

BI projects are organized according to the same six stages common to every engineering project. Within each engineering stage, certain steps are carried out to see the engineering project through to its completion. Business Intelligence Roadmap describes 16 development steps within these stages, as outlined below.

The Justification Stage

Step 1: Business Case Assessment

The business problem or business opportunity is defined and a BI solution is proposed. Each BI application release should be cost-justified and should clearly define the benefits of either solving a business problem or taking advantage of a business opportunity.

The Planning Stage

Step 2: Enterprise Infrastructure Evaluation

Since BI applications are cross-organizational initiatives, an enterprise infrastructure must be created to support them. Some infrastructure components may already be in place before the first BI project is launched. Other infrastructure components may have to be developed over time as part of the BI projects. An enterprise infrastructure has two components :

1. **Technical infrastructure, which includes hardware, software, middleware, database management systems, operating systems, network components, meta data repositories, utilities, and so on.**
2. Nontechnical infrastructure, which includes meta data standards, data-naming standards, the enterprise logical data model (evolving), methodologies, guidelines, testing procedures, change-control processes, procedures for issues management and dispute resolution, and so on.

Step 3: Project Planning

BI decision-support projects are extremely dynamic. Changes to scope, staff, budget, technology, business representatives, and sponsors can severely impact the success of a project. Therefore, project planning must be detailed, and actual progress must be closely watched and reported .

The Business Analysis Stage

Step 4: Project Requirements Definition

Managing project scope is one of the most difficult tasks on BI decision-support projects. The desire to have everything instantly is difficult to curtail, but curtailing that desire is one of the most important aspects of negotiating the requirements for each deliverable . Project teams should expect these requirements to change throughout the development cycle as the business people learn more about the possibilities and the limitations of BI technology during the project.

Step 5: Data Analysis

The biggest challenge to all BI decision-support projects is the quality of the source data. Bad habits developed over decades are difficult to break, and the damages resulting from bad habits are very expensive, time consuming, and tedious to find and correct. In addition, data analysis in the past was confined to the view of one line of business and was never consolidated or reconciled with other views in the organization. This step takes a significant percentage of the time allotted to the entire project schedule.

Step 6: Application Prototyping

Analysis of the functional deliverables, which used to be called system analysis, is best done through prototyping so it can be combined with application design. New tools and programming languages enable developers to relatively quickly prove or disprove a concept or an idea. Prototyping also allows business people to see the potential and the limits of the technology, which gives them an opportunity to adjust their project requirements and their expectations.

Step 7: Meta Data Repository Analysis

Having more tools means having more technical meta data in addition to the business meta data, which is usually captured in a computer-aided software engineering (CASE) modeling tool. The technical meta data needs to be mapped to the business meta data, and all meta data must be stored in a meta data repository. Meta data repositories can be licensed (bought) or built. In either case, the requirements for what type of meta data to capture and store should be documented in a logical meta model. When licensing a meta data repository product, the requirements documented on this logical meta model should be compared to the vendor's meta model, if one is provided. In addition, the requirements for delivering meta data to the business community have to be analyzed (e.g., online help function).

The Design Stage

Step 8: Database Design

One or more BI target databases will store the business data in detailed or aggregated form, depending on the reporting requirements of the business community. Not all reporting requirements are strategic, and not all of them are multidimensional. The database design schemas must match the information access requirements of the business community.

Step 9: Extract/Transform/Load Design

The ETL process is the most complicated process of the entire BI decision-support project. It is also the least glamorous one. ETL processing windows (batch windows ) are typically small, yet the poor quality of the source data usually requires a lot of time to run the transformation and cleansing programs. Finishing the ETL process within the available batch window is a challenge for most organizations.

Step 10: Meta Data Repository Design

If a meta data repository is licensed, it will most likely have to be enhanced with features that were documented on the logical meta model but are not provided by the product. If a meta data repository is being built, the decision must be made whether the meta data repository database design will be entity-relationship based or object oriented. In either case, the design has to meet the requirements of the logical meta model.

