ITERATIONS
THE DW METHODOLOGY

ITERATIONSTM - The Data Warehouse Methodology
by Prism Solutions

Acknowledge the Challenges
Data warehouse projects pose a unique set of analysis, design, technology and management challenges. These challenges are unlike those found during traditional development projects to build operational systems. Successful data warehouse implementations require a methodology that identifies and addresses these differences to efficiently guide project managers and project teams. More than applying a specific set of technologies, successful data warehouse implementations are the result of an effective and repeatable development process.

Evaluating traditional system development life cycle (SDLC) methodologies, business process re-engineering (BPR) methodologies, or rapid application design (RAD) methodologies for approaches to solve the challenges of data warehousing is an exercise in futility. Regardless how many boxes or arrows or levels they may have, they are primarily geared toward developing operational systems that run the business, rather than informational systems that analyze the business. Not only are the analysis and design characteristics unique to data warehouse development absent among them, but worse, concepts that appear useful in a traditional methodology may be contrary to what an experienced data warehouse architect would recommend. The arbitrary use of common methodologies and accepted techniques focused on building operational systems may be largely to blame for the less than optimal data warehouses within many organizations.

ITERATIONS removes the guesswork from data warehouse planning, analysis, design and management by providing development teams with a clearly defined set of Tasks, timeframes, deliverables, and Roles that can be tailored to your data warehouse initiative.

The Evolution of ITERATIONS
Prism Consulting has pursued the best data warehouse talent to staff its consulting practice with most of the consultants bringing experience from some of the “early adopters” of data warehousing throughout the US. Leveraging the years of experience related to the specifics and intricacies of delivering successful data warehouse solutions, Prism has developed a flexible, iterative project roadmap that can be adapted to various data warehouse initiatives across all industries. ITERATIONS is based on many of the tenets articulated by Bill Inmon and is also focused on ensuring successful implementations for a wide range and scope of data warehouse initiatives including data marts, operational data stores, and enterprise data warehouses. This enhanced methodology addresses all facets of a data warehouse initiative:

- Project management
- Data warehouse analysis
- Data warehouse modeling
- Data warehouse design
- End user access design
- Meta data management
- Technical environment design and preparation
- Construction of the data warehouse environment
- Testing of the data warehouse environment
- End user acceptance
- Project reviews
ITERATIONS was not developed in a laboratory or derived by adapting traditional systems development methodologies. Rather, Prism capitalized on its five-year history of delivering “real” data warehouse solutions, the experience of our consulting staff, and the validated data warehousing theories of Bill Inmon, as the sole sources of input for the methodology. The Prism methodology has continually evolved into what is now a practical, adaptable, comprehensive approach for enabling successful data warehouse implementations.

A review of ITERATIONS is a good starting point for discussing Prism’s approach to any type of data warehouse consulting engagement. A brief review of the methodology should provide an appreciation for how Prism Solutions’ experience and knowledge can help to ensure the success of your data warehouse initiative.

**ITERATIONS Framework**

ITERATIONS was designed in a highly structured manner. This approach yields several benefits including ease of understanding, enabling of a common vernacular within the data warehouse project team, and the modular design permits it to be straightforwardly integrated with existing project methodologies.

ITERATIONS is clearly organized into Modules, and all of the Activities required to complete the Module are identified along with their corresponding Deliverable(s). Additional detail at the Task level is provided in the Generic Project Plan. Modules are grouped two ways - within a Track, and within a Phase. The Tracks represent distinct sets of Modules which should occur in parallel. Phases represent a progressive grouping of Modules. Typically, these Modules are completed prior to initiating the next Phase. Projects are managed and monitored by Phase. In this way, clients may enlist ITERATIONS-certified consultants to assist with a particular Track or Phase of Modules, or even specific Modules themselves. Further, these data warehouse development Modules may be easily inserted into or integrated with existing methodologies.
ITERATIONS Tracks
Five parallel Tracks of work efforts help coordinate various Activities, optimize resource utilization, and maximize data warehouse project efficiencies resulting in more timely implementations.

