



The Extraction Process

Let's examine the transformation components and see how data could change while passing through each of them.

- **Extractor and extract structure:** Extractors are programs that collect data from source systems. During the extraction, extractors move the contents of the required fields for data transfer via the transfer structure. The extract structure contains all the source system fields provided by the extractors.

Extractors combine and change data during the collection process. They do so via straightforward data delivery of certain fields of the single table, multiple outer or inner joins between views, tables, or ABAP programs with their own business logic. While investigating data inconsistencies, you could check how extractors gathered the data even before it flows into BW. Transaction **RSA3** in R/3 displays the screen where you perform the primary data quality check. You can see how the extractor presents data (e.g., what values the transfer structure fields display) on the DataSource level. In some cases, you could start looking for your problem from here after you discovered that your BEx or Web AD query brings you wrong results.

- **Transfer rules:** Transfer rules transfer data from the DataSource into the InfoSource and transform or modify data when necessary. We use transfer rules to determine which InfoObjects in the communication structure are filled with data from which fields in the transfer structure and what method BW uses to transfer this data. Five methods deliver this data:

1. Data transfers 1:1 (no manipulation)
2. The rule delivers a constant value for every row of the field
3. An ABAP routine delivers flexible data
4. A formula delivers flexible data
5. A start routine allows transformation and further delivery of data via a global ABAP routine, which could affect data from all fields at once)

All of these data transfer methods work for each field separately.

- **Update rules:** As part of the data staging process, an InfoSource delivers data to the update rules. Update rules

specify how BW posts the data (key figures and types of characteristics) into InfoProviders. They post data to the data targets and derive and modify data. Seven methods update the data into the data target:

1. Data transfers 1:1 from source characteristic (no manipulation)
2. A constant value fills the field's values
3. An ABAP routine delivers flexible (master and transaction) data
4. A formula delivers flexible data
5. The attribute update occurs by deriving another characteristic from the master data table
6. The initial value fills the field
7. A start routine allows transformation and further delivery of data via global ABAP routine, which could affect data from all fields at once)

- **InfoProvider:** InfoProvider is a structure that you can base a BEx query on. There are three types of physical and four types of logical InfoProviders. The physical ones are InfoCube, InfoObject, ODS. These are the actual containers that receive data during the upload. Physical InfoProviders can store the data in an aggregated or granular form. The logical InfoProviders are virtual InfoCubes, InfoSets, RemoteCubes, and MultiProviders. These are just queries that extract data from physical InfoProviders during the query run. You should check all types of InfoProviders to make sure that they use data correctly.

- **BEx report:** BEx report is a query that utilizes SAP BW reporting functions. BEx reports allow users to evaluate the dataset from an InfoProvider according to various characteristics and key figures. By selecting and combining the InfoObjects in a query, you can decide how to evaluate data from an InfoProvider.

- **Web AD report:** Web AD report is a user-built Web template that presents BEx queries online. Usually, a company uses Web AD reports when it doesn't want to install BEx GUI on every user's desktop. A Web AD report lets users access a centralized reports repository via SAP NetWeaver Portal or another Web application.