A general information process begins with defining the intelligence problem in terms of users’ requirements. It is beneficial to think in terms of 5Ws in this context—Who, What, When, Where, and Why. For a military intelligence application, responses to the Ws are in terms of foreign threats to national or international security. In business, requirements may be expressed in terms of competitive position in the marketplace in contrast to that of competitors. For an airline (e.g., see Continental Airlines, TUN, and Chapter 5), clear metrics are used in the industry, such as on-time performance, proper baggage handling, overbooking, and so on. Continental could see that its performance on these metrics needed improvement. Hence, it chose to define requirements in terms of these metrics (“What”), meaning that tracking these performance measures over time and by basic business events became a major requirement for the initiative. Other requirements were less transparent and required tracking a customer’s set of trips as important business events. For example, Continental decided that if it targeted a certain segment of its customers by granting certain privileges, offering special incentives, and so on, then it might be able to both attract and retain the type of customer it seeks—particularly the most profitable ones.

An additional requirement had to be stated to address “When.” Some of Continental’s BI initiatives could be done in batch (i.e., under lower time pressure) using historical data—for example, Demand-Driven Dispatch, Good Will Letter, and Group Snoop. Each of these initial applications had a basic expectation (a “Why”) of delivering value with respect to optimizing flight schedules, encouraging more President’s Club membership, and clamping down on travel agent group overbooking, respectively. Other applications had a more time-sensitive requirement, for example, Marketing Insight, A Personal Touch, Elite Access, and Elite Upgrade Monitoring. This time dimension, or essence of the question of “When,” had major implications for the remaining phases of the general process. In the case of time-sensitive applications, the cycle time of collecting data, processing it, analyzing it for inclusion in reports, producing the reports, and disseminating those reports was very important—it had to be done in what was referred to as “right time” in the Continental case.

Another important aspect of the requirements assessment was “Who,” (i.e., for whom do we produce BI reports in order for them to have the most impact on our business objectives for the initiative?). While many reports are very important in terms of supporting managerial decision making, there are circumstances in which the “Who” is not so clear. For example, Continental’s Marketing Insight application provides customer profiles to sales personnel, marketing managers, and flight personnel (e.g., ticket agents, gate agents, flight attendants, and international concierges). It is possible that a Continental flight attendant can offer an apology to a high-value customer after discovering the passenger experienced serious delays on a prior leg. Relevant to the “Who” question, common sentiment was emerging about the importance of engaging many levels of personnel in BI initiative-supported decision making, as follows:

We judge leaders by how well they make big, strategic decisions. But corporate success also depends on how well rank-and-file employees make thousands of
small decisions. Do I give this client a special price? How do I handle this cus-
tomer’s complaint? Should I offer a seat upgrade to this customer? By them-
selves, such daily calls—increasingly made with the help of enterprise decision-
management technology—have little impact on performance. Taken together,
they influence everything from profitability to reputation. (Rohde 2005)

After requirements assessment, the general model of intelligence creation and use
moves to the planning phase. Here a major question relates to the return on invest-
ment (ROI) of a data warehousing initiative and the subsequent and incremental
costs/benefits of conducting individual projects like the BI applications at Continental,
discussed above. A data warehouse represents a major up-front cost, because it implies
a major change to an organization’s infrastructure. However, infrastructure invest-
ments aren’t new to business as there have already been major investments in net-
working, telecommunications, OLTP databases, servers, and so on. Let’s consider this
major up-front investment first; then we will examine the assessment of return on
investment on individual BI projects.

The return on an investment is calculated by a simple formula: ROI = Present
Value of Benefits/Present Value of Costs. For an infrastructure investment, the tricky
part is getting at the present values. A major study summarized by Kelly (2001)
described benefits in three categories. First, there are “Keepers.” Keepers represent
money saved from a data warehouse investment that is currently being paid to create
and disseminate ad hoc decisions supporting reports from disparate source systems
that may have to be painstakingly integrated for each individual report. The IT staff
required to do each report may be redeployed in a data warehouse infrastructure to do
backlogged decision support reports once a data warehouse is in place. A straightfor-
ward calculation follows: IT Staff Savings = (# of Members Redeployed) * (Salary +
Benefits).