<table>
<thead>
<tr>
<th>TRACK</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project</td>
<td>Modules focused on orientation, commitment, management, administration, training, marketing and strategy</td>
</tr>
<tr>
<td>User</td>
<td>Modules focused on business requirements, departmental/individual designs, end user access, and end user acceptance</td>
</tr>
<tr>
<td>Data</td>
<td>Modules focused on data model analysis and design, atomic level design, source system analysis, data extraction design, data warehouse processing design, construction, and data warehouse population</td>
</tr>
<tr>
<td>Technical</td>
<td>Modules focused on technical component assessment, selection, integration, technical environment sizing, environment preparation, and environment testing</td>
</tr>
<tr>
<td>Meta Data</td>
<td>Modules focused on business/technical meta data integration, and meta data access design &amp; development</td>
</tr>
</tbody>
</table>

ITERATIONS Phases
The Phases of ITERATIONS represent groupings of Modules that are completed in concert and often have many dependencies among Modules within the Phase and from previous Phases. The conclusion of a Phase should result in a checkpoint that reviews and ensures successful completion of expected project deliverables and ongoing progress of the data warehouse initiative.

<table>
<thead>
<tr>
<th>PHASE</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Startup</td>
<td>Ensures the organization is prepared for a data warehouse project in terms of awareness and commitment, and establishes that a data warehouse strategy is in place</td>
</tr>
<tr>
<td>DW Management</td>
<td>Ensures and plans for training, support, project management, change management, data warehouse marketing, and ongoing administration of the data warehouse</td>
</tr>
<tr>
<td>Analysis</td>
<td>Assesses, scopes and models potential data warehouse solutions, source system solutions, data availability, cleanliness, and completeness, along with high level technical recommendations</td>
</tr>
<tr>
<td>Design</td>
<td>Designs the data environment, data access environment, data extraction environment, maintenance processing environment, and the detailed technical environment</td>
</tr>
<tr>
<td>Construction</td>
<td>Builds and unit tests the data extraction solutions, data access solutions, maintenance processing solutions, the technical environment and develops end user training</td>
</tr>
<tr>
<td>Testing</td>
<td>Performs various levels of integrated data warehouse testing and user acceptance</td>
</tr>
<tr>
<td>Implementation</td>
<td>Makes the data warehouse accessible to the end users and being monitoring and optimization efforts</td>
</tr>
</tbody>
</table>
ITERATIONS Roles
ITERATIONS defines and describes the data warehouse project team Roles and responsibilities and cross references these Roles to the Deliverables Matrix, the Module and Activity Narratives, and the Generic Project Plan.

Data Warehouse Roles

![Data Warehouse Roles Diagram]
ITERATIONS Modules

ITERATIONS is comprised of 35 Modules, each with associated, specific Activities. Each Module is reflected in the comprehensive Generic Data Warehouse Project Plan. Both the Modules and their corresponding Activities have significant Narratives describing them, identifying responsibilities, deliverables, and techniques that can be used to complete them (See Example Section to follow). The following high level diagram illustrates the ITERATIONS process.

The ITERATIONS process

ITERATIONS Product Materials

The ITERATIONS product, available in a set of binders and hyper-linked CD-ROM, includes over 1000 pages of integrated data warehousing best-practices techniques and consists of the following materials:

- **Process Narratives** – detailed descriptions of the nearly 200 high-level and mid-level work items. Each process Narrative consists of 8 sections:
  - Purpose
  - Key Participants
  - Description
  - Effort
  - Deliverables
  - Techniques
  - Responsibility
  - Success Indicators

- **Role Definitions** – detailed skill definitions defining the Core and Extended team resources’ responsibilities.

- **Documentation Templates** – defining each of the suggested nearly 200 project deliverables for creating a solid project audit trail and reference for your future data warehouse releases. The templates are provided both on-line in MS Office and in printed format and included sample completed templates.
• **Deliverables Checklist** – a cross-reference of Documentation Templates to each of the Modules & Activities for project managers to use as a completion checklist and reference. The checklist is provided in MS Excel format with tabs for each Phase.

• **Generic Data Warehouse Workplan** – a detailed hierarchical, flexible data warehouse project plan that lays out each of the process steps, team member responsibilities, and deliverable milestones. The project plan consists of approximately 200 work items at the Activity level, and 700 work items at the more granular Task level. The workplan is provided both on-line in MS Project and in printed format.

• **Tech Topics** – a growing series of about 30 whitepapers on various data warehousing topics written by Prism experts. These Tech Topics are cross-referenced throughout the process Narratives.

