



The Devil You Know: Integrating Old Reports Into New BI Systems

Claudia wishes to thank Mike Biensen for his contributions to this month's column.

As you might expect, there are fundamental issues in getting people to think in terms of receiving and reviewing data, as opposed to reports. Historically, reports defined the user/IT relationship. They were what IT departments were asked to provide and what business users expected to receive. This was important in the past because the development of reports was a laborious process that required time, budget and technical resources. Reports were static in nature and difficult to change; therefore, a significant amount of effort was expended in report design to get as close as possible to meeting the business requirement. Today, reports remain the best tangible evidence of the data that business users need to make critical decisions.

However, depending solely upon existing reports, rather than on the data, often results in some inherent problems when it comes to enhancing business intelligence (BI) systems. Let's look at a few of these problems.

Effort: A significant amount of time and effort is required to pore through hundreds, perhaps thousands, of existing reports – each possibly containing redundant and conflicting data – to find the unique data items that need to be in the data warehouse. Much of the time and effort may be spent on data elements in reports that are rarely used in actual decision making. They are simply holdovers from previous requests for data that are no longer needed or even valid.

Meaning: The analyst struggles to understand the meaning of each report data element, even those that seem to need no explanation. For example, “net sales” may appear on numerous reports such as sales, marketing and finance reports. The analyst must determine the algorithm

used for its calculation and its definition for the sales report, and then determine if these are the same for the marketing report, the finance report and so on. Then on to the next element – “customer.” If “customer” appears in the reports, is this ship-to, bill-to, sold-to, payer or parent customer? In many cases, even the person creating or those using the reports may not know or may have an incorrect interpretation of the data elements. Then, there are other fields with more nebulous names such as quantity or count. Resolving these questions requires extensive investigation, perhaps even returning to the original programs producing the data for the reports. At the end of the day, all you have are somewhat better documented existing reports. However, you may not be much closer to getting to the actual data requirements for your warehouse than you were before.

It is likely that old meanings no longer apply in the new BI environment; but once a report has been accepted for review by the BI team, the expectation from the client is that all information on the report will be provided using just the same definitions and calculations – even if the information on the old report was incorrect. Like shooting a rifle with a bent sight, the client may have learned to run the business with the old report. When presented with a corrected report, the client may miss the target because the bias applied to the old report, while no longer needed, is still being applied.

Applicability: An existing report may, in reality, not be what the requestor actually wants. Perhaps it was originally developed as a stopgap measure or its utility over the years has diminished because of operational or business changes. The business user may have adapted to it, perhaps by supplementing it with spreadsheet manipulations or other personal data inputs. This additional and very meaningful

step is not likely to be discovered with just an analysis of reports. Applicability is especially relevant when a data warehouse is being developed in conjunction with new operational systems or ERP implementations. Business process changes resulting from the new systems usually make some existing reports obsolete. New information requirements needed by the new processes are not captured; and, therefore, the warehouse misses the mark.

Facilitating the New: In today's changing business environment, new questions are the rule rather than the exception. Your focus should be on providing new capabilities or methods for businesspeople to interact with the data, slicing and dicing, analyzing and manipulating it in their own ways to gain the insight necessary to make sound business decisions. The key to enabling this is to determine the real data requirements needed to answer the key business questions – not simply replicate old reports. Let's look at some techniques that work well to determine the real data requirements for your data warehouse.

First, you need to weed out all reports that have little or no relevance in today's decision making. The business users, in conjunction with those knowledgeable about the new environment, must step up to the plate to identify and prioritize both the reports and the fields which are really of use to them. This should eliminate a significant amount of effort spent analyzing elements that no one wants or needs.

Of even greater importance is education for the business community (and IT implementers!) to focus on the business requirements for the needed data, not just on the creation of reports. The techniques for doing this include methods such as business process and data modeling, facilitated sessions, analysis of business users' MBOs or other job performance criteria.

Facilitated sessions in which the business users describe why they need certain pieces of data and how that data is used in their everyday decision making yield tremendous insight into what is absolutely mandatory in the data warehouse as opposed to data that is simply “nice to have.” An implementation team that tries to supply “everything” to the business users will not only have a high probability of failing, they will most likely create an environment that is, at best, difficult to use. Everything is not really what the business community wants because much of this flood of data will have minimal business interest or even understanding. A business analyst trying to gain insight from an overpopulated data warehouse struggles to find kernels of usable data amongst the clutter of data items having little or no business meaning or value.

Use facilitated sessions or one-on-one interviews to help the implementers develop an understanding of business objectives and drivers expected from the BI environment. Important KPIs and other information needed to support strategic business decision making can be determined. An added benefit is to get the businesspeople to think outside of the box (e.g., perhaps they should consider external information sources maybe not available in their current environment today that could be effectively combined with internal data to provide a richer information environment). Simply punting by

requiring the implementers to examine and replicate only existing reports is not an acceptable way of gathering true BI requirements.

In addition to these methods, other good sources of data requirements may include existing data warehouses, data marts and personal systems (a.k.a., spreadmarts). These systems often contain data developed outside of the operational systems that add real business value, especially to higher-level personnel. These include alternative hierarchies, relationships or business rules not documented in the operational systems and attributes not easily captured from external sources. Personal spreadmarts exist everywhere throughout an organization and are created by people who are the key information providers to upper management. They may be the main users of your BI environment.

Shifting the main focus of data gathering from analyzing only existing reports to examining the actual data and reports used for decision making is a paradigm change that may be resisted by many within the organization. These resisters may be concerned about your de-emphasis of reports and develop worst-case scenarios or expectations about losing the reports they do use to run the business. It is necessary to work closely with the business community and especially these resisters to identify the truly critical reports and assure them that resources will be available to recreate the needed reports in the new

environment. Again, education may be the key here to convince the business users that all of their needed data will be available. It is also a paradigm shift to convince these same users that they themselves can create any and all of the reports they deem necessary.

The prevailing and constant message should be that a data warehouse is an information source, not necessarily just a reporting source, and that most data for analysis can be retrieved directly by businesspeople with tools requiring only limited technical ability. Most executives and higher-level personnel are receptive to this when they can be shown how easily information can be directly obtained. Following a presentation concerning data warehouse information access including a demonstration scratch, you may be pleasantly surprised to hear your company president remark to others, “I could do that.” 

Claudia Imhoff, Ph.D., is the president and founder of Intelligent Solutions, a leading consultancy on CRM and business intelligence technologies and strategies. She is popular speaker and internationally recognized expert and serves as an advisor to many corporations, universities, and leading technology companies on these topics. She has coauthored five books and more than 50 articles on these topics. She can be reached at CImhoff@IntelSols.com.

Mike Biensen is currently the business intelligence leader of a major project at Celanese, recently featured in CIO magazine, to move all SAP and other legacy systems to one SAP instance. Benefiting from data warehousing best practices learned prior to his experience with SAP's Business Warehouse, Biensen is implementing this system utilizing a hub-and-spoke architecture. Operational and other source data goes through a data hub where integration and business rule resolution occurs before going to the spokes (e.g., data marts or external systems). He can be reached at mbiensen@celanese.com.