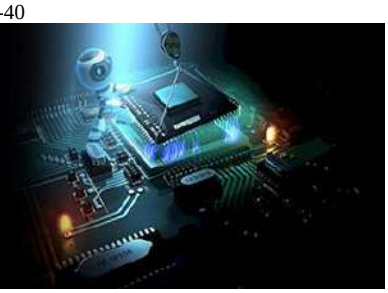
International Journal of Engineering in Computer Science



**E-ISSN:** 2663-3590
**P-ISSN:** 2663-3582
IJECS 2019; 1(1): 36-40
Received: 26-11-2018
Accepted: 28-12-2018

**Subash Somi**
Data Architect, Division of Information Systems, Data Diverse LLC, Washington DC, USA

**Correspondence**
**Subash Somi**
Data Architect, Division of Information Systems, Data Diverse LLC, Washington DC, USA

# Business intelligence efficiency

## Subash Somi

**DOI:** https://doi.org/10.33545/26633582.2019.v1.i1a.9

**Abstract**
A critical success factor for business intelligence (BI) programs is the adoption of a philosophy that implements repeatable processes and promotes reuse. Four of the most significant areas in which this philosophy needs to be applied are architecture, data, methodology, and governance. The BI architects must anticipate future needs but must be careful to deploy the architecture incrementally. BI teams must strive to deliver business intelligence functions every few months, and this requires a blueprint that can be divided so that each project builds on the infrastructure established by the previous ones. This approach must be applied to the data, as well. A subject area model and entity-level business data model can be developed relatively quickly and provide the data blueprint. The model can then be attributed as needed for each data warehouse increment, with each attribute being defined from an enterprise perspective. The methodology is a series of processes, and to the extent that each process is thoroughly defined, it becomes more reusable. Finally, the governance structure must adhere to the repeatability and reuse philosophy. A Center of Excellence provides a good way to capture and disseminate information about best practices, templates, and applications of the data. All of these works together to help organizations leverage their information resources.

**Keywords:** Business intelligence, business process performance measures, data analytics, data warehouse, information systems, business analytics

## Introduction
One of the primary goals of business intelligence programs is to provide consistent business information of its users. This consistency does not just happen. It requires a disciplined approach that incrementally deploys business intelligence while looking beyond the individual project to promote reuse of various aspects of the program – the architecture, the data, and the methodology steps, combined with an effective governance structure.
This article describes each of these areas and explains how an incremental approach and repeatable processes contribute to the success of the business intelligence program.

## Architecture
The business intelligence architecture needs to be incrementally deployed. To ensure that all of the pieces will eventually fit together, creation and enforcement of standards are mandatory. The scope of the first business intelligence project is often limited. It may address 100 data elements, create a single data mart, and have a limited set of query and reporting capabilities. Scope containment is important to enabling the delivery of business value within 60 – 120 days. With the pressure to meet the time constraints, temptations often exist to defer pieces of the architecture. There is nothing inherently wrong with deferrals as long as the team understands the risks of the omission and has a plan for subsequently including the deferred components.
For example, a global company may have three sales systems – one in the United States where 60% of the business is transacted, one in Europe where 25% of the business is transacted, and one in Asia where the remaining business is transacted. Additionally, the team knows there are significant data quality issues with the Europe installation. An incremental deployment approach is appropriate for such an environment.
The first step is to establish a high-level logical architecture that can encompass the long-term environment. One such architecture is shown in Figure 1. In addition to absorbing the three sales systems, the architecture will be resilient to adding more sources and providing additional query and reporting capabilities. With the architecture mapped out, the business intelligence team can determine the types of tools that will be needed and can initiate appropriate tool selection efforts.