• **Training and Case Study** – workbooks and exercise materials from the formal three-day ITERATIONS education class.

**ITERATIONS Licensing**
For approximately the cost of one person-month of consulting, you can now significantly increase the likelihood of building data warehouses that:

- Solve a specific business challenge
- Deliver value within a reasonable timeframe
- Achieve high return on investment
- Meet or exceed expectations
- Meet user requirements
- Deliver a data warehouse solution on schedule, within budget, and effectively utilizing the resources available
- Minimize the impact on operational systems
- Maximize information availability and analytical capabilities
- Design toward flexibility to ensure future decision support needs can be accommodated

Depending upon the scope of your implementation, you can select a “Mart License”, “Departmental License”, or “Enterprise License”.

Activity Definition Example
Following is an excerpt from just one of the nearly 200 ITERATIONS Narratives:

**ACTIVITY A3.2**
**Assess Technical Environment Characteristics of Candidate Source Systems**

- **Phase:** Analysis
- **Module:** Source System Analysis
- **Track:** Data

**Purpose**
Perform analyses on the technical environments of candidate systems of record.

**Description**
In most organizations, there are multiple candidate source systems with the same or comparable data. Candidate source systems are the ones whose data is most likely to be loaded into the data warehouse. The focus of this Activity is to identify the technical system characteristics of similar or identical data within the same or different candidate systems that indicate which is a preferable or optimal source for populating the data warehouse. At this point, all potential source systems have been identified. Most of this effort is spent on the most likely candidate systems. These candidate sources are evaluated based on the following characteristics of their data (data quality analysis is performed in A3.3, Evaluate Quality of Legacy Data):

- Timeliness
- Nearness to the source
- Degrees of granularity
- Batch windows

Methods used to gather this information are:

- Observation of data entry into legacy system
- System staff interviews
- Queries against source systems
- Source code analysis
- Network path analysis
- Source system capacity analysis
- Source system support considerations
- Future implementation considerations

Systems of record that are serious contenders as sources to the data warehouse should be evaluated in depth. Comprehensive technical environment evaluation of the candidate source systems provides a means for selecting between systems that may be able to provide similar or identical information to the data warehouse. For selected candidate systems, the Data Quality Administrator needs to evaluate the quality of source system data by completing the next Activity, A3.3.

Regardless of the best operational data for the data warehouse, technical limitations of operational systems may prohibit them from being acceptable or usable sources of data. This Activity also assesses the feasibility of the current computing environment of a candidate source providing data to the data warehouse environment. For example, limitations that may prohibit a system from optimally supporting the data needs of the data warehouse include, but are not limited to:

- Accessibility
- Capacity
- Availability
- Support personnel

In addition, for the initial implementation of the data warehouse, this Activity serves to increase understanding of the organization's overall computing environment.

**Deliverables**

The Data Warehouse Data Architect completes the following template during this Source System Analysis Activity:

**ACTIVITY &nbsp; Filename Deliverable**

A3.2 &nbsp; A32t.doc &nbsp; Candidate Source Systems Assessment

**Responsibility**

Data Warehouse Data Architect

**Key Participants**

- Source System Expert
- Source System DBA
- Data Acquisition Developer

**Effort**

The number of source systems, their geographic dispersion, and their complexity can affect the time needed to complete this Activity. Additionally, availability of source system support personnel who understand the system and can discuss source system characteristics will have a significant impact on duration. As a guideline, this Activity will require up to one week per major system (such as a billing system or a financial system) and one to three days per minor system (such as a system populated from a major system with a minimal number of tables).

**Techniques**

Using the information obtained in the previous Activity, gather additional information to assist during analysis of the most likely candidates, to include:

- *Processing Narratives*
- *Existing Reports* supporting the business requirements identified during Business Requirements Analysis

Meet with the appropriate System Managers, Subject Matter Experts, Key System Users and Application Support Analysts to discuss system specifics, such as:

- *Required fields*
- *Options for field extraction*

Identify the point in the system processes in which the fields exist in the nearest state to those required in the data warehouse. For example, in a billing system, the data may be required after it has been rated versus
before. However, should usage information be required, extracting the data prior to rating may be advantageous.

- Estimated data volumes
- Extractions already developed against the data

It must be determined why these extracts exist, what current production systems they support, what their stability and level of flexibility are (i.e., can they be changed within reasonable timeframe), support levels, and so forth.

