METADATA

V. Vaishnavee

AP-AI-DS

SNSCE

Metadata

- Information about enterprise data

- Data about data stored in the warehouse
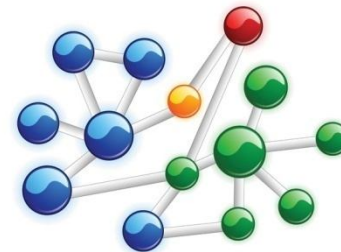
Metadata contains
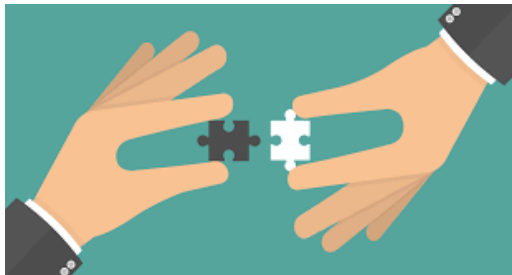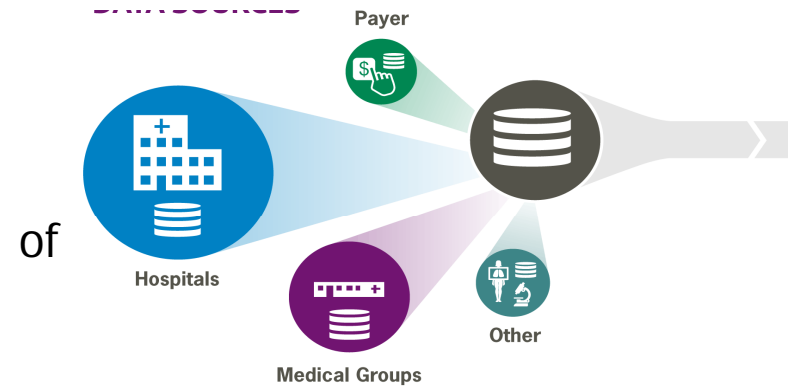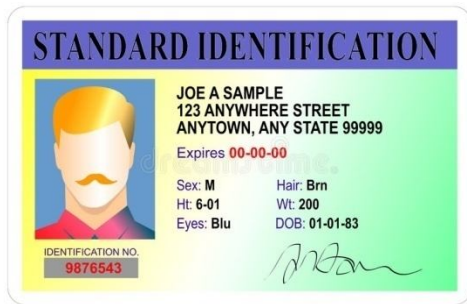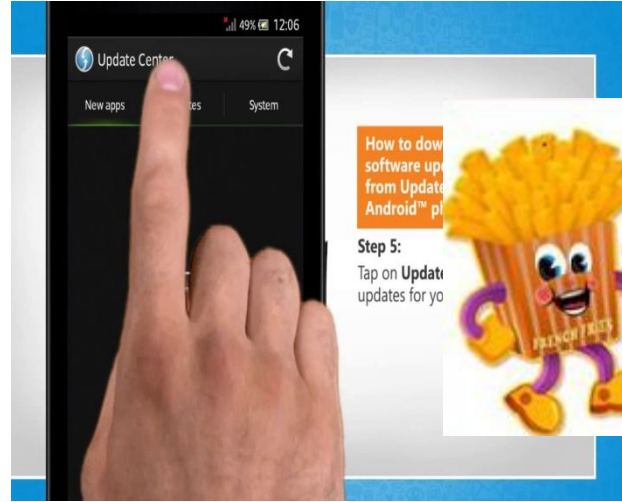
love

[luv] *noun.*

giving them the last piece
of cake. No matter how
much you want it

Metadata contains

of

Metadata contains

Metadata contains

- Location and description of warehouse system & data components (warehouse objects)

- Names, definition, structure and content of the data warehouse and end user views

- Identification of authoritative data sources

- Integration and transformation rules used to populate the data warehouse – includes algorithms used to convert, enhance transform data

- Integration and transformation rules used to deliver data to end-user analytical tool

- DW operational information – includes a history of warehouse updates, refreshments, snapshots, ownership authorizations

Metadata Interchange Initiative

- Launched by Arbor Software, Business Objects, Cognos, ETI, Platinum technology and Texas instrument - to bring vendors and users together to address a variety of difficult problems & issues with regard to exchanging, sharing and managing metadata.