The next category in the study suggests looking at the “Gatherers” category. Here,
we are trying to get at the money saved due to the automated collection and dissemi-
nation of information. Managers who are used to receiving multiple reports, con-
structed from legacy decision support initiatives take time to reconcile those reports,
glean what they can from cross-referencing key performance indicator data, and so on.
For this gathering category, the managerial time savings = (# of Managers) * (Salary +
Benefits) * (Percent Time Saved) * (Productivity Rating), where that final rating
should reflect the fact that all time saved isn’t equivalent to the amount of increased
workload that can be assigned.

The third category is “Users.” This category is a bit more fuzzy, as it represents the
money saved or gained by supporting improved decision making. One important
aspect of this measure is information timeliness. For example, if reports can be issued
within a time in which action can be taken to avoid business costs/issues, then an
advantage of a data warehouse infrastructure is the opportunistic decision-making
behavior it can provide. Most ROI calculations for a data warehouse initiative would
relate to faster production of needed reports and fewer costs associated with reconcil-
ing disparate reports, but the quality and impact of decisions made will alter the way
business is conducted throughout an organization. Kelly (2001) reports that typical
ROI contributions of data warehouse projects are 20 percent for Keepers, 30 percent
for Gatherers, and 50 percent for Users. This means that users definitely need to be
involved in data warehouse infrastructure projects, but it doesn’t explain some of the
intangible benefits, such as managing the total customer relationship, adding value to
the customer, empowering all levels of decision makers in the organization, and being
able to manage the big picture as well as specific issues.
It has also been pointed out that cost analysis for a data warehouse investment often misses Total Cost of Ownership (TCO) issues, including the fact that the yearly annual cost of maintaining a data warehouse often runs from 40 to 60 percent of the initial costs (Adelman and Moss 2001). In addition, it is true that many organizations now evaluate the cost of capital for an investment, that is, the return that could be expected if the same investment dollars were placed elsewhere. Payback period (or break-even analysis) and rate of return are two additional measures organizations use. A data warehouse is always implemented in conjunction with an existing relational database management system (RDBMS) that handles OLTP. For this reason, costs and benefits must reflect the price of add-ons to existing infrastructure, and this will vary by RDBMS vendor. For example, some vendors will price by the number of users or even by server. Most products have add-ons that increase price, but some add-ons aren’t necessary, for example, if extract, transform, and load (ETL tools) can be done in-house. Additional tools may also be needed for data cleansing and performance monitoring. There may also be increases in networking infrastructure costs as more demand is placed on these services. The nature of the architecture of the data warehouse will place different demands on the network infrastructure and bandwidth in particular. One mustn’t leave out training costs, help desk/support costs, and system administration costs. Many of these costs are higher initially, and they decline with the time the warehouse is in place. One measure of the impact on decision-making is the potential for cash flow acceleration. In areas like supply chain management, increasing inventory turns, backlog control, more accurate demand forecasting, and so on might be likely sources of cash flow acceleration. This aspect again points out the need for user involvement in estimating the TCO of a data warehouse.


**ONLINE FILE W1.2: BI GOVERNANCE**

*Governance* refers to defining and implementing an infrastructure and strategy that will support the goals of the enterprise. According to Matney and Larson (2004), BI governance builds in flexibility by creating robust processes capable of scaling to any size and scope, and all aspects related to the BI efforts are clearly defined. Along with a mechanism to manage BI strategy, BI governance provides measurements for gauging success. One nontangible benefit: improved morale as the staff involved builds its knowledge base and shortens its learning curve.

Matney and Larson (2004) proposed that BI addresses the following three challenges:

- **Challenge #1:** How Do You Get Your Arms Around Your Current BI Environment?
- **Challenge #2:** How Do You Get Your Arms Around Your Current BI Strategy?
- **Challenge #3:** How Do You Get Your Arms Around BI Value?

By answering these challenges, BI can reach maturity and a company can then utilize it to its advantage. For example, Harrah’s (TUN) has become a leader in gaming due to its customer loyalty program that is supported by DW and BI analysis.
Matney and Larson (2004) also identified the following four components of BI governance:

- **BI governance committee** (for project direction, alignment, and prioritization)
- **BI life cycle** (to ensure consistent project delivery and meet end-user expectations)
- **End-user support** (ongoing improvements to user effectiveness and empowerment)
- **BI review** (evaluating maturity levels and tracking metrics)

They viewed the BI life cycle as composed of six cyclical steps: data integration, analytics, custom tool work, user acceptance, training, and tool selection. This cycle includes all the processes needed to implement BI initiatives.