- Timing of extractions, including batch windows

Identify the specific point in batch processing in which the extraction should occur.

Using tools currently available, obtain sample data to verify the data quality, stability, definition, structure, integrity and usability.

**General Guidelines for Selecting the Optimal Source System**

After selecting the optimal source system, it is a good idea to update the template, A3_2t.doc, Candidate Source System Assessment. The optimal source system should be determined based on the following criteria:

**Timeliness**

Accuracy and completeness of data is often relative to the point in time the data is extracted from the current operational environment. Extraction during specific times in a process may prove one source more accurate and complete than another, yet this may not hold true of the same source at another time in the process. Identify and understand the different time implications of the data in the existing systems environment, then identify the best source of data to satisfy the business requirements of the data warehouse.

Examples of considerations for timeliness include:

- How often the source data is created, updated and deleted
- Relative timeliness of sources to be combined in the data warehouse. For example, if the customer master file is updated weekly while the sales master file is updated daily, this can result in sales with no apparent customer when a new customer places an order.

**Structural Compatibility**

While accuracy, completeness and timeliness are very important in deciding on the best data, how well the source data structurally conforms to the data warehouse data model is also an issue. Structural compatibility or conformance occurs when the source system data can be extracted and mapped into the data warehouse table structure with minimal creation of new keys or redesign of the data warehouse data model. If there is a gross mismatch between the data model and the optimal current system data, some compromises are usually made. In the next Activity, A3.3, the Data Quality Administrator performs this evaluation in more detail.

**Nearness to the Source**

Based on the theory that data becomes increasingly corrupt as it passes through an organization's computing environment, consider how far removed the data is from its originating source. Often, the further from the source, the more modifications have been made to the data, and the more difficult it will be to maintain as changes occur to the source system environment. On the other hand, sometimes the best system of record is the output of a downstream process. For example, an output file of an operational system that is carefully controlled, audited and loaded into the General Ledger may be a more appropriate source of data for the data warehouse than creating a separate extract process from the original source that mimics the same results.

**Other Techniques**

Following are techniques and considerations that aid the source system selection:

**Degrees of Granularity**

Similar source data tables may consist of differing degrees of granularity. For example, one sales table may contain line item detail while another contains order summary data. Identifying the differences and business rules applied to both candidate sources will help facilitate the selection of the best system of record.

**Observation of Entry into System**

If time allows, observing the Activities of users of the operational systems at the point of entry is an excellent method for identifying the system's data characteristics.

**System Staff Interviews**

Developing a set of questions concerning source system data and characteristics that can be used when interviewing data creators, data users, data stewards, and data base administrators is often effective.
Network Path Analysis
Accessibility of various source platforms (via local area or wide area networks) often differs significantly because of different network connectivity solutions. Identifying the network paths to be used, and testing the network performance for moving data from the candidate source system platform to the data warehouse platform can be advantageous during the assessment process. In addition, future plans for network upgrades should be identified.

Source System Capacity Analysis
Source system capacity (including processing, storage, and batch windows) is often a major limitation that will affect decisions surrounding the optimal source. Often, preferred source systems have the lowest available capacity because the data resident in the system is in high demand throughout the organization. In an attempt to ensure the capacity is sufficient, time and effort should be expended:

- Evaluating and testing the capacities for compiling and executing data extraction programs
- Investigating the batch queues and priorities are available to the data warehouse project
- Identifying whether acceptable data staging disk space is available
- Assessing plans for hardware upgrades or retirements

Batch Windows
Typically, source extract applications process during the late evening or early morning when operational activity is lowest. For the likely source systems, identify the available batch windows and ascertain whether the extractions can complete processing during these windows. At this time also identify the scheduling anomalies (e.g., scheduled down-time, developmental freeze periods, or scheduled purges/archival).

Source System Support Considerations
The availability and commitment of source system staff to work with the Data Warehouse Project Team can also have an effect on the selection of data sources. Since the source system staff will be relied upon to manage and monitor job queues, assist in coding job control programs, establish system access priorities, and other environment-specific Activities, their commitment is critical.