- Objective:
  - To develop standard specifications for metadata interchange format
  - Define mechanism that allow vendors to exchange common metadata

Initial Goals

- Creating a
  - vendor-independent,
  - industry-defined and
  - maintained standard access mechanism and
  - standard API

  for metadata

- Enabling users to
  - control and manage access and manipulation of metadata through use of interchange-standards-compliant tools

- Defining a
  - clean, simple interchange implementation infrastructure that will facilitate compliance & speed up adoption by minimizing the amount of modification
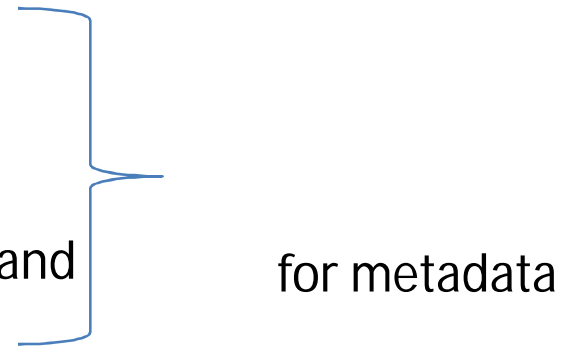
Metadata Interchange Standard – Two distinct metamodels

- ## Application metamodel
  - Tables used to hold the metadata for a particular application
  - Example, the set of tables used to store meta data in developer may differ significantly from those used by the Data Analyst

- ## Metadata metamodel
  - Set of objects that the metadata interchange standard can be used to describe
  - Information that is common (i.e., represented) by one or more classes of tools, such as data discovery tools, data extraction tools, replication tools, user query tools, etc.

Brain Teasers

- https://www.youtube.com/watch?v=rqvadvWZCgU

Metadata Interchange Standard Framework

- Metadata itself store

- any type of storage facility or

- format such as
  - relational tables,
  - ASCII files ,
  - fixed format or
  - customized formats

- the Metadata interchange standard framework will translate the an access request into interchange standard syntax and format
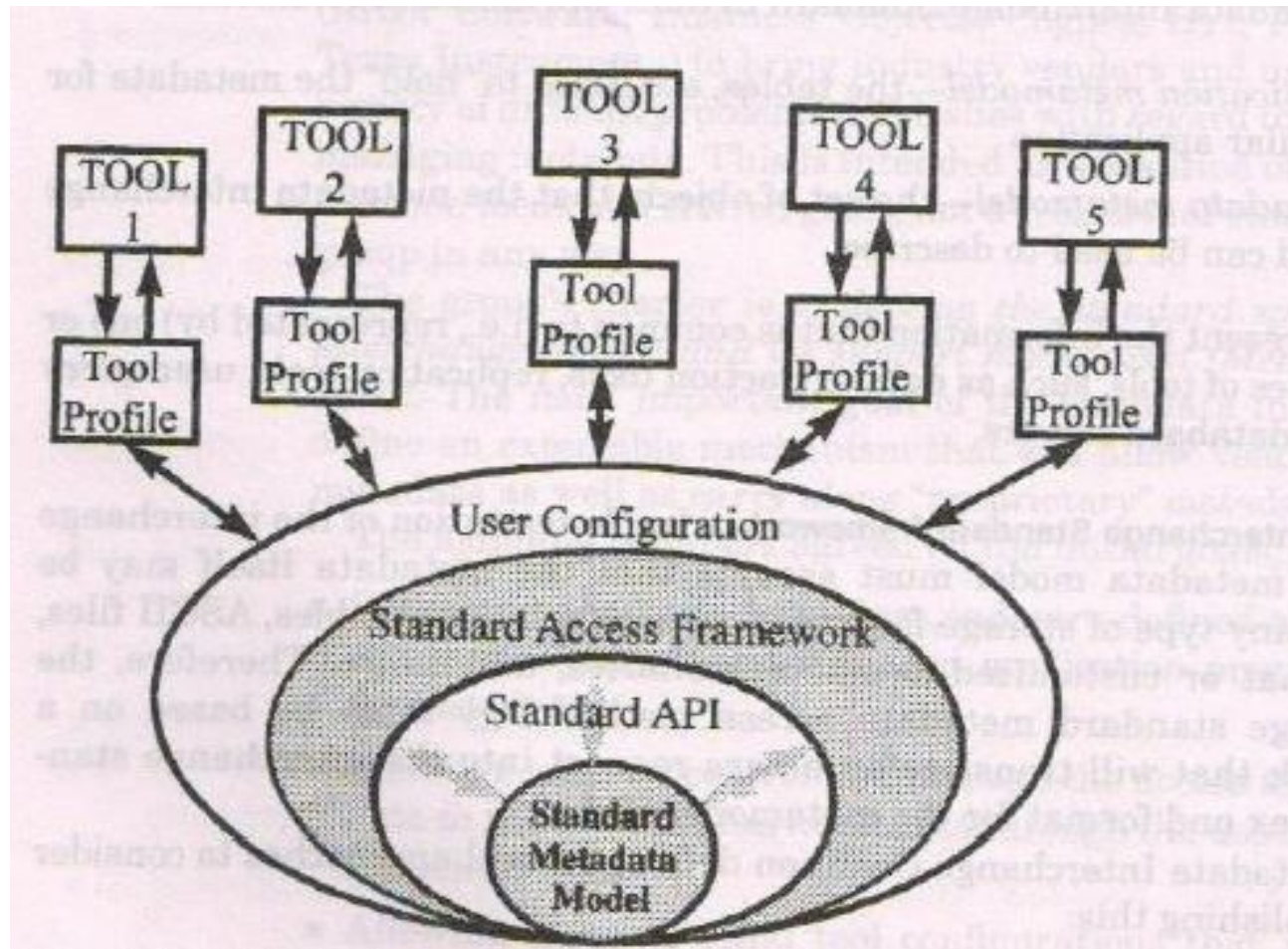
Metadata Interchange Standard Framework