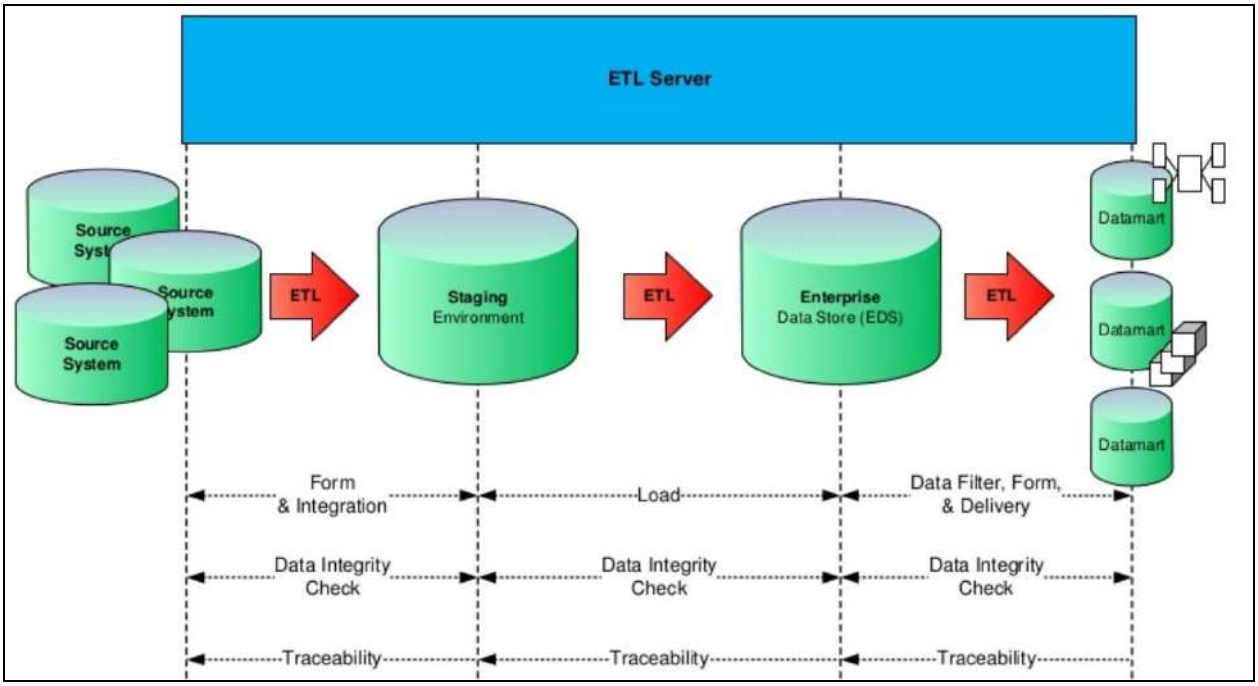



**Fig 1:** Architecture

The deployment sequence for the architecture can then be developed. The approach that follows is an example of an incremental deployment for this architecture. The approach presumes that the enterprise data warehouse has been designed based on the business data model.

**Initiative 1:** Deploy a data warehouse with the U.S. sales systems only. This project will include the establishment of the staging area and the data warehouse. It will include critical data cleansing activities as well as the generation of surrogate keys where needed. A data mart should also be created so that U.S. sales information is delivered to the business. Metadata should be provided with the data mart to ensure that the users understand that only U>S> data is included.

**Initiative 2:** Integrate data from the Asian system. This step introduces the complexities of data integration but avoids dealing with the suspected quality problems of the European system. The ETL programs developed for the first iteration can be reused or enhanced to meet the integration needs. The first data mart can absorb Asian data with no change. Additional data mart may be developed to provide new functionality.

**Initiative 3:** Integrate data from the European system. This system reuses the approach previously developed for data integration and introduces the needed data cleansing complexities. The ETL programs developed for the first iteration can be reused or enhanced to meet the integration needs.

**Iteration 4-n:** Incorporate additional source systems data based on business needs or build new data marts using data already in the data warehouse to support business intelligence applications.
The incremental deployment of the architecture requires the development and enforcement of relevant standards. These standards address the tools, naming standards, integration approaches, cleansing approaches, metadata capture, maintenance, and distribution, etc. To the extent that these are developed in advance, rework is minimized as the environment is expanded to absorb additional data or to deliver new capabilities.

**Data**
An effective data stewardship program is at the root of data consistency and reuse. Within this program, the data stewards exercise responsible care over the enterprise's data assets. When fully deployed, the stewardship program includes a group of data stewards, each of whom has responsibility for a specific set of structured and unstructured data. These stewards act as a group in setting overall policies and standards, and the individual stewards then apply these rules over the data for which they are responsible. The stewardship responsibilities directly contribute to the repeatability within the business intelligence program through their involvement in the business data model, data profiling, use oversight, etc.

**Business Data Model**
The business data model portrays the information of interest to the enterprise. Without this model, it is very difficult to achieve data consistency and reuse. Information Technology, in its role as the custodian for the data assets of the enterprise, provides data analysts to perform the mechanics of creating the model. The model content is provided by the data stewards. This content includes the major groupings of information (entities) and their definitions, the individual data elements (attributes) and their definitions, and the relationships among these. Development of a complete business data model can easily take six to twelve months, and few companies are willing to wait that long before applying it. A better approach is to develop an entity-level model quickly and then to develop the details in conjunction with business intelligence initiatives that apply the model to the design of the data warehouse. This approach provides the benefits of the

business data model without the delays required to develop the full model.