Future Implementation Considerations
As the data warehouse continues to incorporate new data elements in successive implementations, consider future requirements when selecting appropriate data sources. Although two data sources may satisfy the target data requirements in a relatively equal manner, one may also include data that will likely be needed for the next anticipated subject area.

Success Indicators
Success is achieved upon completing and documenting a comprehensive assessment of the candidate source systems. Success will also be demonstrated, when based on this assessment, minimal reconsideration of additional data sources in this (or a future) implementation is required.

Upon implementation of the data warehouse, this Activity will have been successful when extract performance is satisfactory, the data extracted from the systems is of sufficient quality, and additional data can be included rapidly.

Deliverable Example
Following is an excerpt from the deliverable template corresponding to the Narrative above.

<table>
<thead>
<tr>
<th>Candidate Source Systems Assessment</th>
<th>A3_2t.doc Analysis Phase Source System Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author:</td>
<td>Create Date:</td>
</tr>
<tr>
<td> </td>
<td> </td>
</tr>
<tr>
<td>Review Date:</td>
<td> </td>
</tr>
<tr>
<td>Complete Date:</td>
<td> </td>
</tr>
</tbody>
</table>
**Overview**

| Individual Expected to Complete Template: | Data Warehouse Data Architect |

**Intended Use:** Identify and document the characteristics of the potential source systems. These characteristics will then be used to select the “best” source of data for the data warehouse.

---

**Potential Source System Identification**

<table>
<thead>
<tr>
<th>Potential Source System:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Structure:</td>
<td></td>
</tr>
<tr>
<td>Primary Function:</td>
<td></td>
</tr>
<tr>
<td>Contact Names:</td>
<td></td>
</tr>
</tbody>
</table>

---

**System Profile**

<table>
<thead>
<tr>
<th>Operating System:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform:</td>
<td></td>
</tr>
<tr>
<td>Typical Programming Language Used:</td>
<td></td>
</tr>
</tbody>
</table>

**System Accessibility and Availability**

<table>
<thead>
<tr>
<th>Batch Window Availability:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Batch Processing Capacity:</td>
<td></td>
</tr>
<tr>
<td>Physical Proximity to Expected Data Warehouse Platform:</td>
<td></td>
</tr>
<tr>
<td>Data Storage Capacity:</td>
<td>(temporary files)</td>
</tr>
<tr>
<td>Communications Protocol:</td>
<td></td>
</tr>
<tr>
<td>Communications Availability:</td>
<td></td>
</tr>
<tr>
<td>Support Resources:</td>
<td></td>
</tr>
<tr>
<td>System Stability:</td>
<td></td>
</tr>
<tr>
<td>Security:</td>
<td></td>
</tr>
<tr>
<td>Enhancement &amp; Upgrade Plans:</td>
<td></td>
</tr>
<tr>
<td>Ease of Access:</td>
<td></td>
</tr>
</tbody>
</table>

---

**Data by Source System**

Develop a rating scheme that reflects the priorities of the organization to evaluate the data within the potential source system according to major considerations identified below.
"Acceptable to meet data warehouse requirements?"

**Project Plan Example**

The Iterations Data Warehouse Workplan contains four distinct levels of detail: Phase, Module, Activity and Task – each level expanding upon the previous one. The plan also incorporates the complete set of suggested deliverables (as milestones), and provides Role assignments at both the Module and Activity level. Finally, the plan includes dependencies (links) at both the Module and Activity level.

A small portion of the entire detailed ITERATIONS project plan is illustrated on the following page:

**ITERATIONS Detailed Project Plan (partial)**

Items in blue italics represent suggested Deliverables and their completion dates (documentation templates) for this Module. In practice, project managers often elect to modify the plan to manage projects at the
Activity level rather than the detailed Task level and/or prefer to Track deliverables via the Deliverables Matrix only.

**ITERATIONS Deliverables Matrix Example**

Project managers and other team members may also make free-form use of the on-line Deliverables Matrix that cross-references Deliverables to their associated Activities and Modules. There is one matrix for each Phase. A partial matrix is illustrated below:

Sample Deliverables Matrix (partial)

**ITERATIONS Education**

**ITERATIONS Data Warehouse Methodology (4 day course)**

The ITERATIONS Data Warehouse Development Methodology provides the data warehouse project team a set of guidelines and techniques to enable the development of a successful data warehouse. ITERATIONS addresses the unique project management, architectures, design approaches, technologies, and analytical techniques necessary to develop a successful data warehouse within a reasonable timeframe. Using these methods, data warehouse development teams can greatly reduce the risks of failing to meet user requirements and management expectations. Based on Prism Consulting’s years of experience, the ITERATIONS education class packages and presents these "best practices" to the class participant.

