Business Intelligence – Not a simple software development project

By

V. Ramanathan

Executive Director, Saksoft

Date: Mar 2006

India	USA	Asia	Europe
Phone: +91 44 2461 4501	Phone: +1 212 286 1083	Phone: +65 6224 2550	Phone: +44 207 754 7154
Email: info@saksoft.com	Email: info@saksoft.com	Email: info@saksoft.com	Email: info@saksoft.com

The contents in this document are updated constantly and meant for understanding and appreciating the issues discussed. They cannot be used as commitment from Saksoft in any contracts.

SAKSOFT

TABLE OF CONTENTS

1. Introduction

One of the critical errors organizations commit very often is to treat a Business Intelligence (BI) project like any other software project, even if they understand that the project is complex and the scope is large.

In this article we will try to explore the challenges of a BI project, by defining what they are, and by summarizing the areas to focus on, if we have to successfully deploy a BI solution.

The Top 10 Critical Challenges for Business Intelligence Success

More than half of all BI projects fail — make sure yours isn't one of them

Let's start with the bad news: More than half of all Business Intelligence projects are either never completed or fail to deliver the features and benefits that are optimistically agreed on at their outset. While there are many reasons for this high failure rate, the biggest is that companies treat BI projects as just another IT project. Face it: Business Intelligence, or BI, is neither a product nor a system. It is, rather, a constantly evolving strategy, vision and architecture that continuously seeks to

Organizations must understand and address these 10 critical challenges for BI success. BI projects fail because of:

1. Failure to recognize BI projects as cross-organizational business initiatives, and to understand that as such they differ from typical stand-alone solutions.
2. Unengaged business sponsors (or sponsors who enjoy little or no authority in the enterprise).
3. Unavailable or unwilling business representatives.

#Computer world white paper dated 30th June 2003

2. BI Solution – A Perspective

Before we explain what a BI Solution is, it is important to understand several wrong conceptions that people have about them. There are several simplified assumptions made about a BI project, each of which is incorrect. Some of them are:

- Since a BI project is a top-level information providing solution, the objective of integration is kept very broad or lacks focus, or has the wrong sponsors.
- Management Information System and BI are considered similar except that MIS summarizes operational data whereas BI looks at summarizing data for decision-making.
- A BI solution is assumed to be the same as a Data Warehouse project
- There is a failure to recognize BI Projects as cross-functional ones involving both IT and the Business. In some organizations, the IT Department sponsors the project while in others, the Business sponsors it.
    - IT sponsored projects have a broad scope, but run longer …
    - Business sponsored projects complete faster but the solution lasts for a shorter period
- There is a lack of recognition that data source identification and mapping are business-consulting issues and not IT issues.

First we must recognize that BI projects are about Information Integration and Presentation. The number and variety of such operational systems depends on the age of the business, the variety of products & services offered and the evolution of the enterprise. But for the people managing the business, the operations have to be seen in a consistent context – a context that enables them to view the business with a set of key business indicators. Then we must understand that information presented in a BI solution is a 'business key indicator' and not a traditional functional performance indicator like ROCE (Return On Capital Employed). Examples of key indicators could be 'Time to process a redemption request' for a Mutual fund, 'Components of Customer Balance' for a credit card business etc. There is a need to build these key indicators to know the health of business at any time. The critical step towards getting values for these indicators is integration of information from multiple operational systems.

From the above it is fairly clear that a BI solution is not the same as a functional MIS solution or a project on building a Data Warehouse and reporting numbers. It is a cross functional project that must have comprehensive business user representation,

with users who can understand the meaning and usefulness of data/information. Then the role of the IT team comes in, to get the work done with suitable tools or technology.

