Guía de desarrollo, Anexo 24.02

Proceso de exportación al almacén de datos PWC

Almacén de Datos

GERENCIA DE INFORMÁTICA

JOSEFA VALCÁRCEL, 44
28027-MADRID

Índice General

1 NATURALEZA DE LAS TABLAS DE INTERCAMBIO 3
2 PERIODICIDAD Y NATURALEZA 3
3 ESPECIFICACIÓN DE LA ESTRUCTURA DE LOS REGISTROS 4
4 MECANISMO DE TRANSFERENCIAS DE LOS FICHEROS 5

Índice de Ilustraciones y Tablas

Tabla 1. Ejemplo Periodicidad y Naturaleza 3
Tabla 2. Especificación de la estructura de los registros 5

1 Naturaleza de las tablas de Intercambio

La Interfaz de Entrada del *Data Warehouse* está compuesta por las tablas en Oracle que utilicen los Sistemas Operacionales existentes en la DGT, para que PWC se conecte a la BDDD es necesario que proporcionen un usuario con permisos de lectura.

2 Periodicidad y Naturaleza

Como parte de los requerimientos se precisa definir la periodicidad y naturaleza de carga que tendrán los datos. La periodicidad se refiere al intervalo de tiempo en el cual se descargara cada bloque de información, su naturaleza se refiere al contenido de dicha descarga, ya sea registros incrementales (altas, bajas y modificaciones) o descarga completa, todos los registros de la tabla origen.

Se definirá la profundidad histórica de la carga inicial junto con la historicidad misma de la tabla. (Ej. 5 años de carga inicial, 10 años de información online)

La periodicidad y naturaleza definida para las descargas del área "FUENTE" depende de las características de negocio de las tablas de origen y de las necesidades de disponibilidad en el sistema de destino.

En la siguiente tabla se definen dichos criterios a un nivel de entidad, las entidades están agrupadas por área del modelo para su mejor entendimiento.

Se incluyen todas las entidades dentro del modelo lógico de datos.

	Tabla	Periodicidad	Naturaleza	Ventana de Histórico
1	Tabla 1	Diario/Semanal/mensual/etc	Completa/Incremental	Si aplica
2	Tabla 2	Diario/Semanal/mensual/etc	Completa/Incremental	Si aplica
3	…	…	…	…

Tabla 1. Ejemplo Periodicidad y Naturaleza

3 Especificación de la estructura de los registros

La definición de los registros debe indicar:

- Posición del campo dentro de la tabla (1,2,3…n).
- Nombre del campo.
- Descripción del campo.
- Si el campo es obligatorio o no.
- El formato exacto del campo en el fichero de exportación:
    1. El formato de las fechas debe ser 'YYYY/MM/DD' o 'YYYY-MM-DD'.
    2. Si la hora es representativa dentro de la fecha en el sistema de origen ('datetime'), el tipo de dato en destino es 'Timestamp(0)' y su formato 'YYYY/MM/DD hh24:mi:ss' o 'YYYY-MM-DD hh24:mi:ss'.
    3. Si se exportan horas (time) el formato será 'hh24:mi:ss'.
    4. Si hay campos numéricos, se debe indicar el formato (escala, precisión, ceros por delante o no, signo ('+/-').
    5. Si hay campos numéricos con decimales, se utilizará el punto '.' como separador de decimales y NO se utilizará separador de grupos (millares, millones etc…)
    6. Si hay valores de dominio, es decir, un campo asociado a un diccionario de datos, todos los códigos y significados deben detallarse (valores admitidos. p. ej.: 'S'=sí, 'N' = no).

Esta información se detalla en una tabla como la del siguiente ***ejemplo***:

Funcionalidad: Contiene las notificaciones relacionadas con el expediente

	Nombre Campo	Descripción	Tipo (BD)
1	Campo1	**Campo obligatorio / No obligatorio** Descripción	Tipo de Dato
2	Campo2	**Campo obligatorio / No obligatorio** Descripción	Tipo de Dato

	Nombre Campo	Descripción	Tipo (BD)
3	Campo3	Descripción	Tipo de Dato
4	…	…	…

Tabla 2. Especificación de la estructura de los registros

4 Mecanismo de transferencias de los ficheros

La transferencia de la información se realizará utilizando PWC que se conectará al origen (Oracle) y al destino (Teradata) siendo el encargado del proceso de ETL, es decir, será el que procesa y carga la información en la base de datos del Almacén de Datos a través de sus propias herramientas.

La fuente será la responsable de proporcionar todos los datos para la conexión y transferencia de Oracle.

NOTA ACLARATORIA: las fuentes de datos una vez integradas en el DWH deben de mantenerse actualizadas en el tiempo, así comunicando el cambio de interfaz o cambio funcional que sufran dichas fuentes a la jefa de proyecto de almacén de datos de DGT