ENSEMBLE MODELING

Hans Hultgren
May 2016
Ensemble Modeling

Ensemble Modeling represents a family of modeling forms that share a common purpose and a common modeling paradigm. Ensemble forms address our need for data integration, historization, auditability and modeling agility. Today over 1400 models exist using this technique with the majority leveraging the Data Vault modeling pattern. This session will cover the need, the approach, the underlying premise and the current flavors of Ensemble Modeling. Attendees can expect to understand why organizations should consider Ensemble Modeling for their DWBI program.
AGENDA

1. Modern Data & the EDW
2. Ensemble Modeling
3. Data Vault Modeling
Session Objectives

• Ensemble Modeling
  – What is EMF?
  – What are the Benefits?
  – How does it fit in my DW?

• EDW New Paradigm
  – Modeling Agility
  – Ensemble Modeling
  – CBC, Process & Architecture

• Data Vault Modeling
  – DV backbone, Colors, Model Review
  – Hubs, Links, Satellites and Big Data
DATA WAREHOUSING
Data Warehousing

*Integrated, Non-Volatile, Time-Variant, Subject Oriented Data in support of [the Business]*

+ Agile
+ Auditable
+ Predictable
+ Scalable & Repeatable
About Change

Operational Systems

Changes are an exception. The result of some unexpected event or perhaps poor management of business requirements.

Changes are to be avoided or at least minimized.

Data Bases

Changes are rare. They do happen as business rules change or subject areas expand. Mostly these should be foreseen changes were there is adequate time to plan for them.

Changes should be minimal and predictable.

Data Warehouse

Changes are innate and constant. Changes occur all the time. The data warehouse is in a state of perpetual change – adapting to new and changing sources, new subject areas, new context and new and changing business rules.

Change is the only constant in the Data Warehouse.
Accept Change / Embrace Change

• Recognize that Change is ever-present in your data warehouse
• Accept and Embrace Change

Agility is “The measure of our ability to adapt to Change”
Agility is the primary feature of the enterprise data warehouse
Ensemble Modeling

• Data Modeling for the Data Warehouse

• A family of data modeling approaches that are optimized for the data warehouse

• Data modeling forms that are particularly strong at accommodating Change...
How do we Engineer for Change?

• Engineering for Change in Ensemble Modeling begins with one common premise; one common approach...

  Every time anything changes it impacts the whole thing!  
  Hmm, why don’t you separate the things that change from the things that don’t change?
ENSEMBLE MODELING
Unified Decomposition™

To engineer for change (agility) means that we break things out into component parts (Decomposition). This allows for a data warehouse that facilitates the capture and integration of things that are either interpreted in different ways or changing independently of each other.

To maintain the Key Dependency that holds our Core Concepts together, engineering for agility also means we keep these parts together (Unified).
Ensemble Modeling™

• The constellation of component parts acts as a whole – an Ensemble. 

All the parts of a thing taken together, so that each part is considered only in relation to the whole.

• With Ensemble Modeling the Core Business Concepts that we define and model are represented as a whole – an ensemble – including all of the component parts. An Ensemble is typically based on all things defining a Core Business Concept that can be uniquely and specifically said for one instance of that Concept.
Ensemble Modeling Forms

- Anchor
- Focal Point
- Data Vault
- Hyper Agility
- Temporal

- Your Style
- Anchor Vault
- 2G
- Matter
- DV2.0
DATA VAULT MODELING
The Data Vault Ensemble

- The Data Vault Ensemble conforms to a **single key** – embodied in the **Hub** construct.

- The component parts for the Data Vault Ensemble include:
  - **Hub**  *The Natural Business Key*
  - **Link**  *The Natural Business Relationships*
  - **Satellite**  *All Context, Descriptive Data and History*
Ensemble & **thinking** differently

- The minimal construct then for an “entity” such as “Customer” is now (in data vault) a **Hub with a set of Satellites**
Applying the modeling pattern
Data Vault Ensemble Benefits

• **Business**
  - Ability to adapt quickly to new business needs
  - Data is traceable allowing for a fully auditable, integrated data store
  - Allows the EDW to absorb all data all of the time
  - Easily adapts to new data sources and changing business rules – without expensive re-engineering
  - Results in an Data Warehouse with lower total cost of ownership (TCO)

• **Projects**
  - Ideal for agile development techniques resulting in lower project risk and more frequent deliverables
  - Can be built incrementally without compromising the core architecture

• **Architecture**
  - Parallel loading
  - Architecture that supports future expanded scope
  - Can scale to virtually any size
Ensemble

- An **Ensemble** is a Core Business Concept
  - Person, Place, Thing, Event
  - Business-Driven
  - Atomic Level
  - Identifiable

- An **Ensemble** contains only
  - One **Hub**
  - Directly attached **Satellite(s)**
  - All related **Link(s)**
Hub

- A **Hub** Construct in Data Vault
  - contains Business Key
  - only the Business Key
  - contains No Context
  - is always 1:1 with EWBK

- A **Hub** Table contains only
  - Business Key
  - Surrogate Key (Data Warehouse)
  - Load Date / Time Stamp
  - Record Source
- A **Link** Construct in Data Vault
  - contains Relationship
  - only a Relationship
  - contains No Context
  - is always 1:1 with Relationship

- A **Link** Table contains only
  - 2-n FKs for the Relationship
  - Surrogate Key (Data Warehouse)
  - Load Date / Time Stamp
  - Record Source

- Unique
- Specific
- Natural Business Relationship
Satellites

- A **Satellite** Construct in Data Vault
  - contains Context only
  - has no FKs (no relationships)
  - Designed by * Rate of Change
    * Type of Data * System...

- A **Satellite** Table contains only
  - Business Key FK +
  - Load Date / Time Stamp
  - Context Data...
  - Record Source
Sales DV Model - Backbone
Business Driven Modeling with DV
Data Vault Model – How it Looks
Applying Data Vault

- Financial Institutions
- Telecommunications
- Retail
- Manufacturing
- Technology
- Energy & Utility
- HealthCare
- Consultancy
- Transportation
- Government
- Education
- Etc.
About Data Vault Ensemble

Estimated 1400 + Data Vault based Data Warehouses around the world
Links and Information

CDVDM Training & Certification
www.GeneseeAcademy.com
Hans@GeneseeAcademy.com
gohansgo

Book  DataVaultBook.blogspot.com
HansHultgren.WordPress.com
HansHultgren

Online video-lesson training
DataVaultAcademy
DataVaultAcademy.com