3. Information Integration

As a first step to creating a BI Solution, let us analyze the issues on hand. Given the variety of technologies involved in software and hardware that exists in these operational systems (from where data is sourced), the integration is not a simple activity. Various factors complicate this integration activity like:

- Identification of data[1] sources
- Elimination of duplicate data
- Cleansing data
- Filling in the information gaps (identifying surrogates)
- Identifying the integration platform (hardware & software)
- Identifying common / Single delivery channel for information delivery
- Ensuring data availability on time, consistently
- Ensuring the huge volume of data is processed on time
- Setting up processes to consistently perform the integration steps taking care of errors and exceptions that can creep in
- Cost of information integration is commensurate with business benefits
- Identifying key indicators and presentation of information on the key indicators
- Protecting the future – changes in business, business processes and technology

Finding solutions to each of the above issues is neither a pure business/ management consulting exercise nor can new information technologies provide a simple one-stop solution. The issues have both Business and IT dimensions. Every BI solution project requires a comprehensive Business Information Integration Framework[2] (BIIF) that gives the issue a common structural shape, so that there is a systematic approach to identifying the issues, analyzing them, and defining solutions.

We can broadly state the steps involved in information integration as:

- Raw data collection

[1] Throughout this document data and information are used in a general & interchangeable sense, however data in most places is used to refer to the raw form and information is used to refer to the refined /processed form for final use.

[2] Veri-sens is one such framework from Saksoft

- Data to Information
- Information Store
- Information Pre-cooking

Many technological solutions exist for each of the activities listed above. But no methodology exists on what method to use and how to use it.

Raw data Collection

Let us examine the raw data collection step. In a credit card business, for a million cards there will be a minimum of ten million transactions in a month. Do we pick all the information about the customers (e.g., multiple addresses) or do we pick just the card number alone? Does name or address have any use in the information analysis? Typical software companies want the users to help them answer these questions. Users in most organizations don't have adequate time to think, examine and find answers. And invariably the user tends to say, " I think I will need it" – for the user is not worried about the implications of the same on processing time or machine size. Here is where a good business analyst[3], well versed with operations of banking can really help. The analyst can structure the information requirement and work backwards to identify the source of data and availability of relevant data.

Similarly if we have the raw form of data coming from different sources, how do we get it transported to the final destination? The volume of data could be very large and normal information transfer bandwidth used for inter networking countries and branches may not be adequate for this volume transfer. For example one trans-national bank may want to transmit, from its operational center in Europe to its centralized processing center in Asia, a daily volume of 5GB of raw data. Due to operational systems constraints the time window available for the same may be as low as 2.5 hours.

[3] Business Analyst in this note implies a person who have experience / exposure to operations of the customers business and has excellent domain knowledge. He is not just an analyst, but has significant domain knowledge to act on behalf of the user.

Data to Information

When it comes to 'what information' is required, users struggle to define the scope and boundaries. In an IT sponsored project the tendency is to transport most of the data - the IT team typically thinks they can meet all of the users requirements in the future, forgetting the issue on volumes explained earlier. This is one reason why IT sponsored projects tend to go on for a long time without showing any tangible business result. The point to note is that a BI solution differs from any general reporting system (MIS or EIS) as the MIS or EIS systems present indicators, which are very general in nature whereas the BI systems present indicators, which are pointers to business health trends. Unlike MIS, the information here is not used for day-to-day decision making, but working out the strategic action plan or initiating corrective steps to stem a negative trend.

Information Store

Once we decide on possible data, which is to be converted to information by suitable processing, then the next issue is how to store this information and how much of it to store. In the earlier example, we discussed a daily raw data volume of 5GB. How many days of such information needs to be stored? A simple answer is "for a month" – but that leads us to a *volume of 300* GB. At 25 cents an MB, the cost of storage alone will now be $76,800 – which is really high when compared to the accruing benefits. Similarly in the Credit card industry, a 36-month trend is typical. But keeping 36-months data again implies a significantly large database. So a very prudent decision, that balances requirements and cost, is critical.

Information Pre-cooking

Another critical issue that keeps coming up is whether we should have pre-cooked information for delivery or provide a very flexible query tool. The extreme is to create reports and keep them pre-packaged, ready for presentation. This will be useful if not only presentation, but contents and selection options are agreed upon across users. The other extreme is to provide for a flexible query tool like Business Objects and let the user query the available raw information. While this gives users a lot of freedom, it fails in two ways. One failure is that there will be no standard report to compare notes across users and the second is that senior managers typically do not have the

time to use a query tool. They prefer reports coming to them at a few clicks of a button. An Actual solution may lie somewhere in between or a composite of all options depending on the nature and objective of the application.

