Whitepaper

Business conditions that drive the need for subject oriented datamarts and business intelligence

Executive Summary

Businesses today are driven to manage enterprise performance but face the reality that key business information at the enterprise level is still largely inaccessible. Regardless of the fact that organizations have spent countless millions on integrated systems in the past decade, the promise of integrated, enterprise level reporting and analysis capabilities is only now becoming a reality in corporate America.

Providing integrated enterprise management systems for reporting and analytics requires that systems be designed and built in distinctly different ways from systems that capture daily business transactions.

This paper explores the business and technology conditions that drive the need for data warehousing and business intelligence.

Terminology

Consumer or Information Consumer. A person, at any level of the organization, who has a need for information.

Operational System. A system used primary to capture information necessary to run a business. These systems are designed to *capture* the transaction rapidly. Leading ERP systems, accounting systems, custom billing systems, etc are all examples of operational systems.

Informational System. A system specifically designed for reporting and ad-hoc query. These systems are designed specifically to provide information *about* the transactions.

Business Conditions That Drive Datamart Projects

Business Issues

At a strategic level the need to effectively manage performance drives the need for datamarts in all organizations. Throughout the past decade corporations both large and small alike bought off on the promise that having an integrated system would provide an excellent reporting solutions." In

this case, *integrated* includes ERP systems, best-of-breed financial systems, packaged software systems, and custom build applications. Many believed that their new systems would streamline processes, integrate data, and provide instant management reports.

In many cases, the unfortunate truth was that the *promise* was a *myth*. While the new systems streamlined processes and integrated data, seldom did they provide the reporting that was so important in the selection process. The reality was that the
company now had a large data asset

With many integrated systems now installed, we find that consumers still need access to the information which drives the business. Consumers still seek a single version of "the truth" to answer the key business questions which must routinely be addressed.

Strategically there is a need to organize information into integrated enterprise management systems and make it generally available to consumers.

At a tactical level organizations also face a variety of issues that surface when interviewing users about their daily routine. These issues are addressed as questions below.

1. Do I spend more time collecting and preparing data than I do analyzing it?
2. Do meetings sometimes stall because of a disagreement between two or more people about the validity of numbers?
3. Can I access both summary and detail information quickly from within the same application?
4. Am I ever frustrated by my lack of ability to access data that I know exists within my organization?
5. Do I or my staff spend time rekeying data from reports into databases or spreadsheets?
6. It is impossible to predict all of the reports I will ever need?
7. Do I currently spend time trying to interpret data in reports?
8. Can I easily retrieve and review information that is critical to the success of my functional area?

While predicting the variety of ways that data must be formatted and presented is impossible, one can predict the universe of data needed to answer key business questions.

Enter the "data warehouse" or "datamart"– technology buzzwords which define both the universe of data and the computer location where answers to key business questions are stored. Its goal is to improve business analysis and decision making capability. For the purpose of this document, we will use the term *datamart* to define specific data subject areas (i.e., finance, sales, marketing, purchasing, etc.) which are stored in a database.

The datamart:

- holds information which addresses both strategic and tactical information needs.
- provides information which allows key operating functions to effectively manage performance.
- unifies information from various databases into a single database. By unifying the information, consistent business views may be produced.

Beyond storing information of interest to consumers, corporations also need to address their business reporting needs in a consistent manner. Organizing information in a datamart addresses half of the data availability issue. The remaining data availability issues are addressed by providing access tools for reporting, analytics, and ad-hoc query. These access tools fall under a banner commonly referred to as *Business Intelligence*. Business Intelligence will not be discussed in detail here; however, it is strategically equally as important as is organizing information into a datamart.

From a business perspective the need for the datamart is driven largely by the need to manage performance. Whether managing profit loss, reviewing last month's sales, increasing cross-selling capability, or structuring a volume purchasing agreements – these examples all highlight an organization's need to provide relevant and timely access to business information. The need to consolidate information after a merger can also drive an increase in datamart usage.

Technology Issues

When looking through a prism to understand technological reasons for datamarts, the reasons focus to a single point: *consistently managing information*. Datamarts provide a consistent means of addressing the following:

- operational vs. informational data
- business rules
- consistent business views
- production windows
- information availability

The remainder of this section provides additional detail to the above points.

Operational vs. Informational Data. It is commonplace for operational systems to fall short when it comes to reporting on informational data. This is because operational systems are designed with rapid data entry in mind. These systems *capture,* the transaction in real-time, are designed for high throughput, and require high availability. By contrast, informational systems can be real-enough

time. They provide information *about* the transactions. Updating informational data periodically meets most consumer needs. Consumers need the ability to flexibly query data in different ways – capturing the transaction is not important here. High availability of information is important but not critical to daily operations.

From a systems design perspective it follows that operational systems and informational systems have entirely different systems designs. The data itself, degree of normalization, data availability, and indexing requirements are completely different.

Business Rules. Only a few people in the organization may truly understand the business rules which govern how to properly construct queries or how to massage data so as to properly reflect *how* the business actually views the data. A key aspect of datamart design is addressing "the rules" and assuring that the rules are consistently implemented as data is placed into the mart. By addressing the rules upfront, consumers stand a much better chance of retrieving the right answer the first time.

Consistent Business Views. The requirement for users to have consistent business views of the data is critical. Consistent business views help to minimize risk. In the absence of a consistent business views, users often create queries that generate incorrect results. Having consistent business views minimizes risk and creates an environment where data may be made available to a broader consumer base.

Production Windows. When considering production windows, we must review both peak and off-peak periods. During peak periods, querying against an operational system negatively impacts system performance when optimal performance is needed most. During off-peak periods, it is common for longer queries to collide with nightly backup processes. In both cases, the move to a datamart lessens the impact of either of these conditions.

Information Availability. We briefly touched on the concept of *real-enough time* data as it related to information availability. By this we mean that information has *periodicity* (i.e., a frequency with which it is needed). Informational data is most often reported daily, weekly, monthly, etc – it is *not* needed up to the minute. While the operational system provides *real*-time data, few are the users, other than customer service types, who require it for business reporting.[1]

These are major technology drivers which lead organizations to build datamarts. Some additional reasons to consider include: (1) information is expressed in business terminology, (2) formatting is in business terms, (3) sort orders reflect those most often needed.

[1] Some die-hard managers will not let the idea of *real-time data* out of their sights. When this is so, a process walkthrough often exposes key process deficiencies that are usually sufficient to make the idea of real-time go away. For example, how real-time is an order when it sits on an employee's desk awaiting input?

Summary

Although companies have completed large, multi-million dollar integrated system projects, the need for integrated information persists to this day. In many cases providing consumers with timely answers to critical business questions requires a two pronged solution. First, that organizations build datamarts to house and integrate key business data. Second, that reporting systems be built in order to provide consumers with access to timely information. This would also include tools for reporting, analytics, and ad-hoc query.