The ITERATIONS training course applies a case study approach to ensure the class participant understands the practical application of the methodology in a realistic environment. It involves discussions of topics such as:

- Iterative and parallel project management
- Negotiated data warehouse design approaches
- Cooperative analysis efforts
- Data warehouse implementation and feedback mechanisms
- Data warehouse orientation, commitment and expectation management
- Project budgeting
- Data warehouse marketing and support
- Meta data collection/integration/management
- End user access development

The course presents the materials in an easy-to-follow fashion, consistently correlating the Module or Activity with the associated Phase in the iterative process.

Also, to ensure maximum value, the course is highly participatory and discussion-oriented, and class participants have ample opportunity to frame their own organization’s culture, standards and technologies within the ITERATIONS methodology.

**ITERATIONS Adaptation (2 consulting days - Optional)**

Because ITERATIONS will be utilized in a very specific cultural, technical, and organizational environment, some aspects of the methodology may need to be tailored to the specific environment. In this two day (minimum) session, an ITERATIONS-certified consultant will work with your project manager or project
team to define how ITERATIONS will be implemented in your organization and on your project. Activities include, but may not be limited to, review of the ITERATIONS deliverables, review of the Roles and responsibilities and correlating them with the skills and capabilities of your project team, and high level development of a project plan.

A lengthier but similar effort may include integrating ITERATIONS with your existing methodology. In this situation, the consultant will work to either extend your current methodology with the applicable Modules or Activities within ITERATIONS, or incorporate necessary components of your methodology into ITERATIONS.

ITERATIONS-Based Consulting Services

Accomplishing the objectives of a data warehousing initiative and moving beyond may require an experienced hand throughout the project, or in particular Phases. Prism Solutions Consulting brings a wealth of experience and expertise to the initiative, offering both a data warehouse specific methodology and industry-recognized experts in data warehousing.

In general, data warehouse projects pose a unique set of analysis, design and management challenges, which are very unlike traditional development projects. Combine these data warehousing challenges with the information architecture challenges that an organization faces, and the advantages become clear of enlisting the services of consultants who possesses a significant breadth and depth of data warehouse development skills.

Prism Solutions’ consulting staff is comprised of internationally recognized data warehouse professionals with several levels of experience. Clients will be assured the appropriate consultant will be matched to the project demands. Consulting services are available for all types of data warehouse design, development and implementation projects, with specific expertise available in the following areas.

Data Warehouse Readiness Assessment

Moving forward to build a data warehouse requires a great degree of technical and organizational preparedness. In this limited engagement, Prism will draw upon its experience to evaluate a client’s readiness to proceed with a data warehouse initiative. Each major component of the ITERATIONS data warehouse development methodology will be addressed (based on the client’s business requirements and objectives.) A Senior Consultant will review project plans, system documentation and project requirements and incorporate their findings in the formal Readiness Assessment Report.

Data Warehouse Enterprise Strategic Planning

Companies are often anxious to analyze how the data warehouse can benefit their entire organization. Typically, they select to first implement an initial data warehouse to evaluate it’s broader potential. Prism’s experience in delivering data warehouse solutions bridging departmental boundaries across industries can be applied to the strategic planning process. Prism will perform a detailed data warehouse implementation review and work with your management to generate a feasible, refined enterprise-wide data warehouse strategy and implementation plan.

Data Warehouse Project Management

Managing the construction of a data warehouse can demand a level of expertise and attention not currently available in many organizations. For this reason, customers may wish to entrust the project management to an outside resource who is familiar with the various development Phases and issues related to data warehouse implementation. Prism will manage all or part of the data warehouse project including: work planning, work supervision, status report preparation, leading design reviews, end user training, and other related Activities as specified by the client.

Data Warehouse Source System Analysis

Organizations often have several similar or identical data sources for populating the data warehouse. One of the first steps in building a data warehouse is to identify the legacy data most appropriate for populating the data warehouse. By assessing the integrity, volatility and accessibility of a client’s data sources, Prism will
produce a recommendation of the types of transformation, filtering, integration, summarization and retention needed to construct the data warehouse according to specifications.