4. Information Presentation

The next possible problem is to identify the key indicators for the BI Solution to present. One common set of indicators which are used are financial ratios like 'Operating Margin' or 'Return On Capital Employed' – but these ratios lack the ability to help build a strategic action plan as there can be numerous (divergent) points that contribute to them. **One critical ingredient for identifying a key indicator is that it must easily relate to an operating area** or a focus point. For example, 'Number of new applications' is a key indicator for a credit card business as it indicates efficacy of the sales & marketing function.

In identifying the right indicators, we need to address the following points:

- Does the set of key indicators together reflect comprehensively the status of business?
- What is the level of detail? If indicators are too detailed, the exercise of integrating information becomes unwieldy and we may miss the forest for the trees. If we have very macro or high level indicators, then the chances are that these indicators themselves are not helpful in devising action plans for correcting the course of the business.
- Are the definitions clear? All users must agree upon the indicator definition and its derivation must be clearly spelt out. (A good practice is to build a glossary of terms and formulas)

With the key indicators identified, each indicator must be dimensioned. Dimensioning is the process through which we close in on an action area. For example if we take new applications as an indicator, we must agree whether value for the day or month is more appropriate. Though daily figures enable quick actions once a negative trend is observed, if there are too many fluctuations on a daily basis, then the daily trend may be misleading. The result of acting on daily trend may be an over-reaction. Similarly we can dimension that we need the data for Region/Country, Product (e.g. Visa or Master) or Channel (direct mailer, through agents etc.) etc. Just like having the right number of indicators is prudent, the right number of dimensions is equally important.

To report the key indicators we must also check the quality of the source data. Quality of data refers to the availability of a consistent and correct representation of data. Wherever alternates are available to data, one needs to evaluate which alternate is

simple, consistent and correct. For example, we can determine that a credit card is closed by either looking at status flag or closure date. The status flag may not be consistently set when ever the account is closed. People, instead of setting the flag and closure date, may do only one of them.

The use of surrogates (e.g., if status is not set to any value, it is an open account) and the use of more than one form of data (both closure date and status code) are some possible alternatives to improve data quality. Thus we will have quality raw data that is going to be used in the business information integration exercise.

Once the key indicators are identified, the next step is to map how these indicator values will be derived from the available source data. The formula or logic of deriving the key indicator value must be consistent with the available data, and the dimensioning that was finalized for the key indicators should be realizable.

In many Information integration exercises, this mapping is considered a detailed technical design issue. In the BI solution context there are two critical differences. Firstly, it is business user driven – the business user must certify that the source to target mapping is purposeful. Secondly, such an exercise precedes any technical design (like DB design). Since data values (e.g., charge off reason code values) may change dynamically, designers choose a meta table option to store these mappings. The usage of a meta table provides flexibility in configuration. This can however, not delay the mapping exercise to a later phase. Quality output from information integration implies that we have clarity in the conceptualization of how output is to be derived from source right at the starting point.

5. Need for a framework

Given the challenges of a BI solution, we must have a defined framework (BIIF) for successfully completing a BI project. The issues we identified in earlier sections are just quick examples, and there are many more that need to be addressed such as 'What features and functionalities to have?" "What tools and technology to use?" "What security and access control aspects to consider?", "What data build-up and stages/storage aspects to consider?" etc.

A well-defined framework will answer or at least provide directions to take on these issues. The advantages of a good BIIF are:

- Structured approach to the information integration exercise
- Wealth of knowledge assembled from similar project experience
- Time to deploy shrunk by minimum 50%
- Proven methodology – no trial & error
- Modular implementation architecture enabling best mix & match strategy with 3rd party/ Existing products
- Help in selecting toolset & architecture – Right choice makes SW engineering efficient

6. Summary

A BI Solution is vital need for any large organization. The first requirement of a BI solution is information integration and it is not a simple issue of identifying the need, possible technology choices and constructing application software to meet the demand. The steps in the information integration are, to identify the source of data, identify required key indicators to present to business users, map the input and output data, build processes to collect and collate data, select technology to be used, manage the project of development and lastly, implement the solution successfully.

Finding solutions to the typical issues in a BI project is neither a pure business/management consulting exercise nor are new information technologies able to provide a simple one-stop solution. The issues have both Business and IT dimensions. A well-defined Business Information Integration Framework could provide a systematic approach to identifying the issues and solutions.