- ## Procedural approach:
  - Individual tool's interaction with defined API.
  - Intelligence to communicate with API – built into tool –wherever tool may need to create, update, access or interact with metadata

- ## ASCII batch approach:
  - Entire ASCII file containing metadata interchange standard schema and access parameters is reloaded whenever a tool accesses metadata through standard API.

- ## Hybrid approach:
  - Follow data-driven model
  - Implementing table-driven API that support only fully qualified references for each metadata element, a tool could interact with the API through the standard access framework

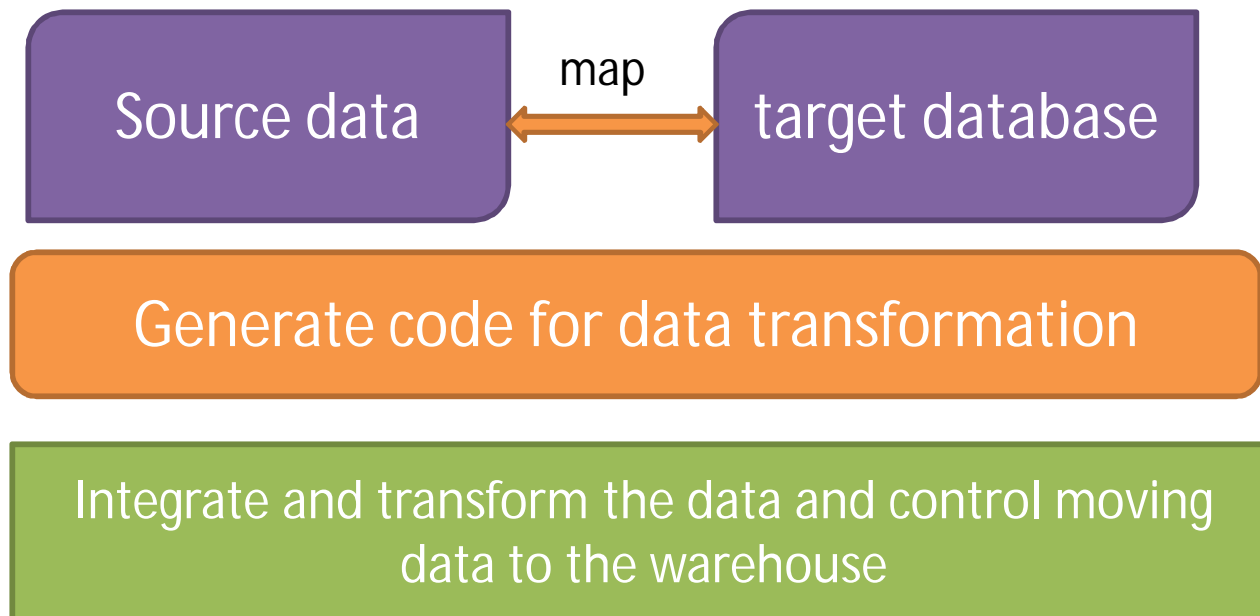# Proposed layers of proposed data interchange standard