The Construction Stage

Step 11: Extract/Transform/Load Development

Many tools are available for the ETL process, some sophisticated and some simple. Depending on the requirements for data cleansing and data transformation developed during Step 5, Data Analysis, and Step 9, ETL Design, an ETL tool may or may not be the best solution. In either case, preprocessing the data and writing extensions to supplement the capabilities of the ETL tool is frequently required.

Step 12: Application Development

Once the prototyping effort has firmed up the functional requirements, true development of the access and analysis application can begin. Developing the application can be a simple matter of finalizing an operational prototype, or it can be a more involved development effort using different, more robust access and analysis tools. In either case, the front-end application development activities are usually performed in parallel with the activities of back-end ETL development and meta data repository development.

Step 13: Data Mining

Many organizations do not use their BI decision-support environment to the fullest extent. BI applications are often limited to prewritten reports, some of which are not even new types of reports but replacements of old reports . The real payback comes from the information hidden in the organization's data, which can be discovered only with data mining tools.

Step 14: Meta Data Repository Development

If the decision is made to build a meta data repository rather than to license one, a separate team is usually charged with the development process. This becomes a sizable subproject in the overall BI project.

The Deployment Stage

Step 15: Implementation

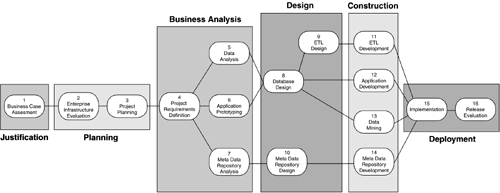
Once the team has thoroughly tested all components of the BI application, the team rolls out the databases and applications. Training is scheduled for the business staff and other stakeholders who will be using the BI application and the meta data repository. The support functions begin, which includes operating the help desk, maintaining the BI target databases, scheduling and running ETL batch jobs, monitoring performance, and tuning databases.

Step 16: Release Evaluation

With an application release concept, it is very important to benefit from lessons learned from the previous projects. Any missed deadlines, cost overruns, disputes, and dispute resolutions should be examined, and process adjustments should be made before the next release begins. Any tools, techniques, guidelines, and processes that were not helpful should be reevaluated and adjusted, possibly even discarded.

You do not need to perform the development steps in sequence; most project teams will likely perform them in parallel. However, because there is a natural order of progression from one engineering stage to another, certain dependencies exist between some of the development steps, as illustrated in Figure 0.6. Steps stacked on top of each other in the diagram can be performed simultaneously , while steps that appear to the right or left of each other are performed relatively linearly (with less overlap) because of their dependencies.

Figure 0.6. Development Step Dependencies



While some development steps are clearly project-specific, most development steps must be performed from a cross-organizational perspective. Thus the focus of those project activities takes on a cross-functional dimension, and the reviewers of those activities should include business representatives from other lines of business. The main task for the business representatives from the other lines of business is to validate and ratify the strategies, policies, business rules, and standards either being used or being developed during the BI project. Table 0.1 indicates which steps are project-specific and which ones are cross-organizational.

Table 0.1. Project-Specific versus Cross-Organizational Steps

| **Development Step** | **Project-Specific versus Cross-Organizational** |
| --- | --- |
| 1. Business Case Assessment | Cross-organizational |
| 2. Enterprise Infrastructure Evaluation (technical and nontechnical) | Cross-organizational |
| 3. Project Planning | Project-specific |
| 4. Project Requirements Definition | Project-specific |
| 5. Data Analysis | Cross-organizational |
| 6. Application Prototyping | Project-specific |
| 7. Meta Data Repository Analysis | Cross-organizational |
| 8. Database Design | Cross-organizational |
| 9. ETL Design | Cross-organizational |
| 10. Meta Data Repository Design | Cross-organizational |
| 11. ETL Development | Cross-organizational |
| 12. Application Development | Project-specific |
| 13. Data Mining | Cross-organizational |
| 14. Meta Data Repository Development | Cross-organizational |
| 15. Implementation | Project-specific |
| 16. Release Evaluation | Cross-organizational |