**Data Profiling**

The business intelligence environment does not create raw data; it captures data from existing data stores and transforms it into information to support analysis and decision making. Data profiling recognizes that the data within the source databases are not necessarily accurate, complete, or consistent. The process analyzes the source data to identify deficiencies in this area so that the ETL (Extraction, Transformation, Load) processes can perform the necessary adjustments to provide consistent data to the business intelligence user community.

Information Technology analysts and programmers work in conjunction with the data stewards to understand the source data. Specialized tools can speed up this process, but ultimately its effectiveness is dependent on the rigor (e.g., repeatable process) that is applied and, on the data, stewards' knowledge and analytical skills. Data profiling consists of determining the sources, tables/files, and elements that need to be analyzed, analyzing the data, and determining the appropriate data quality actions that need to take place during the ETL processes. Data profiling does not actually fix the problems – it uncovers them. The need for additional corrective actions within the source systems or business processes may be identified, but pursuing these is outside the scope of the business intelligence project.

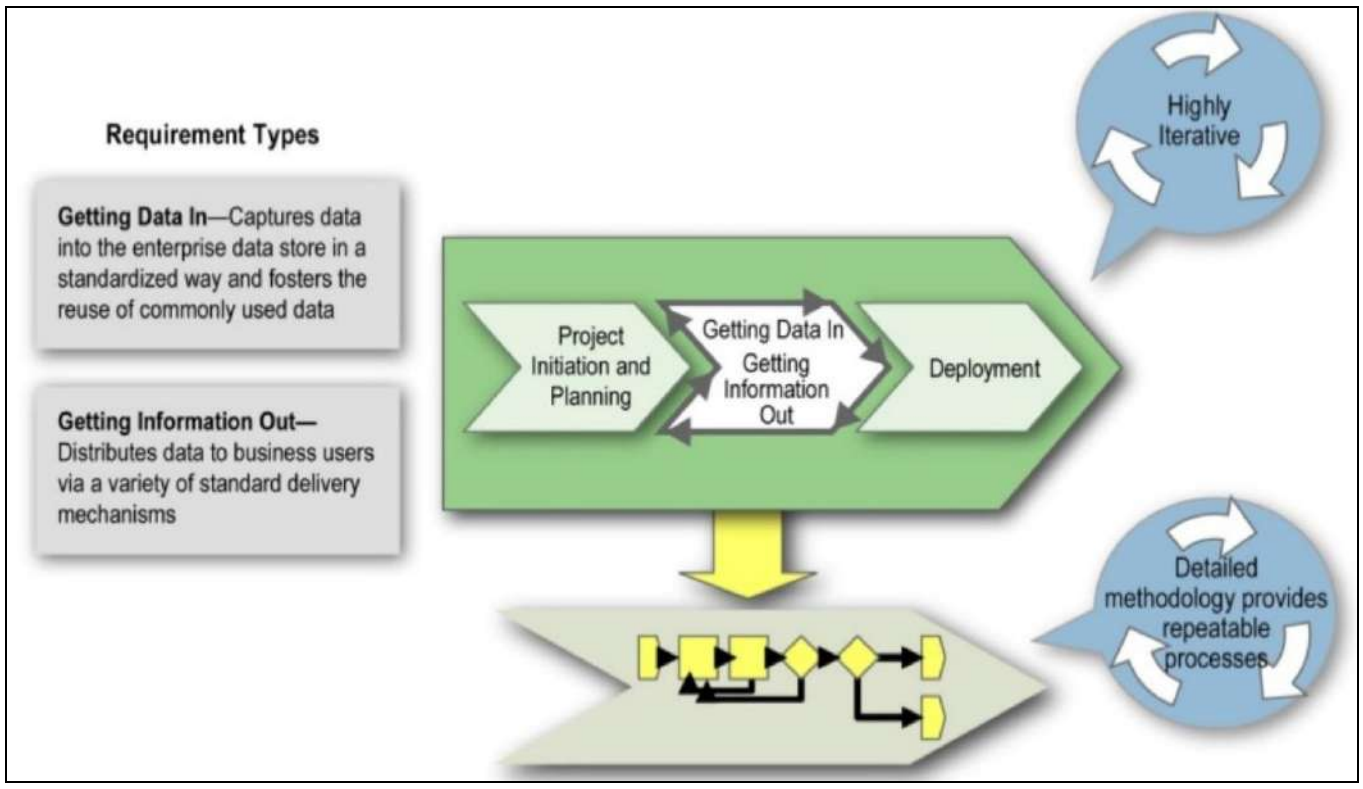
**Data Use**

A third responsibility of the data stewards is to determine and communicate appropriate use of the data, to ensure that appropriate security measures are in place, and to understand how the information is actually being applied. It is in this last area that there are significant opportunities to increase the value of the business intelligence environment through reuse. For example, if a business analyst applies the information in an innovative way, the stewards, working in conjunction with the business intelligence team, should be cognizant of the potential for replicating the application. Once the opportunity is identified, the team can analyze the application and "productionalize" it so that others can gain the same benefits without needing to reinvent the wheel.

