

METADATA MANAGEMENT

Metadata previously known as Data Dictionary is a collection of definitions and relationships that describe the information stored. When used in context of a data warehouse it is the data that defines the warehouse objects. Few examples of metadata are timestamp at which the data was extracted, the data sources from where metadata has been extracted and the missing fields that have been added by data cleaning or integration process. Metadata mgt involves tracking assessment and maintenance of metadata.

.ktr – Kettle transformation

Notepad – as XML Tags

Metadata can be divided into four groups:

1. Business Metadata
2. Process Metadata
3. Technical Metadata
4. Application Metadata

Business Metadata

Capture information such as business definitions structure and hierarchy of the data, aggregations, rules, ownership characteristics, subject areas and business rule, based description of transformation rules and definitions of business.

Process Metadata

Constitutes of all the metadata generated for various processes which includes

a) Source/Target Maps

b) Transformation Rules

c) Data cleaning Rules

d) Extract audit Trails

e) Transform audit trails

f) Load audit trails

g) Data quality audit

Technical Meta Data

This data consist of

- a) Data Location
- b) Data Format
- c) Technical Name
- d) Data Size
- e) Data Type
- f) Indexing

Application Metadata

The Company dashboard is build on the top of the datawarehouse. The dashboard is accessible by all the branch heads and senior executives. The following is

Data access history

Who is accessing

Frequency of access

When was it accessed

How was it accessed. Etc