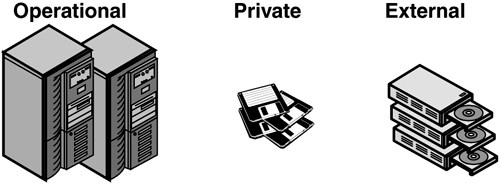
Requirements for:

* Ease of use
* Level of data granularity
* Timeliness
* Data quality
* Security
* Amount of external data
* Historical requirements
* Tool capabilities

# **Business Analysis Issues**

One of the challenges in building a BI decision-support environment is to merge data from different types of data sources. There are three major types of data sources: operational, private, and external (Figure 1.2).

##### Figure 1.2. Three Major Data Sources



##### Operational Data

Online transaction processing (OLTP) and batch systems provide internal operational data about subject areas, such as the following:

* Financial
* Logistics
* Sales
* Order entry
* Personnel
* Billing
* Research and engineering

##### Private Data

This internal departmental data usually comes from the desktops and workstations of business analysts, knowledge workers, statisticians, and managers. Examples include the following:

* Product analysis spreadsheets
* Regional product usage spreadsheets
* Prospective customer databases

##### External Data

Organizations often purchase external data from vendors that specialize in collecting industry-specific information available in the public domain, such as the following:

* Health care statistics
* Customer profile information
* Customer catalog-ordering habits
* Customer credit reports

External data is usually clustered around the following categories:

* Sales and marketing data: lists of prospective customers
* Credit data: individual credit ratings, business viability assessments
* Competitive data: products, services, prices, sales promotions, mergers, takeovers
* Industry data: technology trends, marketing trends, management science, trade information
* Economic data: currency fluctuations, political indicators, interest rate movements, stock and bond prices
* Econometric data: income groups, consumer behavior
* Demographic data: age profiles, population density
* Commodity data: raw material prices
* Psychometric data: consumer profiling
* Meteorological data: weather conditions, rainfall, temperature ( especially for agricultural and travel industries)

#### Source Data Quality

Merging and standardizing data is usually a requirement of every BI application but one that is not so easy to accomplish. One of the difficulties in merging and standardizing data from different types of data sources is that the data is stored in different file structures on different platforms. What makes the process even more difficult is that the keys for the same objects on different data sources usually do not match, the definitions for the same apparent data are often inconsistent, and the values are often missing or conflicting. In addition, different people in the organization have authority to determine business rules and policies for data from different types of data sources, and resolving data conflicts among them or getting clarification is often all but impossible .

Standardizing data from internal operational data sources is difficult enough, but standardizing data from private and external data sources is a major challenge and could be costly. This cost should be calculated and included in the cost-benefit analysis.

# **Cost-Benefit Analysis**

All BI decision-support initiatives should fulfill at least one of the five benefit categories listed below (Figure 1.3).

1. **Revenue increase, possibly in the form of:**
   * Identification of new markets and niches
   * More effective suggestive selling
   * Faster opportunity recognition
   * Faster time to market
2. Profit increase, including possibilities for:
   * Better targeted promotional mailings
   * Early warning of declining markets
   * Identification of under-performing product lines or products
   * Identification of internal inefficiencies
   * More efficient merchandise management
3. Customer satisfaction improvement through:
   * Improved understanding of customer preferences
   * Improved customer-to-product matching
   * Up-selling to customers
   * Increased repeat business
   * Faster resolution of customer complaints
4. Savings increase through:
   * Reduction in wasted or out-of-date merchandise
   * Reduction in requests for customized reporting
5. Market share gain through:
   * Increased numbers of customers who defect from the competition
   * Much higher customer retention rate as compared with previous years and with the competition

##### Figure 1.3. Benefit Categories



In addition to determining ROI, a business case assessment must include an appraisal of risk. Any project is bound to involve some risks and, given the high costs of BI projects, performing a risk assessment is a high priority.

