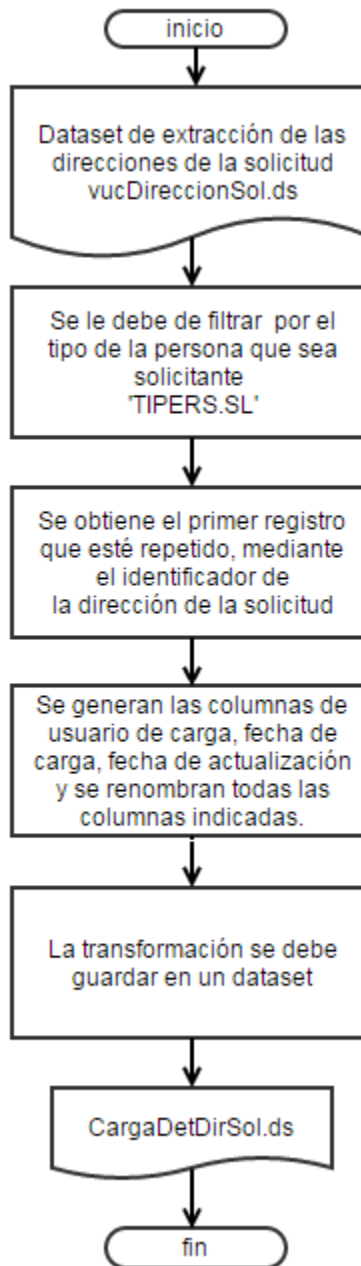
ENTRADA

Tipo	Nombre	Información Adicional
Archivo	Dataset: vucDireccionSol.ds	Dataset que contiene la información extraída de las direcciones de la solicitud de Unidad 1.

SALIDA

Tipo	Nombre	Regla de Actualización
Archivo	Dataset: CargaDetDirSol.ds donde almacena la información de la transformación lista para la carga.	Reescribir el archivo.

DIAGRAMA DEL PROCESO



DESCRIPCIÓN DETALLADA DEL PROCESO

1. A partir del dataset que se obtuvo en la extracción, "vucDireccionSol.ds" que contiene los campos que se especifican en el documento, **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**, se le debe filtrar por el tipo de persona que sea solicitante "TIPERS.SL".
2. A continuación, se obtiene el primer registro que esté repetido, mediante el identificador de la dirección de la solicitud.

3. Al resultado anterior, se le agregan las columnas de usuario de carga, fecha de carga y fecha de actualización y se renombran las columnas indicadas y especificadas en el documento, **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**.
4. Por último, el resultado de la transformación se guardará en el dataset “CargaDetDirSol.ds”, localizado en la ruta de carga de archivos para Unidad 1.

PROCESO – [TRANSFORMAR FIRMAS DE LA SOLICITUD]

Descripción	Transformar firmas de la solicitud de Unidad 1
--------------------	--

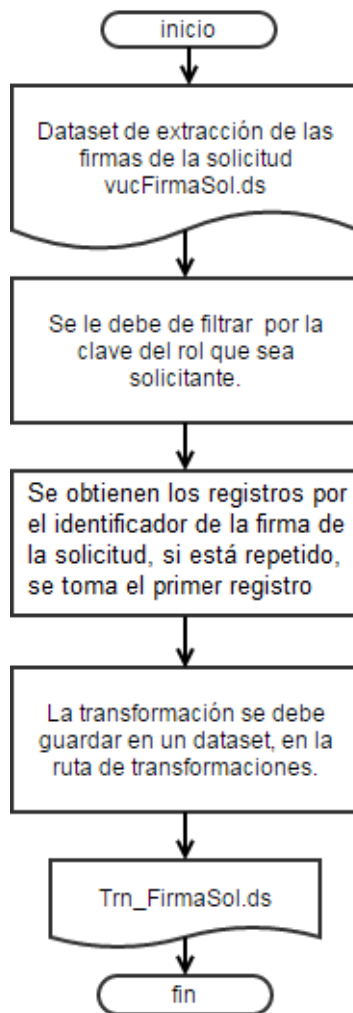
ENTRADA

Tipo	Nombre	Información Adicional
Archivo	Dataset: vucFirmaSol.ds	Dataset que contiene la información extraída de las firmas de la solicitud de Unidad 1.

