

tdwi's  
**Best of Business  
Intelligence**  
*A Year in Review*

VOLUME 5

THE VERY BEST OF  
TDWI'S BI ARTICLES,  
RESEARCH, AND  
NEWSLETTERS



# 2007: THE YEAR OF LIVING DANGEROUSLY

**PLUS Forecast: The 2008 Technology Market**

**CONTEMPLATING CONSOLIDATION**

**RESEARCH EXCERPTS: OPERATIONAL BI  
AND BI SOLUTIONS FOR SAP**

**INSIGHTFUL ARTICLES ON:**

- Predictive Analytics
- Tailoring Dashboard Content

**tdwi**  
THE DATA WAREHOUSING INSTITUTE

# Data Integration Concerns?



## Make data transparency a reality with ETI

### Top 5 Reasons Customers Choose ETI Solution® for High Performance Data Integration:

1. Lowest Total Cost Solution – saves time and labor; no need for additional software and hardware
2. High performance and scalability – no limits on data volume or project complexity
3. Native mainframe support – benefit from existing assets; legacy to SOA and back
4. Bi-directional integration – any solution that's not bi-directional is only half a solution
5. Flexible – complements and adapts to your corporate standards and BI environment

**eti**®

High Performance Data Integration

**Click here to download ETI's newest whitepaper:  
"The ROI of Data Transparency"**

To learn more, go to [www.eti.com](http://www.eti.com) or  
call: (512) 383-3000 or (800) 856-8800

## SPONSOR INDEX

Baseline Consulting

DataFlux Corporation,  
a SAS Company

ETI

Information Builders

MicroStrategy

myBiQ LLC

ParAccel

SAND Technology

Syncsort Incorporated

## TABLE OF CONTENTS

### FEATURES

- 5 2007: The Year of Living Dangerously  
**STEPHEN SWOYER**
- 11 Forecast: The 2008 Technology Market  
**WAYNE W. ECKERSON, PHILIP RUSSOM**
- 14 Contemplating Consolidation  
**WAYNE W. ECKERSON**

### 2007 BEST PRACTICES REPORTS

- 17 Best Practices in Operational BI  
**WAYNE W. ECKERSON**
- 23 Business Intelligence Solutions for SAP  
**PHILIP RUSSOM**

### TEN MISTAKES TO AVOID SERIES

- 29 Ten Mistakes to Avoid in Data Quality Management  
**ARKADY MAYDANCHIK**

### TDWI FLASHPOINT

- 35 Reinventing Your BI Program  
**STEVE WILLIAMS**
- 39 Striking a Balance in Agile BI Development  
**EVAN LEVY, FERNANDO MARTINEZ-CAMPOS**