# **Risk Assessment**

Risks are factors or conditions that may jeopardize a project. Risks should be assessed for the following six major variables :

1. **The technology used for implementing the project**
2. The complexity of the capabilities and processes to be implemented
3. The integration of various components and of data
4. The organization and its financial and moral support
5. The project team staff's skills, attitudes, and commitment levels
6. The financial investment in terms of ROI

Table 1.1 depicts a basic risk assessment matrix for these six variables, using the colors of a traffic light to indicate the severity of the risk:

Green = low risk ”go ahead with the project

Yellow = medium risk ”caution, proceed slowly

Red = high risk ”stop, reevaluate before proceeding

Each organization should develop its own appropriate variables and risk conditions for analyzing the risks most likely to impact its BI project. In developing that detailed risk assessment matrix for your organization, expand on the questions listed below.

* Technology risk

- How mature are the selected technologies within the marketplace ?

- How mature are the selected technologies within the organization?

- How many different technologies will co-exist?

- Do we have incompatible operating systems?

- Do we have incompatible database management systems (DBMSs)?

Table 1.1. Basic Risk Assessment Matrix

|  | **Level of Risk** | | |
| --- | --- | --- | --- |
| **Variable** | **Green (Low)** | **Yellow (Medium)** | **Red (High)** |
| Technology | Experienced with mature technology | Minimal experience with technology | New technology, little experience |
| Complexity | Simple, minimal workflow impact | Moderate, some workflow impact | Mission critical, will require extensive reengineering |
| Integration | Stand-alone, no integration | Limited integration required | Extensive integration required |
| Organization | Solid internal support | Supportive to a large extent | Little internal support |
| Project team | Business experience, business-driven, talented, great attitude | Some business experience, business-driven, talented, fair attitude | No business experience, only technology-driven, limited talent, bad attitude |
| Financial Investment | Possible ROI within a very short time | Possible ROI within a moderate time frame | Possible ROI after a few years |

* Complexity risk

- How complex is the overall IT environment?

- How complex is the BI application itself?

- How extensively will workflow have to change? Will it have to be completely reengineered?

- How many sites will be supported?

- What is the degree of distribution of data, processes, and controls?

* Integration risk

- How many interfaces will the BI application have?

- Are there external interfaces?

- How much source data redundancy exists?

- Can the primary keys from various data sources be matched?

- Do we have incompatible standards? No standards?

- Do we have "orphan" records as a result of referential integrity problems?

* Organization risk

- How much risk will business management tolerate ?

- How much risk will IT management tolerate?

- How much financial and moral support can we expect when the project encounters hurdles?

* Project team risk

- How much experience does the team have with successful implementations of BI applications?

- How broadly based is that experience?

- How well balanced is the team?

- How is team morale ?

- How likely is it that we may lose one or more team members ?

- Do our team members' skills cover all the basic disciplines?

- Will the business representative be an active player?

- How strong is the project manager?

* Financial investment risk

- How fast can ROI be expected?

- How likely is it that the costs will outweigh the benefits?

- Can financial risk be mitigated by using only proven technologies?

|  |  |  |
| --- | --- | --- |
| |  |  | | --- | --- | | graphics/hand_icon.gif | The combination of high complexity and greater integration often results in a higher risk of failure to the organization. | |

Expand each of these risk categories with organization-specific detailed variables and detailed conditions for each of the three severity rankings (low, medium, high). Table 1.2 shows an example of a detailed risk assessment matrix taken from a case study.

