# Scholarly Response 1: Data Warehousing Paradigms

**Rohan Purohit**

**MSCD 681 Data Warehouse Design**

# College of Computer & Information Sciences

# Regis University

**Scholarly Response 1: Data Warehouse Paradigms**

Bill Inmon and Ralph Kimball are considered two of leading authorities in data  
warehouse research. Compare and contrast each approach to data warehouse design  
and implementation. Be sure to include the major guidelines for each paradigm  
Also, research data warehouse best practices and provide critical success factors in  
implementing a data warehouse solution.

**Introduction**

The data warehousing and business intelligence industry has evolved considerably since its inception in the early 1970’s. Today there are many organizations that use data warehouses and business intelligence systems to derive business relevant insights from petabytes of historical consumer data. A traditional data warehouse contains a massive amount of highly detailed, time series data for decision support. (Breslin, p.2) With time not only has there been a change in DW/BI technology but also DW related logical modelling. Dimensional modeling popularized by Ralph Kimball has emerged as a leading architectural practice for building DW/BI systems.

Today there are two main methods of data warehousing. The first method is attributed to Bill Inmon. This style is a top down development approach that adapts traditional normalized relational database tools to develop an enterprise wide data warehouse. From the enterprise wide data store, individual databases are developed to support decision support needs. (Breslin, p.1)

The second method is attributed to Ralph Kimball. This style is a bottom up approach that uses dimensional modeling. And rather than building an enterprise wide database we are advised to build one data mart per major business process. Individual data marts are then integrated using a data bus standard. (Breslin, p.1)

**Analysis & Discussion**

Inmon proposes three levels of data modeling. The first level is the Entity Relationship Diagram (ERD). Inmon notes that the development team should create a set of ERD for each department that will come to utilize the data warehouse. The second level is the data item set (DIS) which is constructed for each department. The DIS contains a primary data grouping, a secondary data grouping and a connector signifying the relationships of data. (Inmon, p.4) The final level of Inmons data model implements the physical model. In this final model the mid-level data model is extended to include keys and physical characteristics of the model.

Inmon recommends that we should denormalize tables in the physical level. The data warehouse is rarely updated as data is historical and this makes it convenient to bypass requirements that would otherwise require normalization in an operational data scenario. The next steps in Inmons spiral development methodology involves selecting the granularity of transactions. Once granularity is resolved the subject areas or department specific databases are selected. Inmon stresses that once a set of processes is implemented its subsequent iterations should be easier to create other departmental databases.

Kimball proposes to start with dimensional modeling. This involves identifying fact tables which hold numerical data and the related dimensional tables which answer qualitative details about the numeric data. An aggregation of these fact tables and dimension tables is collectively called a star schema. Dimensional modeling bypasses normalization in order to make query performance optimum. Kimball formally recommends the four step dimensional design process. First, a business process is picked to reflect a process that embodies a measurable physical action in the real world. Second we establish the grain of the business process, the grain can be either periodic snapshot, transactional or accumulating snapshot. Third, we identify the dimensions which describe the data qualitatively. Fourthly we identify the facts that provide numerical data for the model. (Breslin, p.8)

Both Inmons and Kimball’s methods use time stamped data and ETL to prepare the data from the source systems and mediate it to the front room where business intelligence applications operate.

**Summary & conclusion**

Inmon sees the data warehouse as an integral part of the corporate information factory (CIF). Much of his methodology derives from the operational database world and therefore his methods are somewhat evolutionary. (Breslin, p.6) However, Kimball’s methods are more agile and revolutionary. Kimball notes that normalization or snow flaking would cause the analytical queries slower and maintains that redundancy is a key feature to get performance from a analytical query system.

According to Kimball it is important to know the distinct components that makeup a Data Warehouse/Business Intelligence environment. These are operational source systems, ETL system, data presentation area and business intelligence applications. (Kimball p.18)

Kimball makes a key point that when you use conformed dimensions and conformed facts of a set of dimensional model, you have a practical and predictable framework for incrementally building complex DW/BI systems that are inherently distributed.

An organization uses Inmons approach if it has a large data warehousing team and has a large project scope with enterprise wide access needs. Inmons approach takes longer to implement and is usually implemented in projects where we can wait to see results. Kimball’s approach is faster as it sports agile and specialized teams to develop a data mart that is expected to store mostly business metrics. (Breslin, p.13)

In addition an organization needs to consider preparation of the high level dimensional model, the preparation of detailed dimensional model development, concurrent model review and validation and comprehensive final design documentation by the project team to successfully implement a data warehousing project.

# References

Breslin, M. (2004). Data Warehousing Battle of the Giants: Comparing the Basics of the Kimball and Inmon Models. Business Intelligence Journal. p 1-15.

Kimball, R. & Ross, M. (2013). The Data Warehouse Toolkit: Third Edition. Indianapolis, IN. Wiley Publishing.