The data stewards have other responsibilities, several of which promote consistent information use and application of repeatable processes.

**Methodology**

By its nature, a methodology is a series of interrelated processes, as shown in Figure 2. These processes address the work to be performed, the deliverables, the dependencies, and the project participants. The value of the methodology is apparent right from the start.

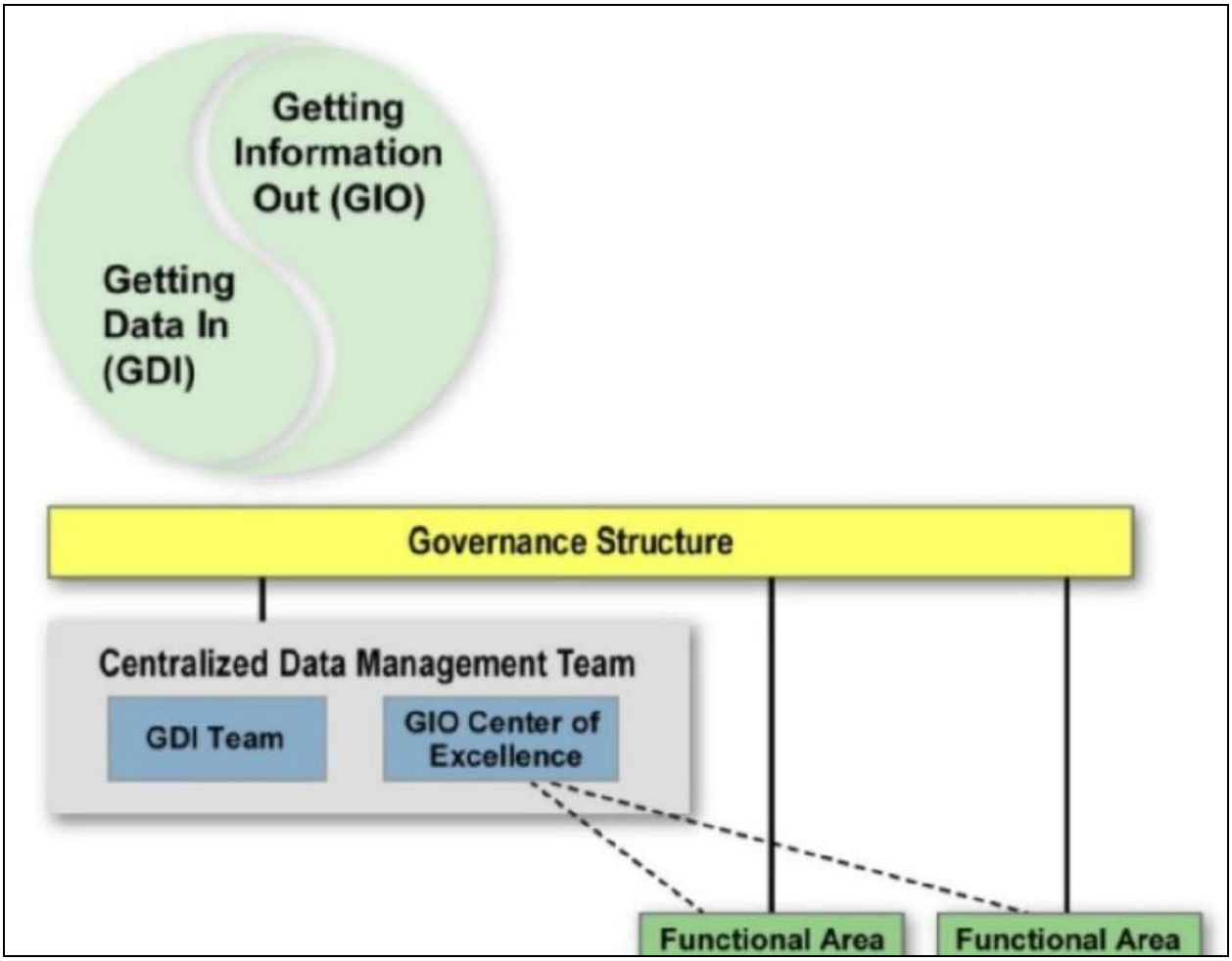



At the beginning of each project, the project manager develops the project plan. With a methodology, the manager begins this activity with a plan template that includes all of the project activities and deliverables. The manager then tailors it to the specific project.