SALIDA

Tipo	Nombre	Regla de Actualización
Archivo	Dataset: Trn_FirmaSol.ds donde almacena la información de la transformación	Reescribir el archivo.

DIAGRAMA DEL PROCESO



DESCRIPCIÓN DETALLADA DEL PROCESO

1. A partir del dataset que se obtuvo en la extracción, "vucFirmaSol.ds" que contiene los campos que se especifican en el documento, **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**, se le debe filtrar por la clave del rol, de tal manera que sea solicitante.
2. A continuación, se obtienen los registros por el identificador de la firma de la solicitud, si está repetido, se toma el primer registro.
3. Por último, el resultado de la transformación se guardará en el dataset "Trn_FirmaSol.ds", localizado en la ruta de transformaciones de archivos para Unidad 1, debido a que es un insumo para la transformación del Detalle de Trámites de Altex.

PROCESO – [TRANSFORMAR EL DETALLE DE LOS TRÁMITES DE UNIDAD 1]

Descripción	Transformar el detalle de los trámites de Unidad 1
--------------------	--

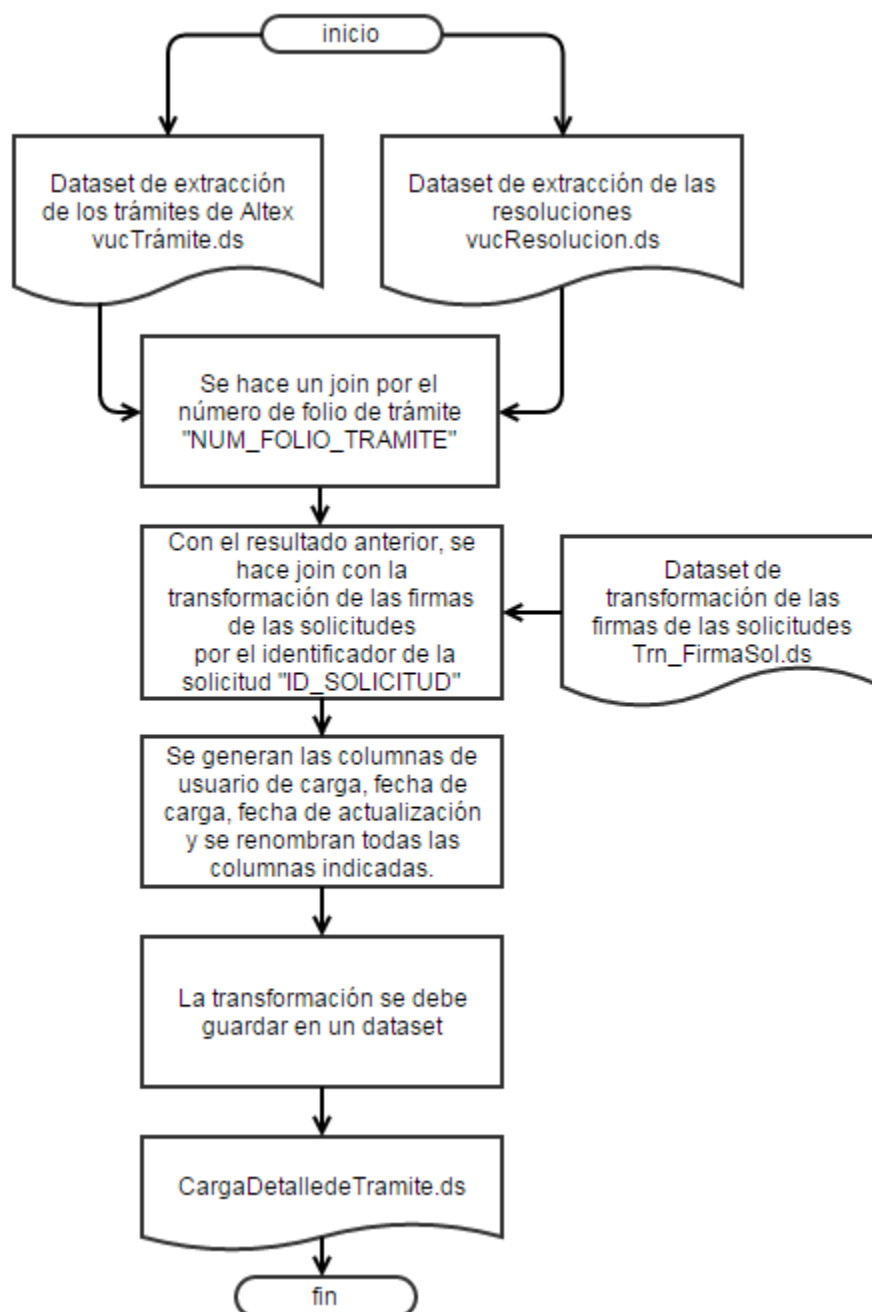
ENTRADA

Tipo	Nombre	Información Adicional
Archivo	Dataset: Trn_FirmaSol.ds	Dataset que contiene la información transformada de las firmas de la solicitud de Unidad 1.
Archivo	Dataset: vucResolucion.ds	Dataset que contiene la información extraída de las resoluciones de Unidad 1.
Archivo	Dataset: vucTramite.ds	Dataset que contiene la información extraída de los trámites de Unidad 1.