Components of metadata interchange standard framework

- Standard metadata model
  - ASCII file format used to represent metadata that is being exchanged
- Standard access framework
  - Describes minimum number of API functions a vendor must support
- Tool profile
  - Provided by each tool vendor
  - File that describes what aspects of the interchange standard metamodel a particular tool supports
- User configuration
  - File describing the legal interchange paths for metadata in the user's environment
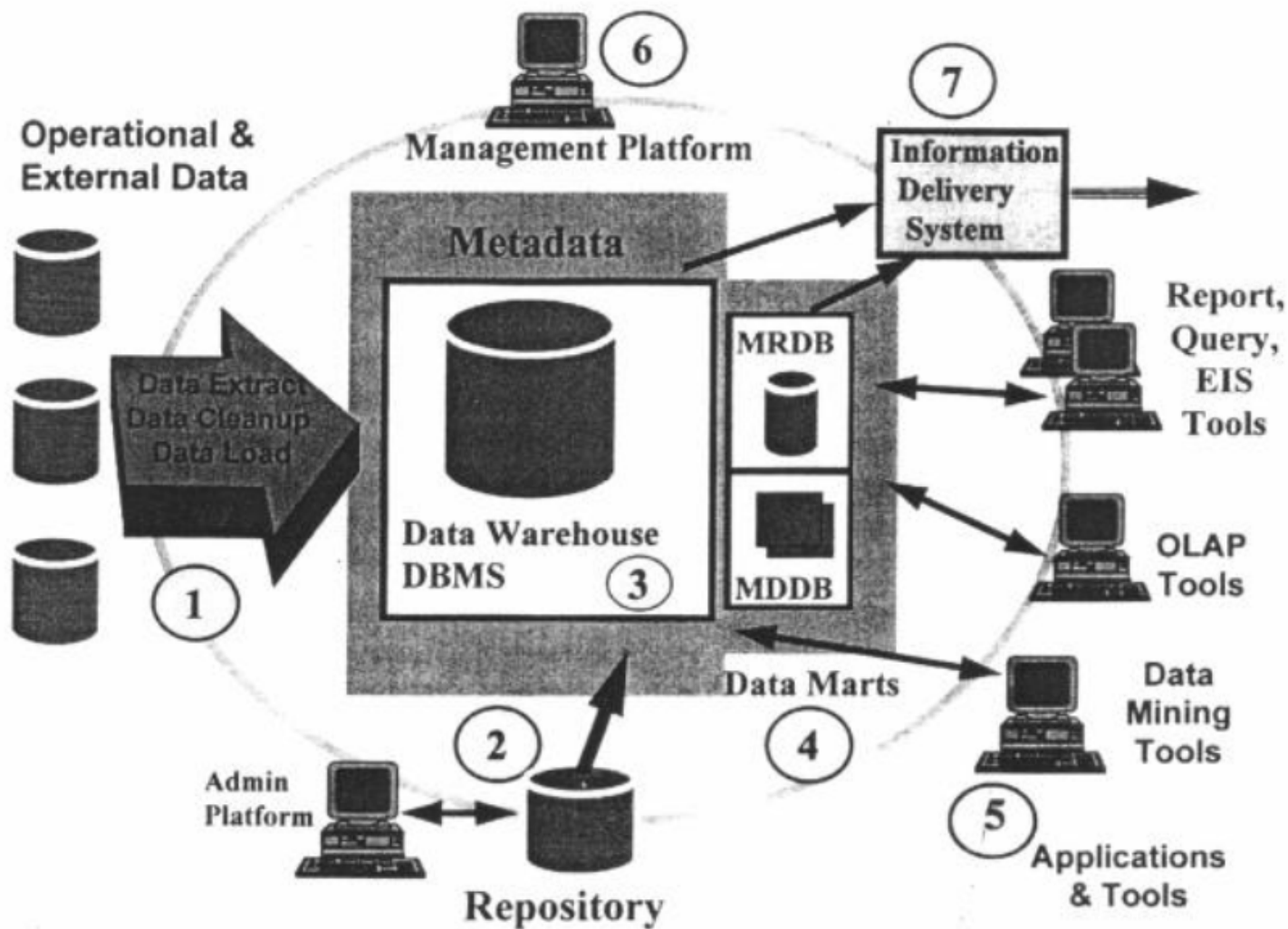
Metadata Repository

- Metadata is inside metadata repository and managed by metadata repository.

