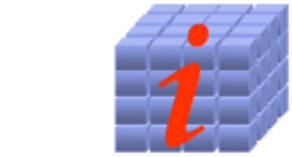

Skills and Mindsets of Data Warehousing Consultants

It is often said that hiring is not hard, but hiring the right candidate is. This is most relevant to a data warehousing and business intelligence environment. To successfully manage, architect, develop, and implement your enterprise-wide reporting solution, you need candidates with great technical and communication skills and, most important, the right mindset for data warehousing and business intelligence.

A data warehousing project succeeds or fails based not only on the integration tools, technology, and database management systems, but also on whether the people working on the project capture the true spirit of such an application or miss it by miles. Data warehousing is a very subjective matter based on who is designing it. As the manager of a data warehousing and business intelligence project, it is on your shoulders to assemble the right team. It will be the collective performance of this team that will create a great data warehousing application.

Before you start recruiting such talent, it is a good idea to know whom you will need on your team, what roles they will play, and what their responsibilities will be. As the old expression says, "If you don't know where you are going, any road will take you there." Based on our experience, the most critical resources required for a data warehousing projects are as follows:

- Data Warehouse Project Manager and Architect
- Data Modeler
- Extract, Transport, and Load (ETL) Developer
- Query and Report Developer

Data Warehouse Project Manager and Architect

The secret to great data warehouse project management and architecture is communication and planning skills, overall understanding of the big picture, and the mindset of treating all users as the ultimate customers of the data. The project manager should provide users with what they are looking for and at the same time educate them about what they can have. This is a delicate situation, but the project manager's mindset should be focused on delivering as well as setting up the right expectations.

The system development life cycle (SDLC) and project management between a data warehouse project and an operational project may have some overlap, but many tasks and deliverables are specific and unique to a data warehouse. While project management is critical for operational projects, it is absolutely essential for a data warehouse, and use of a proper methodology can mean the difference between success and failure. A successful project manager should be able to efficiently:

- Define objectives and purpose of the data warehouse
- Define and implement a road map
- Prepare business cases for the data warehouse
- Lead formulation of work, staffing, and quality plans
- Initiate and oversee assimilation of all project staff members
- Coordinate efforts of all four teams: Extract Transport and Load (ETL), architecture, Decision Support Systems (DSS), and business analysis
- Assist in the measurement of project success against budget, original scope, and business objectives
- Review all deliverables documentation before submittal and participate in quality reviews
- Conduct subject area gap analysis
- Understand business requirements and educate customers to articulate them
- Translate company's "vision" into an integrated solution
- Assert "best practices" and choose the right methodology
- Lead information requirements analysis
- Perform role of a business analyst, data modeler, and/or ETL developer, if required

Data Modeler
A data warehouse can be a competitive advantage dream or a costly nightmare. The difference lies in the way data models have been designed. Data modeling for a warehouse is significantly different than modeling for an operational system. An operational system is optimized for data storage whereas a data warehouse is optimized for data retrieval. A data modeler should clearly understand this difference and design the data models accordingly. The data modeler should:

- Contribute to the information requirements analysis
- Standardize data modeling naming convention, guidelines, and best practices
- Create logical and physical data models
- Generate all database objects and define dimensions, hierarchies, and measures
- Contribute to post-implementation assessments and determine future requirements
- Understand ETL requirements
- Identify data sources and perform gap analysis
- Determine and develop business rules
- Determine criteria for what data belongs in a data warehouse
- Define subject partitions for data warehouse
- Define the role of normalization and, more important, denormalization
- Use star modeling techniques to define dimensions and facts

ETL Developer
According to Wikipedia, "Garbage In, Garbage Out" is an aphorism in the field of computer science. It refers to the fact that computers, unlike humans, will unquestioningly process the most nonsensical of input data and produce nonsensical output. Very often, large corporations go through extensive business intelligence, data integration, and other data consolidation initiatives only to experience low-quality data output. The Data Warehousing Institute estimates that data-quality problems cost U.S. businesses more than $600 billion a year.

The ETL developer is a key resource on the team who is responsible for extracting, transporting loading, mapping, cleansing, scrubbing, and integrating data from disparate source systems into an enterprise data warehouse. This resource must have a mindset of explorers and adventurers. The ETL developer must understand the data warehouse data models and should document the source columns mappings and business rules that need to be applied. The skill set possessed by this resource includes:

- Ability to understand the big picture of the architecture
- Understanding of dimensions and fact tables
- Knowledge of source systems and data columns
- Comprehensive knowledge of business rules required for data transformation
- Understanding of summarization and data aggregation requirements

Query and Report Developer
In data warehousing, seeing is believing. The data residing in a data warehouse is not of much use unless it is presented to users in a format they want with the accessibility and availability they desire.

The query and report developer must understand the reports users are expecting and must be able to speak users' language. More than any other resources on the team, the report developer plays a significant role in ensuring the success of the data warehouse. For this reason, the query and report developer should have a mindset of understanding the business terminology and technical know-how of converting the requirements into great-looking reports. This resource should:

- Comprehend the reporting requirements
- Be creative in presenting the data to users
- Explain gaps in data to users
- Design and build queries and reports
- Design and build end user layer