



# TECHNOLOGY AREA

Last Updated: 6-7-06

DEFINITION	
<i>Name</i>	Extract Transform and Load (ETL)
<i>Description</i>	Extract Transform and Load (ETL) is a process that involves extracting data from multiple sources in various formats, transforming it to fit business needs, and ultimately, loading it into a target system. The target system will generally be configured as a data warehouse or data mart, though ETL can refer to a process that loads to any type of data storage structure. The structure itself will typically be a relational database, but may also be an application, file or other storage facility. The purpose of ETL is to reformat, cleanse and standardize data so that it can be analyzed or exchanged to address business needs and/or promote interoperability. Note that ETT (extraction, transformation, transportation), ETM (extraction, transformation, move), ELT (extraction, load, transform) may be used synonymously with ETL.
<i>Rationale</i>	The proliferation of messaging technologies, electronic reporting, GIS systems and geodatabases, and various applications aimed at collecting data, has created significant challenges for agencies attempting to manage data effectively. Adding to this influx of data is the transfer of huge amounts of data from legacy file systems into relational databases. Fast, comprehensive analysis of this data can provide vital information to allow agencies to better meet their goals or federal mandates requiring exchange of data. However, the volume of data can be so large it becomes extremely difficult for agencies to provide timely, useful information. Many of the current data warehouse systems are based on older architectures that weren't designed to handle enormous amounts of data. As a result, many data warehouse solutions are abandoned or severely under-utilized. ETL processes can be quite complex, and significant problems can occur if ETL systems are improperly designed or poorly implemented. An effective ETL system is vital to a successful data warehouse application. The third generation of ETL products allows for ELT (extract, load, transform). With the ELT approach data is transformed on the target after being loaded. It has been found that if the target database is powerful enough, it can be used to perform all transformations and optimize both performance and investment.
<i>Benefits</i>	A good ETL tool is able to communicate with the many different relational databases and read the various file formats used throughout an organization. ETL functionality may include the ability to read data from any electronic source, standardize, and write it to the target repository. ETL functions are being incorporated into Enterprise Application Integration, or even Enterprise Service Bus systems. Additionally, many ETL vendors now have data profiling, data quality and metadata capabilities to ensure the integrity of the data being processed, regardless of the source. These are addressed in other Technology

	Area documents.		
ASSOCIATED ARCHITECTURE LEVELS			
<i>Specify the Domain Name</i>	Information		
<i>Specify the Discipline Name</i>	Knowledge Management		
KEYWORDS			
<i>List Keywords</i>	Extract Transform and Load (ETL), Extract Load and Transform (ELT), Extract Transform and Transport (ETT), Extract Transform and Move (ETM), Data Warehouse, Data Mart, Data Cleansing, Extraction, Enterprise Application Integration, Enterprise Service Bus, Business Intelligence, Data Exchange, Information Exchange, National Information Exchange Models (NIEM), Metadata		
ASSOCIATED COMPLIANCE COMPONENTS			
<i>List the Compliance Component Names</i>	ETL-Extraction ETL-Transformation ETL-Data Load ETL Best Practices ETL Design ETL Administration ETL Tools ETL Data Structures ETL Logical Data Map		
ASSOCIATED PRODUCT COMPONENTS			
<i>List the Product Component Names</i>			
CURRENT STATUS			
<i>Provide the Current Status</i>	<input type="checkbox"/> <i>In Development</i> <input type="checkbox"/> <i>Under Review</i> <input checked="" type="checkbox"/> <i>Approved</i> <input type="checkbox"/> <i>Rejected</i>		
AUDIT TRAIL			
<i>Creation Date</i>	2/15/2006	<i>Date Approved / Rejected</i>	6/13/06
<i>Reason for Rejection</i>			
<i>Last Date Reviewed</i>		<i>Last Date Updated</i>	
<i>Reason for Update</i>			