

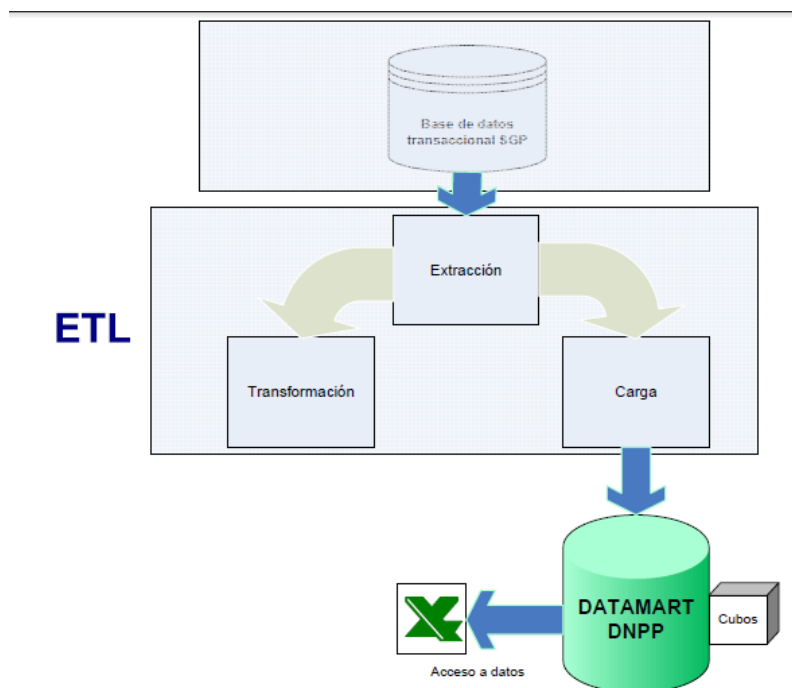
### **3.1.4 HERRAMIENTAS PARA EXTRAER, TRANSFORMAR Y CARGAR FUENTES DE DATOS**

La estructura de un DataMart (DTM) incluye:

1. Datos transaccionales: un origen o fuente de datos para poblar el componente de almacenamiento físico DTM. El origen de los datos son los sistemas transaccionales internos de la organización como también datos externos a ésta. En este caso, es la base de datos SGP.
2. Extracción de Datos: selección sistemática de datos operacionales usados para poblar el componente de almacenamiento físico DTM.
3. Transformación de datos: procesos para agregar y realizar otros cambios en los datos operacionales para reunir los objetivos de orientación a temas e integración principalmente.
4. Carga de Datos: inserción sistemática de datos en el componente de almacenamiento físico DTM.
5. DataMart: almacenamiento físico de datos multidimensionales.
6. Herramientas de acceso al almacenamiento físico DTM: herramientas que proveen acceso a los datos. Estas herramientas pueden ser herramientas específicas de mercado para visualización de bases multidimensionales como MS-EXCEL.

Los puntos 2,3 y 4 son conocidos como ETL, que significan Extract, Transform & Load (extraer, transformar y cargar). La extracción consiste en leer los datos de la base de datos original.

La transformación se refiere al hecho de limpiar la información incoherente o no válida, adaptar los tipos de datos, los valores de los campos, etc. Y por último la carga se refiere al hecho de guardar los datos en el formato adecuado dentro del DTM.



## CONCEPTOS DE BODEGA DE DATOS

### *Extracción, Transformación y Carga – ETL*

#### *\_ Identificar las fuentes.*

Hace referencia al proceso de identificación de las fuentes y datos que harán parte de la bodega o almacén de datos.

#### *\_ Extracción, transformación y carga (ETL – Extract, Transformation and Loading).*

Definición de las herramientas de extracción de datos desde las diferentes fuentes, el tiempo y la frecuencia con que los datos deberán ser cargados en la bodega o almacén de datos.

En la fase de **Extracción** se utilizan herramientas que facilitan el acceso y la obtención de datos almacenados en múltiples plataformas y bases de datos. En la fase de **Transformación** los datos extraídos son convertidos de su forma original a la forma que se requiere presentar para los usuarios de la bodega o almacén de datos. La transformación involucra la definición de reglas de conversión de los datos y/o combinación de datos de diferentes fuentes. La **Carga**, consiste en trasladar los datos transformados a la bodega o almacén de datos.

#### BIBLIOGRAFIA:

Bodegas de Datos: Consideraciones de Implementación, Seguridad y Control

Alfonso Luque Romero. Noviembre de 2006

<http://www.softein.com/phpBB2/temas/Datawarehouse.pdf>