**ONLINE FILE W1.3: THE BI USER COMMUNITY**

Farmers are typically managers and analysts who are most likely to have responsibilities for financial information (including providing organizational reports on revenues, costs, profit margins, etc.) These managers and analysts likely gather the information to track each product or service, and they may, for example, evaluate the impact of sales promotions or other types of advertising and customer outreach campaigns on sales. Cost-cutting campaigns are another example where farmers will be involved in preparing reports from the DW. A farmer likely sees the organization from such dimensions as product, service, market segment, promotional campaign, or geographic location. They often like to drill down into details of each dimension and compare revenues and costs across dimensions in order to glean as much as possible about succeeding and failing products, stores, campaigns, and/or services. In contrast to Farmers, **Tourists** are often from the executive level of the organization. This tourist is therefore very important because they make decisions about the value of the DW/BI to the organization. Each Tourist may take a somewhat individualized tour of the warehouse. Because they have responsibility for the entire organization, the information they desire will be culled from highly aggregated data. While the more technically savvy Tourist might enjoy making more detailed use of the interfaces to the data warehouse, there is often a need for a standardized set of preplanned executables ready for a Tourist to launch easily, perhaps from a common executive portal or a dashboard-type interface.

Operators are the most common users of the DW, and their requests for information are typically routine, with information needs scheduled on a regular basis. They may need some ad hoc use of the warehouse, but the types of reports they rely upon to do their job are typically standardized. Typical Operators represent an organization’s clerical or administrative staff. They have a tactical focus with perhaps a role as a lower-level manager, product line manager, logistics manager, or customer service representative who needs to assess short-term inventory trends. Note that such Operators could also include front-line employees who work with customers, such as the marketing
managers for an airline (see the Continental case in Chapter 2) who can make use of the warehouse to track high-value clients. The time dimension is usually limited in the information normally provided from the warehouse to an Operator, for they are typically interested in what is happening today or what happened last week or perhaps as long as last month. Most of the interfaces to the warehouse that Operators use are fairly standard, with only several keystrokes necessary to receive the information required. Typical reports will also be standardized so that information can be interpreted rapidly for tactical action. Operators mainly use the reporting and querying capability of OLAP as well as simple statistical analysis (e.g., averages, trends), as will be demonstrated in Chapter 3.

Another user community is represented by the Explorers, occasional, unpredictable users of the data warehouse; they may wonder about a certain hypothesis, seek information to verify or contradict it, dig deeper, or move on to another hypothesis. These people often ask ad hoc questions about the organization’s business that usually haven’t been considered before. In essence, they can be viewed as persons digging for some unknown truth using random and unconventional types of questions. On occasion, an Explorer may discover an important clue about the business that will have a high impact—perhaps even an incredible impact. They may massage a great deal of data in gathering the information they are seeking, often looking for a pattern or some basic truth that has thus far eluded the business. Explorers are the classic nonconformists in their thinking and resultant use of the data warehouse, and they use mostly data, text, and Web-mining tools.

In contrast to explorers, Miners are often ruthlessly meticulous as they slice and dice data to glean significant insights. They are looking for the “nuggets” in the data gold mine. Miners prefer large amounts of detailed historical data as they look for hypothesis-confirming patterns, but unlike the Explorer, they typically have a mission in mind. They might be looking for an instance of fraud, characteristics of the most profitable customers, or some other detail that links a successful outcome with a specific market segment or customer profile. These folks need to discover correlations—the insights that can help the organization be more profitable. The data warehouse will be very heavily accessed by the Miners, and the BI tools to support them are among the most sophisticated, ranging from predictive analysis (Chapter 3) to data mining. Miners who seek to classify are looking for ways to cluster customers, products, services, or some other meaningful business entity. Similarly, they may be seeking to classify some business entity according to some set of criteria. A predictive mining mission might provide a model that can be used to establish a norm or rule of behavior expected from a customer, product, or service. That model can then be used in making or supporting decisions when, for example, new customers possess characteristics similar to those of prior customers; that information can be used to target an ad, upsell, and so on. Mining is difficult, and therefore Miners must have patience and persistence. This category of BI user may well include statisticians, supply chain specialists, and actuaries.

