Name: Ahmed, Zubair

Professor: Virginia

Subject: Comp 440

Date: May 6, 2017

**Inmon vs Kimball**

There is an ongoing data revolution as more organizations have seen the need to mine their data for business intelligence to survive tough competition. They have therefore resulted to building data warehouses. In data warehousing, Inmon and Kimball have created a big debate concerning the approach that data warehouse architects should use to design data warehouses (Breslin 6). They have both come up with different conceptualizations of the ways in which the data warehouses need to be designed for optimal information management (Breslin 6). They have approached data warehouse design with different design techniques and strategies (Breslin 6).

Inmon's approach starts with the creation of a large centralized data warehouse for the whole organization (Breslin 7). This is then followed by the creation of several data marts for each of the organization's departments to better serve the specific needs of each type of users (Breslin 7). Kimball, on the other hand, believes in starting with the creation of small data marts to achieve departmental level analysis and reporting (Breslin 10). These data marts can then combined to come up with a large data warehouse (Breslin 10). The two also differ on the most desirable structure of a data warehouse. Inmon emphasizes on the creation of a relational model of a data warehouse (Breslin 7). The reasoning behind this is that, if data is in it’s a relational model, it can easily achieve enterprise-wide consistency (Breslin 7). Kimball, on the other hand, believes that the best database structure is a multidimensional one consisting of a star schema and snowflakes (Breslin 10). His reasoning, quite different than that of Inmon, is that users can themselves understand, analyze and explore data in consistency if data is structured in a dimensional model (Breslin 10).

The use of a central data repository for the whole enterprise and then the establishment of data marts to handle department-specific needs has earned Inmon’s method the name ‘top-down’ approach. The data warehouse stores atomic data, this is data that has the least amount of details (Breslin 8). Departments therefore make use of data marts to draw meaningful information from the atomic data (Breslin 8). Inmon's framework is therefore quite logical, having a central store with all sorts of data from where departments can mine their own intelligence. There are four core features of this data warehouse. To begin with, it is subject oriented. The data in the warehouse is not just scrambled, instead, it is organized in such a way that closely related data is linked together (Breslin 9). This makes it easier for departments to create data marts and pull data that relates to their tasks. Another feature is that the data warehouse is time variant. Therefore, changes made on the data are logged and the warehouse can give reports of what type of data has been changed and when (Breslin 9). Non-volatility is another feature of Inmon's data warehouse design. This means that the data in the warehouse is not disposable, users cannot overwrite or delete it. It is stored as read-only and users can only read it but make no changes to it. This is so because all departments rely on it and inconsistencies would be introduced if users could delete/overwrite data in the warehouse that could probably be used by others. Lastly, the warehouse is integrated. It is connected to all organizational applications and collects their operational data and puts it together in one place (Breslin 9).

In Kimball’s approach, a lot more importance is given to the individual departments in a business and therefore, their data marts are created first (Breslin 10). These handle data concerning their specific departments. The data marts can be combined, if need be, to form a large data warehouse (Breslin 10). Kimball says that the purpose of a data warehouse is to provide a copy of data that is already structured for querying and doing an analysis (Breslin 10). His data warehousing architecture focuses majorly on easing the accessibility of data to end users and at the same time giving a data warehouse the highest level of performance (Breslin 11). Therefore, data that users from different departments need is stored nearest to them in data marts. When the data marts have already structured it, it can then copied to the large data warehouse. Users from other departments that want to access it can easily do so by querying already structured data in the warehouse. Kimball tries to remove all unstructured data from the data warehouse to ensure that it can achieve high performance. He also removes the need to normalize data in the warehouse.

Choosing the better of these two designs is tricky as it mostly depends on the nature of business an organization is in (Breslin 17). Inmon's design is more suitable for stable businesses that can afford the high initial costs for making the warehouse and can patiently wait through the long periods of time taken to build it (Breslin 17). It is complex to build but assures of more consistency than Kimball’s design. An insurance business, for example, is more suitable for this kind of warehouse design as it enables all departments to access information about their clients, agents, and plans from a single point. Kimball's approach is, however, more preferable for organizations that want local optimization for their departments, a low startup cost and to set up in the shortest time possible (Breslin 19). An organization in the marketing business, for example, can benefit most from Kimball’s design.

Work Cited

Breslin, Mary. "Data Warehousing Battle of the Giants: Comparing the Basics of the Kimball and Inmon Models." *Business Intelligence Journal* Winter 2004: 6-20. *ProQuest.*Web. 14 May 2017.