**Data Warehouse Modeling**

The point of departure for building the data warehouse is the data model. The data model serves as the blueprint for organizing the structure and content of data and metadata in the warehouse. Often existing corporate data models must be transformed and extended for use as the data warehouse data model. Prism’s expertise in data modeling techniques and tools can be applied to creating corporate data models, designing data warehouse data models, or adapting Prism’s Inmon Generic Data Models to specific requirements. This Activity involves working closely with the end user and Information Systems communities to ensure the data model is designed to meet their needs.

**Data Warehouse Business Subject Area Analysis**

The overall success of the data warehouse project can be contingent on the selection of the initial business subject areas. Experienced Prism consultants work with clients to identify the subject areas to be populated in the data warehouse and the method in which they should be implemented. The output from this Activity is an understanding of the scope of the effort required for each subject area, so the development stages of the data warehouse are properly estimated and planned for. For each selected subject area, the analysis involves choosing the triggering events for data capture, identifying data relationships, determining naming conventions and planning the frequency of data transfer into the data.

**Data Warehouse Systems Planning**

Constructing a data warehouse involves much more than the movement of data from one platform to another. Information system planning encompasses several Activities needed for building effective decision support systems. This process typically includes building a data model for the data warehouse, determining the optimal computing environment for the data warehouse, selecting the triggering events for data capture, identifying data relationships, resolving naming conventions and planning the frequency of data transfer into the data warehouse.

**Data Warehouse Construction**

Creating the physical data warehouse involves extracting, transforming, transporting, and loading the data. The development of these processes demands a specialized understanding of data warehouse tools and techniques across hardware platforms, operating systems, DBMSs, and networks. Prism data warehouse architects and specialists specify the data mappings and transformations, and develop, test, and maintain the programs.

**Data Warehouse Capacity Planning**

A significant cost of the data warehouse is the computing platform. Accurate capacity planning can help to mitigate this expense. Further, the related financial decisions and capital expenditures should be made as early as possible in the data warehouse project to ensure adequate time for installation. The capacity planning engagement systematically evaluates a client’s data warehouse requirements and preliminary designs to determine the amount of disk storage, processing resources and telecommunication capacity required for the data warehouse environment. Prism’s experience in data warehousing can be drawn upon to overcome the challenges of capacity planning for decision support systems in which the computing workload fluctuates and is difficult to predict. Also, since the data warehouse typically grows in stages over a period of several years, it is critical to project the impact of that growth on corporate resources. Capacity planning is also beneficial for companies who may wish to transfer the data warehouse over time from a mainframe to a distributed environment.

**Data Warehouse Design Review**

Only properly designed and implemented data warehouses will meet end users’ informational needs. A Prism design review will verify that the business requirements and organizational objectives can be achieved via the data warehouse as designed. After each major subject area has been designed, it is studied thoroughly in relation to the corporate data model, the data warehouse data model, the data delivery mechanisms, and the business requirements. This ensures design completeness, accuracy and feasibility. Having an experienced
consultancy such as Prism lead the design review offers the advantages of a broad, external perspective and qualified, constructive input.

**Meta Data Management**
With the abundance of information available to business analysts and other data warehouse users, it is paramount to consider the meaning behind the data. It’s easy to interpret the same information in different ways or questions how the data was derived — this can be a dangerous byproduct of improved information access. Meta data, or data about data is the answer. In helping construct a usable data warehouse solution, Prism will design and deliver a corresponding meta data management solution. This will integrate both the business and technical meta data and provide an access mechanism to search and view the meta data. In addition, this solution will even allow users to launch queries into the data warehouse via one of many popular data access tools.

**Data Warehouse Access Development**
Data warehouse projects driven by business analysis or user access requirements demand complimentary desktop solutions. Different user communities may require unique views of the data warehouse or they may need sophisticated analysis and reporting functionality. Prism Solution’s consulting experience extends into the end user access design and development Activities that precede a successful data warehouse rollout. We will consider not only the informational requirements, but also the data presentation requirements in helping evaluate data access tools, design the internal indexing structures, develop specific reports and queries, train users, and install the software.

Web: http://www.prismsolutions.com/ White paper provided by:
http://www.prismsolutions.com/