SALIDA

Tipo	Nombre	Regla de Actualización
Archivo	Dataset: CargaDetalledeTramite.ds donde almacena la información de la transformación lista para la carga.	Reescribir el archivo.

DIAGRAMA DEL PROCESO



DESCRIPCIÓN DETALLADA DEL PROCESO

1. A partir de los datasets que se obtuvieron en la extracción, "vucTrámite.ds" y "vucResolucion.ds", así como del dataset "Trn_FirmaSol.ds" que se obtuvo en un proceso de transformación y los cuales contienen los campos que se especifican en el documento **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**, se hace un left join entre los trámites de Altex y las resoluciones, por el número de folio de trámite "NUM_FOLIO_TRAMITE".

2. Al resultado anterior, se le agrega la información de las firmas de las solicitudes mediante un left join por el identificador de la solicitud "ID_SOLICITUD".
3. A continuación, se le agregan las columnas de usuario de carga, fecha de carga y fecha de actualización y se renombran las columnas indicadas y especificadas en el documento, **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**.
4. Por último, el resultado de la transformación se guardará en el dataset "CargaDetalledeTramite.ds", localizado en la ruta de carga de archivos para Unidad 1.

PROCESO – [CARGAR DETALLE DE UNIDAD 1]

Descripción	Cargar la información transformada relacionada a los detalles de Unidad 1.
--------------------	--

