



Building A Better Business Intelligence Blueprint

Imagine constructing a skyscraper without a blueprint.

No real structure could ever be built without the aid of a complete and comprehensive plan. The Design Phase of a Business Intelligence project creates an essential “blueprint” that instructs developers how to construct the deliverables specified in the Requirements Phase.

REVIEWING REQUIREMENTS

In any discussion of the Design Phase it is important to review the artifacts that were the outputs from the preceding Requirements Phase.

Data Objects Analysis Matrix (DOAM)

The DOAM documents each data element specified in existing or proposed reports and queries. In this document, data elements are defined and traced back to the source tables from which they originate. Validation of the DOAM by the stakeholders is a clear indication of agreement that the requirements for the project have been properly captured.

Cardinality Diagram

The Cardinality Diagram illustrates the one-to-one, one-to-many, or many-to-many relationships in the data. Understanding these relationships is an important prerequisite to the Design Phase since erroneous assumptions about cardinality will result in a flawed design. As a test of the cardinality, basic joins are made with sample source data validate the assumptions in the diagram.

Preliminary Star Schema

A first cut at the data architecture is an important output of the requirements phase. It gives the Design and Development team an idea of what the possible architecture might be for the data mart to be constructed. This preliminary design is then used to create prototype reports that serve as the final Requirements Phase artifacts.

Bus Matrix

The Bus Matrix is an artifact described in the writings of Business Intelligence guru Ralph Kimball. This artifact ensures that the Star Schema is complete as possible given the available data. In this artifact the subject areas to be analyzed are listed in a matrix along with data “slicers”. For example, a sales subject area can be sliced by date, sales rep, marketing campaign, geography, or customer type.

Prototype Reports

In order to finalize requirements, simple report prototypes are built. These reports are an easy way for stakeholders to visualize requirements and validate that the project is on track to produce the expected deliverables.

CUTTING OVER TO THE DESIGN PHASE

Once requirements have been signed-off, the Design Phase can begin. “Continue” may be a better word than “begin” since several activities from the Requirements Phase could legitimately be described as design-related. This illustrates how the BI creation process is iterative and overlapping in all of its phases. For example, gaps in requirements are frequently revealed in the Design Phase such that certain requirements need to be revised or added before design can be completed.

There are several outputs of the design process. Some are further iterations of Requirements Phase artifacts while others are created within the Design Phase.

Finalized Star Schema and Cardinality Diagrams

Often the data validation steps from the Requirements Phase uncover serious problems. Examples are data joins that don't work, data quality issues, or cardinality assumptions that are disproved. The Data Architect must then make changes to eliminate or work around these problems. Once these have been rectified, the Star Schema and the Cardinality Diagram can be finalized.

A COMPLETE COMPREHENSIVE PLAN

The point of these early testing steps is to discover and correct as many issues as possible prior to the Development Phase when problem resolution is much more costly and time-consuming than in earlier phases.

Aggregation Specifications

Quick response time is standard requirement for any BI implementation. A powerful tool in speeding up queries is an aggregation table. In situations where certain roll ups of the data are commonly used, the summarization is handled in the ETL and used to populate special aggregation tables. This eliminates the need for the reporting tool to do the summarization, thereby producing faster results.

During the Requirements Phase prototyping step, aggregation opportunities (and necessities) often become apparent. However, a thorough analysis of aggregation demands use of the largest practical volume of sample source data. The purpose of large sample data volumes is to approximate as closely as possible performance in the production environment. The time taken in the Design Phase to shave seconds off queries will yield great benefits in user productivity and satisfaction.

Source- to-Target

The essential Extract/Transform/Load (ETL) process needs a roadmap which is provided by the Source-to-Target (STT) document. Evolving out of the Data Objects Analysis Matrix, the STT first takes the data identified and defined in the DOAM and then specifies the needed transformations to the data (the “T” in ETL). Transformations are such things as data cleansing, calculations, or aggregations.

The final ETL step is for the transformed data to be loaded to target tables in the star schema. The STT identifies a specific target table and column for each data element as well as the frequency at which the target tables are to be loaded.

Data Dictionary

All data elements in the Star Schema are defined in the Data Dictionary. This allows users and developers to have a shared understanding of the meaning of each element.

The “single source of truth” goal of for the data mart is supported by the Data Dictionary. For example, when reporting on “profit”, a common understanding of the underlying calculation of that measure is essential. Otherwise, users may misinterpret reports.

Even organizations that have comprehensive documentation of their source system data, the definitions are not always easily understood by business users. The BI Data Dictionary is written in terms that business users will understand. This is critical since the Data Dictionary entries are typically used as metadata in the reporting layer. For example, in some reporting tools, when a cursor is placed over a data element on the report interface the definition of the element appears in a “text bubble”.

Reporting Specifications

Armed with the final Star Schema, the requirements can then be translated into specifications for reports. A typical step in this process is the design of a metadata layer for the reporting tool (e.g. – Cognos Framework Manager or Business Objects Universe).

Test Plan

The Design Phase is an ideal time to craft a test plan for the data mart and reporting interface. Designers are working very closely with the requirements and can more easily document specific test cases for reports or queries. Where ad hoc queries are required, users and developers collaborate to construct detailed scenarios that will test the ability of these queries to return the desired results.

DESIGN FOR SMOOTH DEVELOPMENT

Just as a thorough Requirements Phase makes design steps easier, producing a comprehensive and complete blueprint for the BI project in the Design Phase reduces risk in the subsequent Development Phase. As mentioned above, fixing problems in the development is relatively resource-intensive and causes delays at a time when stakeholders are especially anxious to start using the new reporting capabilities. Often the ultimate success of a project is determined by the quality of the outputs from the Design Phase.



Author

Pete Keers,

Director, Implementation Services



Delivering Business/IT Strategy, Reporting Solutions, and Managed Services that drive measurable business improvement. Unlike conventional services that emphasize solely on technology platforms or consulting, OnApproach bridges the gap, providing clients with practical, analytical, process based solutions that improve decision making and drive business performance.

3455 Plymouth Blvd., Suite 200 • Plymouth, MN 55447 • Phone: 763-557-7118 • E-mail: info@onapproach.net

www.onapproach.net

© 2009 Copyright. All right reserved. OnApproach is a registered trademark of OnApproach. All other brand names are trademarks of their respective owners. 9/15/09