Modeling Pattern Characteristics

Analyzing Modeling Pattern Characteristics & Approaches

GENESEE ACADEMY, LLC

2013

Authored by: Hans Hultgren
# Index

INDEX .....................................................................................................................................1  
FORWARD ................................................................................................................................2  
CHARACTERISTICS ......................................................................................................................2  
  CHARACTERISTICS SUMMARY ....................................................................................................4  
The Modeling Approaches .........................................................................................................5  
  DATA VAULT MODELING ...........................................................................................................5  
  ANCHOR MODELING ..................................................................................................................5  
  3NF Normalized Modeling .........................................................................................................6  
  Dimensional Modeling ...............................................................................................................6  
Final NOTE ..................................................................................................................................7  
More Information .........................................................................................................................7
Forward

Today there are several new agile data modeling approaches designed to address the unique requirements and constraints of data modeling for the enterprise data warehouse. These approaches deploy certain techniques to help manage these requirements and constraints. While there are several characteristics that differentiate these approaches, this analysis will focus on three specific pattern characteristics/criteria. These are a) Encapsulated versus Decomposed table patterns, b) Abstracted versus Business Concept orientation, and c) Generic versus Attributed context. In the next section these three pattern characteristics will be defined. In the following section we will consider specific modeling approaches and how they can be related to these three criteria.

Characteristics

I. Encapsulated versus Decomposed Table Patterns

Encapsulated table patterns have been the standard in data modeling for the past 30 years. Today the primary forms are third normal form 3NF and dimensional modeling. In each of these forms a core business concept is represented by an Entity (or dimension) that includes the business key and all context attributes within the same table.

Decomposed table patterns. As is now common for data warehouse modeling, there are patterns that practice forms of table decomposition. This means that the business keys and context attributes that define a core business concept are translated to a set of interrelated tables. These tables are parts of the whole – an Ensemble – and together they represent the same information as would be in the encapsulated table forms.

In this diagram we see an Encapsulated table form (Entity) next to a Decomposed table form (Ensemble). An Ensemble complies with Unified Decomposition which means that all parts are considered together and only have meaning in relation to the whole.
II. **Abstracted versus Business Concept** Orientation

**Abstracted** concepts are forms of super-typed or high-level entities. In effect these are classifications of things found at the top levels of taxonomies or hierarchies. So a customer, vendor and employee are all types of **party** at an abstracted level. We tend to use abstracted concepts primarily for information modeling and in that way they are often found in industry reference models as well. There are those who also use abstracted concepts in data modeling (logical and physical).

**Business Concepts**, or **Core Business Concepts**, are the business-driven, natural levels of entities. They represent the level at which the business actually transacts, creates, discusses, uses and reports on these entities. So customer, vendor and employee are common business concepts.

This diagram illustrates Business Concept level in comparison to the Abstracted level. Notice that the level of concept is determined by the frequency of actual usage in the business (the number of times the concept is referenced in business).

III. **Generic versus Attributed** Context

**Generic** Context refers to generic forms being applied in capturing and storing the descriptive information about the concepts in our data warehouse. Commonly the generic form is name/value pair (NVP) although it can also be other forms of n-structured data. So the data stored without a defined and modeled schema (attributed data model). Records/Rows in these tables include two parts 1) the name of the attribute/ or tag, and 2) the corresponding data value for that
instance. Example would be a table that has effectively two attributes: name & value. The records within this table then define the attribute name matched with each data value (fname:Hans; lname:Hultgren; email:hans@edwi.org and etc.). This approach is sometimes referred to as data-driven or vertical.

**Attributed** Context refers to actual modeled schemas that include specific attribute names in the data model. This is the traditional approach to data modeling where the context that defines the entity or concept is analyzed, designed and modeled into the table structure. This approach is sometimes referred to as model-driven or horizontal.

Notice here that the NVP pattern above has effectively only two attributes, Name and Value. This form does not communicate the context attributes that are planned (anticipated, expected, or required). To understand what types of context attributes might exist we need to query the data in the table.

In the Attributed model above right we can see that the schema itself communicates the specific attributes that have been designed into the model. In this case the model itself communicates the specific context attributes that exist.

**Characteristics Summary**

The three (3) features above are among the several modeling characteristics that we need to consider when we establish the modeling pattern for our data warehouse program. The focus of this analysis is to compare various data modeling approaches. So for comparison purposes, this analysis is limited to the three factors discussed above.
The Modeling Approaches

For this first installment of this analysis we are considering only Data Vault Modeling, Anchor Modeling, 3NF Normalized and Dimensional Modeling. The modeling approaches considered in the broader analysis include Data Vault Modeling, Anchor Modeling, Focal Point Modeling, Head & Version, Hyper Agility, 3NF Normalized, 3NF Generic, and Dimensional Modeling.

Data Vault Modeling

Anchor Modeling
Notice that each of these four (4) modeling approaches are very similar in terms of the Business Concept orientation and the Attributed context. None of these forms are inherently moving towards Abstracted Concepts or Generic context. However, there is a major difference separating the first two forms (Data Vault and Anchor) from the last two forms (3NF Normalized and Dimensional). This is a very important observation – that the only major difference between these forms relates to the Decomposition of tables (in the first two) and the Encapsulation of tables (in the last two).
**Final Note**

The patterns that explore Abstracted Concepts or explore Generic Context represent yet another set of modeling approaches.

**More Information**

For Seminars, Training and Certification on Data Modeling, Data Vault, and Data Warehousing please see [GeneseeAcademy.com](http://GeneseeAcademy.com)

For online training including updates and current topics please visit us online at [DataVaultAcademy.com](http://DataVaultAcademy.com)