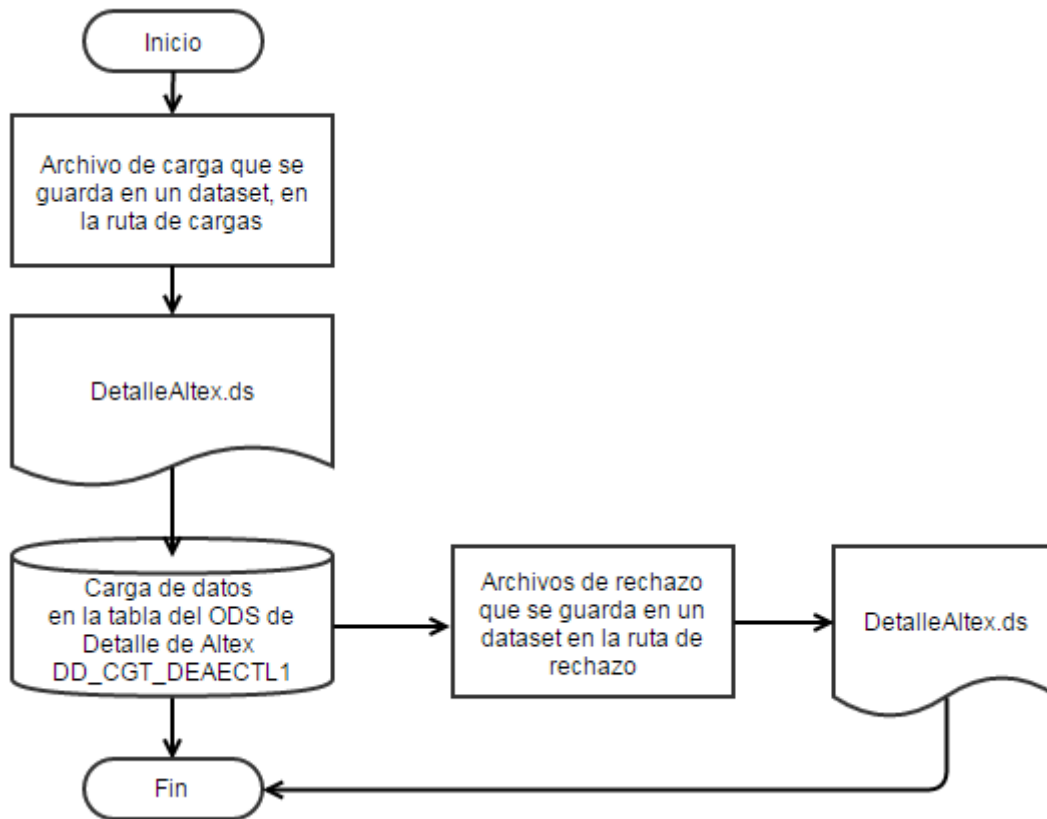
ENTRADA

Tipo	Nombre	Información Adicional
Archivo	Dataset: DetalleUnidad 1.ds	Dataset que conforma los detalles de Unidad 1 transformados.

SALIDA

Tipo	Nombre	Regla de Actualización
Tabla	DD_CGT_DEAECLT1	Reescribir datos
Archivo	Dataset: DetalleUnidad 1.ds archivo donde se almacena la información de rechazo de los detalles de Unidad 1.	Reescribir archivo

DIAGRAMA DEL PROCESO



DESCRIPCIÓN DETALLADA DEL PROCESO

1. Cargar los datos del archivo de carga que se guarda en el dataset “DetalleUnidad 1.ds”, en la ruta de cargas en tabla correspondiente al ODS de los detalles de Unidad 1 que se especifica en el documento, **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**.
2. En caso de que exista algún registro que no cumpla con los detalles de Unidad 1 se guardará en el archivo de rechazo que se guarda en un dataset “DetalleUnidad 1.ds” en la ruta de rechazo.
3. Por último, la carga se hará en la tabla DD_CGT_DEAECTL1 que corresponde a la tabla del ODS de los detalles de Unidad 1 que se especifica en el documento, **VU.Id052.02-MOD-Consultas Trámites Gubernamentales**.

PROCESOS SIMILARES

ID	Nombre del Proceso	Entrada	Salida	Información Adicional
1	Carga detalle dirección de la solicitud	Dataset: CargaDetDirSol.ds	Tabla: DD_CGT_DD SEIOT1	Cargar los datos del archivo de carga que se guarda en el dataset "CargaDetDirSol.ds".
			Dataset: DetalleRechDir Sol.ds	El archivo de rechazo que se guarda en un dataset para el detalle de la dirección de la solicitud es "DetalleRechDirSol.ds". La carga se hará en la tabla DD_CGT_DDSEIOT1 que corresponde a la tabla del ODS de los detalles dirección de la solicitud de Unidad 1 que se especifica en el documento, VU.Id052.02-MOD-Consultas Trámites Gubernamentales.
2	Carga detalle del trámite de Unidad 1	Dataset: CargaDetalledeTramite.ds	Tabla: DD_CGT_DTE AERC1	Cargar los datos del archivo de carga que se guarda en el dataset "CargaDetalledeTramite.ds". El archivo de rechazo que se guarda en un dataset para el detalle del trámite de Unidad 1 es "Rechazado_DetalleTramUnidad 1.ds". La carga se hará en la tabla DD_CGT_DTEAERC1 que corresponde a la tabla del ODS de los detalles del trámite de Unidad 1 que se especifica en el documento, VU.Id052.02-MOD-Consultas Trámites Gubernamentales.

BIBLIOGRAFÍA

- Julius Lern, Paul Christensen. (2010). IBM Infosphere datastage parallel framework standard practices. EUA: IBM RedBooks.
- Salvador Ramos. (2011). Microsoft business intelligence: veo el cubo medio lleno. España: Solid Q Press.
- Josep Lluís Cano. (2012). Business Intelligence: Competir con Información. Bogotá: Escuela Banespyme.