- Metadata repository – used to

| Source data | map ⟷ | target database |

Generate code for data transformation

Integrate and transform the data and control moving data to the warehouse

Metadata Repository

Benefits of having metadata repository in data warehouse framework

- Reduces and eliminates information redundancy, inconsistency

- Increases identification, understanding, coordination, utilization of enterprise information assets

- Simplifies management and improves organization, control, and accounting of information assets

Metadata integration point
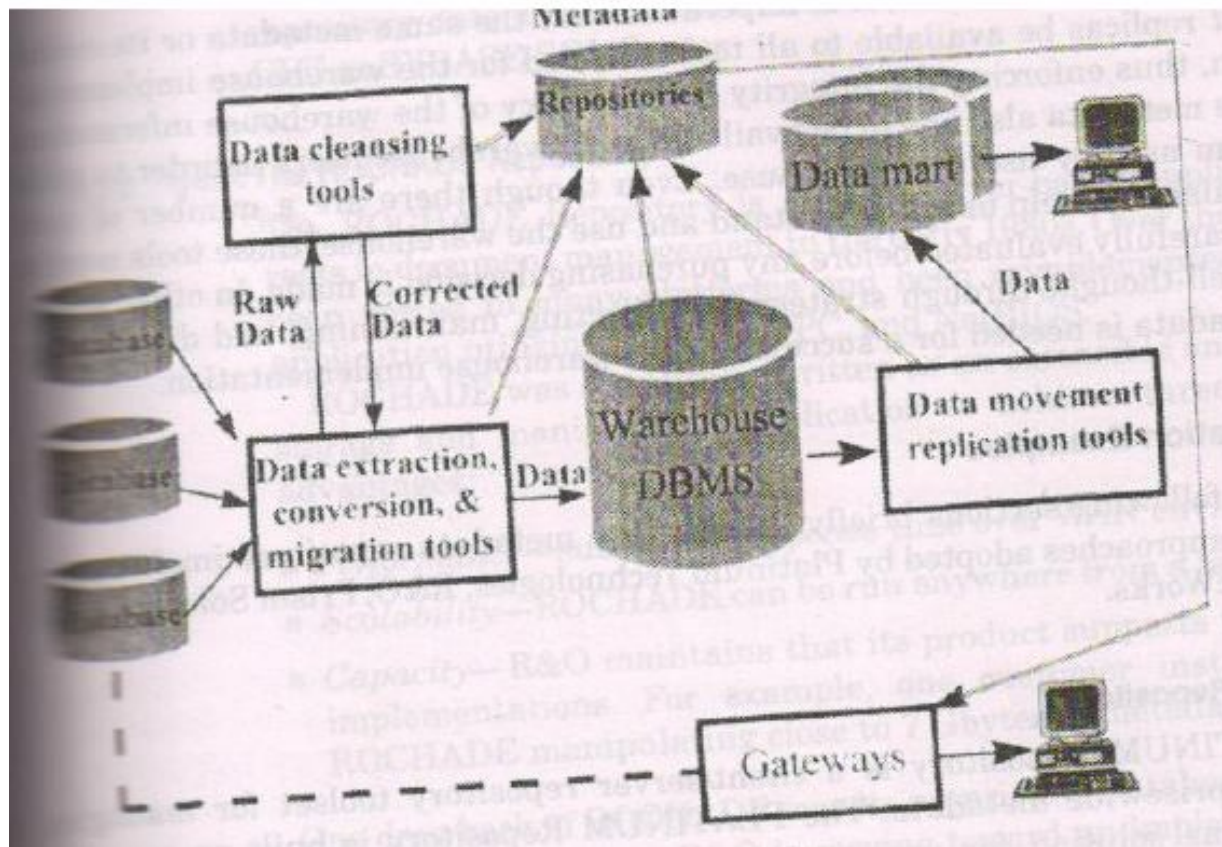
Metadata access and collection is indicated in double lines

Metadata Management

- Problem of DW:
  - inability to communicate to the end user what information resides in DW and how it can be accessed.

- The key to providing users and applications with a roadmap to the information stored in the warehouse is the meta data.

- It defines all data elements and their attributes, data sources and timing, and the rules that govern data use and data transformation. Meta data needs to be collected as the warehouse is designed and built. Must enforce integrity and redundancy.

Assessment