Each step in the project plan should be established as a repeatable process. With repeatable processes, examples of best practices can be accumulated, templates can be created, and training programs established.

**Governance Structure**

The governance structure leverages specific skills to build the enterprise view and enables the functional area teams to focus on information delivery. Figure 3 provides a potential governance structure.




The Governance Structure includes oversight through a cross-functional group that provides direction and oversight to the program. Responsibilities of this group include sanctioning and enforcing guiding principles and policies and enforcing compliance, establishing priorities, allocating funding and resources, promoting the program, and resolving conflicts. Three sets of working teams, under the auspices of the governance structure, carry out the activities needed to deploy business intelligence in the organization.

The Getting Data In Team (GDI) is responsible for gathering the information requirements, understanding the source data quality, developing and maintaining the data models and enterprise data store schema, and designing and developing the ETL processes.

The Getting Information Out (GIO) Center of Excellence is responsible for building the data marts, promoting GIO best practices, coordinating GIO activities performed within functional areas, coordinating GIO education, designing and developing ETL processes for aggregating data from the enterprise data store to the data marts, and facilitating cross area query capabilities.

The Functional Area Getting Information out Teams leverage the data marts and delivery vehicles, gather requirements for information delivery needs, design and build functional area query capabilities and reports, and disseminate appropriate tools, education, training, and best practice information within the functional areas.

**Conclusion**

Business Intelligence programs should leverage repeatable processes and reuse. The architecture should be envisioned based on long term needs but deployed incrementally to provide business value every few months. The data model is developed at the entity level initially, with details added to support each incremental delivery. The methodology uses repeatable processes and supports the iterative development approach. Finally, a governing body promotes the philosophy and guidelines for the program and ensures that business intelligence activities throughout the organization adhere to these.

**References**

1. Azoff M, Charlesworth I. The New Business Intelligence. A European Perspective, Butler Group, White Paper, 2004.
2. Brooking A. Intellectual Capital. Core Assets for the Third Millennium Enterprise, International Thomson Business Press, London, 1996.
3. Casado E. Expanding Business Intelligence Power with System Dynamics, in Raisinghani, M. (Ed.), Business Intelligence in the Digital Economy: Opportunities, Limitations and Risks,Idea Group Publishing, Hershey, PA, 2004.
4. Big Data Analytics with R and Hadoop Vignesh Prajapati, Packt Publishing, 1st edition, 2013.
5. Thota S. Big Data Quality. Encyclopedia of Big Data, 2017, 1-5. https://link.springer.com/referenceworkentry/10.1007/978-3-319-32001-4_240-1
6. Machine Learning with R Brett Lantz, Packt Publishing, 1st edition, 2013.
7. Hadoop For Dummies Dirk deRoos, Paul C.

Zikopoulos, Bruce Brown, Rafael Coss, and Roman B. Melnyk, John Wiley & Sons, Inc., 1st edition, 2014.
8. Cottrill K. Turning Competitive Intelligence into Business Knowledge, Journal of Business Strategy. 1998; 19(4):27-30.
9. Davison L. Measuring Competitive Intelligence Effectiveness: Insights from the Advertising Industry, Competitive Intelligence Review. 2001; 12(4).
10. Degraeve Z, Labro E, Roodhooft F. Total Cost of Ownership Purchasing of a Service: The Case of Airline Selection at Alcatel Bell, European Journal of Operational Research. 2004; 156(1):23-40.
11. Fuld L. New Competitor Intelligence. The Complete Resource for Finding, analyzing, and Using Information about Your Competitors, Wiley, New York, 1995.
12. Gartz U. Enterprise Information Management, in Raisinghani, M. (Ed.), Business Intelligence in the Digital Economy: Opportunities, Limitations and Risks, Idea Group Publishing, Hershey, PA, 2004.
13. Ghoshal S, Kim SK. Building Effective Intelligence Systems for Competitive Advantage, Sloan Management Review. 1986; 28(1):49-58.
14. Gilad B, Gilad T. SMR Forum: Business Intelligence - The Quiet Revolution, Sloan Management Review, 1986; 27(4):53-61.
15. Hannula M. Expedient Total Productivity Measurement, Acta Polytechnica Scandinavica, Industrial Management and Business Administration Series, 1999, 1.
16. Hannula M, Pirttimäki V. Business Intelligence - Empirical Study on the Top 50 Finnish Companies, Journal of American Academy of Business, Cambridge, 2003; 2(2):593-599.