With respect to the up-front assessment of current capabilities, the IS organization (department) that is likely to play an important role in realizing a DW must be carefully examined. In using the term, IS organization, we are referring to the group that is primarily responsible for the company’s computing infrastructure, the day-to-day activities of keeping information systems in operation, and the computer and information technology specialists who may be assigned to the DW and BI project. Important questions in assessing the IS organization include:

- Does the IS organization understand the need for and potential of BI in achieving the company strategy?
- Does the IS organization have the requisite skills and resources to execute the BI requirements?
- Will the IS organization be able to take responsibility for getting the data warehouse up and running?
- Do the people in the company believe that the IS organization can take the lead in supporting change throughout the enterprise?
- Is the IS organization respected enough within the company to achieve the level of change that will be required?

From the perspective of the potential user communities, the following questions are important:

- Do the business users understand why BI is needed and what its potential might be?
- Based on their history, will business users be willing to fund and champion the DW initiative as well as the specific BI projects that make use of the warehouse?
- Has the business user community achieved a historical capability to push the IS organization to get what the community needs?
- Based on recent history, can the business user community be active partners with the IS organization to pull off a project of the necessary scale?
- Will members of the user community be willing to join in efforts to select appropriate technologies and see to it they are properly deployed?

Several alternative scenarios are often discovered as a result of this organizational assessment, each having suggested corrective approaches. First, if the company’s culture is not amenable to the changes likely to be brought on through the BI initiative, there is a high likelihood of project failure. If this is the case, and a major strategic advantage or company survival is at stake, then a considerable effort must be made to educate the organization, including the IS division or department, the potential user communities, and management. If the culture is not amenable to change even with an investment in education, then it may be best to initiate BI-type projects on an ad hoc, case-by-case basis in order to begin building a base of solid successes. A small-scale pilot study—without a full-blown DW investment—might be an appropriate first step. If the IS organization is alone in pushing for a DW and BI initiative, and top management believes it imperative for strategic reasons, then there is a need to promote it to the anticipated, but wary, user communities. Reasons for a lack of interest on the part of the user community could be prior failures of the IS organization, a lack of understanding of the criticality of the initiative with respect to the strategy, no high-level champions of the program in the company, or progress that has stagnated. If any of these reasons are
apparent, but there is a desire to forge ahead, then small pilot projects are recommended, again on an ad hoc basis, with highly successful proof-of-concept projects being widely communicated throughout the company. If no progress can be made—even in projects addressing what might be referred to as “the low-hanging fruit”—then the organization should have in place plans to back away from the initiative.

ONLINE FILE W1.5 TERADATA UNIVERSITY NETWORK (TUN): A PREMIER RESOURCE FOR DATA WAREHOUSING, DSS/BI, AND DATABASE FACULTY

Teradata University Network (teradatauniversitynetwork.com) is a free learning portal designed to help faculty teach, learn, and connect with others in the fields of data warehousing, DSS/BI, and databases. Content-driven and managed by leaders in the field, this portal’s content includes:

1. Teradata software and Teradata business partner software, using an ASP model
2. Course syllabi and course Web sites
3. Cases, projects, and assignments (with teaching notes)
4. Book chapters, articles, and research reports
5. PowerPoint presentations (with speaker notes)
6. Teradata certification information for your students
7. Webinars on timely subjects led by industry leaders
8. The Teradata library and speakers’ bureau
9. Web-based courses
10. Links to related sites

A separate site for students, teradatasudentnetwork.com, contains a subset of the above materials:

1. Teradata software and Teradata business partner software, using an ASP model
2. Cases, projects, and assignments (without teaching notes)
3. Book chapters, articles, and research reports
4. Teradata certification information
5. The Teradata library
6. Webinars on timely subjects led by industry leaders
7. Links to related sites

(This student site excludes materials that are only for course instructors.)

To register for and access the Teradata University Network, go to teradatauniversitynetwork.com and click on Join Now. You will be asked to complete a brief registration form. After your registration is authenticated (to assure that you are a faculty member), you will be able to use this valuable resource.

Teradata University Network members are encouraged to submit content for possible inclusion on the Web site. A review process ensures that all materials on the Teradata University Network are complete, complementary, valuable, and of high quality. Any questions about the network can be directed to Hugh Watson (hwatson@terry.uga.edu) or any of the Advisory Board members (see names at TUN site).

The sponsor of this site, Teradata, a division of NCR, is a leading provider of data warehousing, business intelligence, and data management solutions. Teradata has worked hard with its Advisory Board to create this premier resource for the academic community.