The managers for the organization in this case study listed the detailed risk variables. Then for each variable, they described the conditions for each of the three risk severity rankings. For example, in the category for business workflow support:

* Low risk = Supports business workflow seamlessly
* Medium risk = Requires some manual intervention
* High risk = Requires significant manual intervention

Table 1.2. Case Study: A Detailed Risk Assessment Matrix

|  | **Level of Risk** | | |
| --- | --- | --- | --- |
| **Variable** | **Green (Low)** | **Yellow (Medium)** | **Red (High)** |
| Project requirements: ad hoc reporting | Supports every critical ad hoc reporting requirement | Supports most critical ad hoc reporting requirements | Fails to support critical ad hoc reporting requirements |
| Project requirements: AS/400 | Supports every key business requirement | Supports most key business requirements | Fails to support key business requirements |
| Business workflow support | Supports business workflow seamlessly | Requires some manual intervention | Requires significant manual intervention |
| Architecture evaluation | Well-architected application | Existence of some architectural issues | Poorly architected application |
| Extensibility into subsequent releases | Fully extensible into subsequent releases | Extensible for most requirements | Not extensible into subsequent releases |
| Logical data model: completeness | All information requirements met | Most information requirements documented | Significantly mis-sing information requirements |
| Logical data model: extensibility | Fully extensible | Some extensibility issues | Not extensible |
| Meta data (business and technical) | Complete and easily maintainable | Incomplete or not easily maintainable | Not incorporated |
| Physical data model: completeness | Complete and tuned | Complete but not tuned | Incomplete, cannot be evaluated |
| Physical data model: extensibility for new product types | Fully extensible for new product types | Limited product type extensibility | Incomplete, cannot be evaluated |
| Physical data model: source system feeds | Acceptable design support for source systems | Performance or timing concerns | Incomplete, cannot be evaluate |
| Interfaces (external and internal) | Supports external and internal interfaces | Limited support for external and internal interfaces | Poor support for external and internal interfaces |
| Analysis dimensions and measures: adding new product lines | Easy to add | Can be added, but requires significant cube reconstruction | Cannot be evaluated at the current time |
| Analysis dimensions and measures: adding new tools for data analysis | Proposed cubes and set of dimensions sufficient to support the business analysts | Proposed cubes and set of dimensions provide minimum sufficiency | Proposed cubes and set of dimensions insufficient |
| Use of meta data repository | Fully developed | Limited meta data support | No meta data support |
| Loading of the BI target databases | Load procedures established and perform well | Load procedures poorly documented or perform poorly | Load procedures not developed, cannot be evaluated |
| Physical database issues | Effective and efficient physical database design | Minor issues with physical database design | Physical database design incomplete, cannot be evaluated |
| Performance issues | Conforms to stated performance requirements | Some performance issues | Cannot be evaluated at this time |
| Systems management issues: maintenance | Support procedures well established and documented | Limited support documentation | No support procedures |
| Support issues | Backup and disaster recovery procedures developed and installed | Backup and disaster recovery procedures developed but not installed | No thought given to backup and disaster recovery procedures |
| Security implementation | Satisfies application needs and is easy to maintain | Difficult to maintain | Security design incomplete, cannot be evaluated |

The managers then selected the applicable risk severity ranking for each variable by highlighting the description that most accurately portrayed the condition of their BI project using the colors green, yellow, and red. Out of 21 variables, they rated only two variables low risk, six variables medium risk, and thirteen variables high risk. The managers decided that the overall risk for this BI project was high.

Having a realistic assessment of the severity of potential risks will help the project team create realistic estimates and expectations for the BI project. Conversely, unidentified and unmanaged risks can result in project failure or even jeopardize the entire BI initiative.

1. **Business case assessment report**

The business case assessment report should document the following:

- Strategic business goals of the organization

- Objectives of the proposed BI application

- Statement of the business need (business problem or business opportunity)

- Explanation of how the BI application will satisfy that need (proposed BI solution)

- Ramifications of not addressing the business need and not committing to the proposed BI solution

- Cost-benefit analysis results

- Risk assessment

- Recommendations for business process improvements to the operational systems or to the operational business processes and procedures

The assessment report should also have a one- or two-page executive overview that summarizes the details of the report.