

## Data Warehouse

Data Warehouse (engl., dt. etwa „Datenlager“) nennt man die zentrale Sammlung von Daten in einer Datenbank (Referenzdatenbank), die die Inhalte verschiedener Quellen zusammenfasst (referiert).

Die Inhalte (Daten) werden von den Quelldatenbanken ins „Datenlager“ (Data Warehouse) geladen und stehen dort zur Nutzung (Auswertung) bereit.

Sinn und Zweck des Data Warehouse ist die (virtuelle) Zusammenführung heterogener, häufig unterschiedlich strukturierter Daten an einer Stelle und die Schaffung einer bestände- bzw. quellenübergreifenden Suchmöglichkeit (nur eine Suchmaske für alle „angeschlossenen“ Bestände / Quellen). Zugleich ergibt sich ein Blick auf die Informationslandschaft einer Organisation / eines Unternehmens (welche Bestände / Quellen gibt es wo?).

Man kann die zentrale Referenzdatenbank des Data Warehouse umfassend verwenden; sie kann aber auch selektiv „beschickt“ werden. Die Auswahl-Beschickung (sog. ETL-Prozess) erfordert jedoch einen höheren Aufwand, der nicht immer gerechtfertigt ist. Mehr zum ETL-Prozess (Extraktion – Transformation – Laden [engl. Extract – Transform – Load]) unter: <http://de.wikipedia.org/wiki/ETL-Prozess>.

Mehr Information zu Data Warehouse: <http://de.wikipedia.org/wiki/Data-Warehouse>