### BUSINESS INTELLIGENCE JOURNAL

- 43 Tailoring Performance Dashboard Content  
**ROSS MORRISSEY**

### BI THIS WEEK

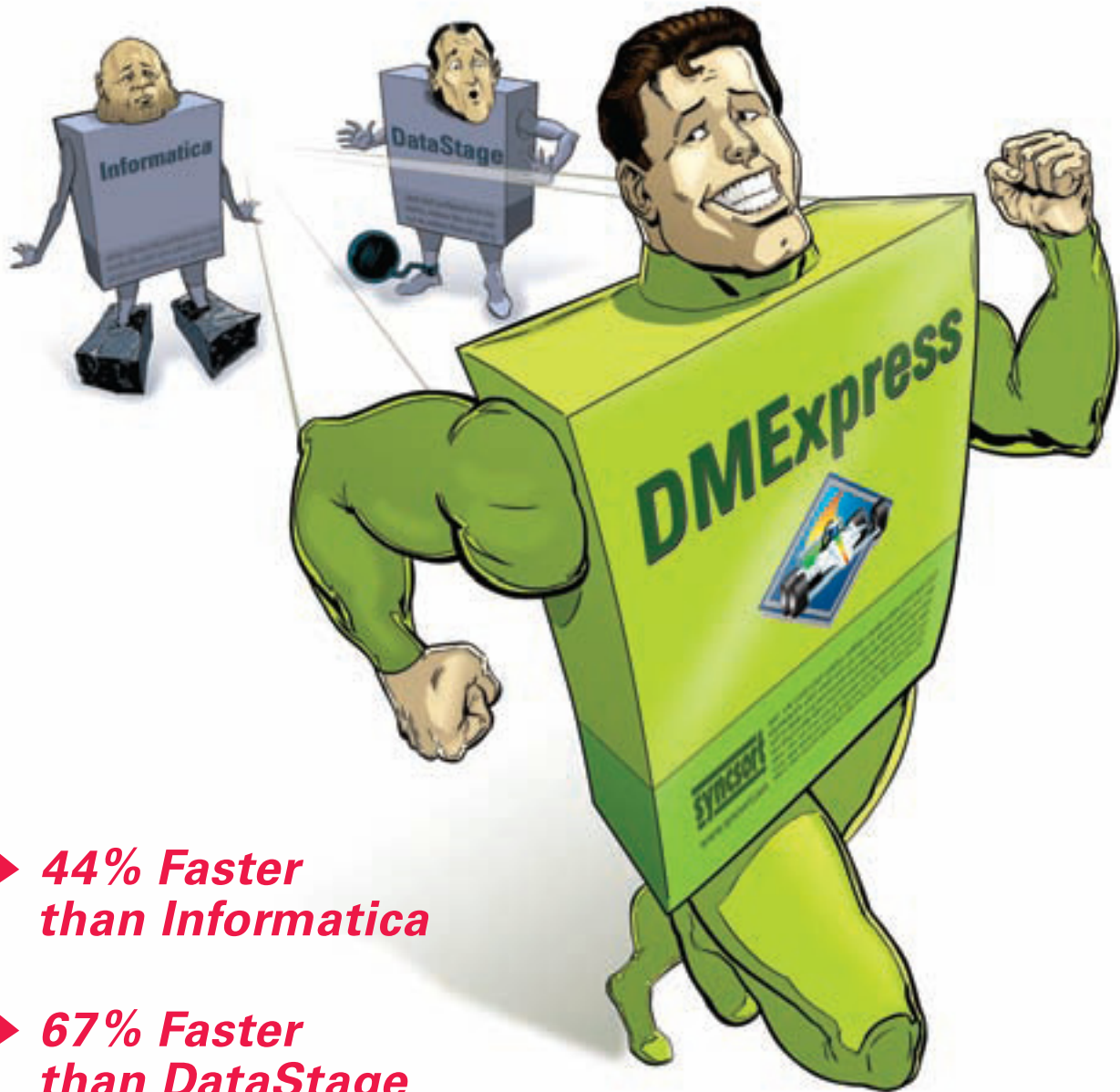
- 53 Predictive Analytics: Slow Adoption Despite Big Benefits  
**STEPHEN SWOYER**

### BI SOLUTIONS

- 56 Transforming Technologies

- 61 **ABOUT TDWI**

# **DMExpress is the Faster ETL!**



▶ **44% Faster than Informatica**

▶ **67% Faster than DataStage**



DMExpress is the high-speed, low-cost ETL tool engineered to speed data warehousing, BI, and other large-volume applications. Simply put, DMExpress can trim many hours, even days off of processing time. For a free consultation on how DMExpress can speed your application visit [www.syncsort.com/bob28ad1](http://www.syncsort.com/bob28ad1)

**syncsort**  
[www.syncsort.com](http://www.syncsort.com)

www.tdwi.org

Editorial Director	Denelle Hanlon
Managing Editor	Jennifer Agee
Production Editor	Susan Stoddard
Art Director	Deirdre Hoffman
Graphic Designer	Bill Grimmer



General Manager	Richard Zbylut
Director of Education	Paul Kautza
Director of Research	Wayne Eckerson
Director of Marketing	Michelle Johnson

## 1105 MEDIA

President & Chief Executive Officer	Neal Vitale
Chief Financial Officer & Senior Vice President	Richard Vitale
Executive Vice President	Michael J. Valenti
President, 1105 Events	Dick Blouin
Vice President, Financial Planning & Analysis	William H. Burgin
Vice President, Finance & Administration	Christopher M. Coates
Vice President, Audience Marketing & Web Operations	Abraham M. Langer
Vice President, Information Technology	Erik A. Lindgren
Vice President, Print & Online Production	Mary Ann Paniccia
Chairman of the Board	Jeffrey S. Klein

### Reaching the staff

Staff may be reached via e-mail, telephone, fax, or mail.

**E-mail:** To e-mail any member of the staff, please use the following form: FirstInitialLastname@1105media.com

**Renton office** (weekdays, 8:30 a.m.–5:00 p.m. PT)  
Telephone 425.277.9126; Fax 425.687.2842  
1201 Monster Road SW, Suite 250, Renton, WA 98057

**Corporate office** (weekdays, 8:30 a.m. – 5:30 p.m. PT)  
Telephone 818.734.1520; Fax 818.734.1528  
9121 Oakdale Avenue, Suite 101, Chatsworth, CA 91311

### Advertising Opportunities

Denelle Hanlon, dhanlon@tdwi.org, 425.277.9130.

**List Rentals:** 1105 Media, Inc., offers numerous e-mail, postal, and telemarketing lists targeting business intelligence and data warehousing professionals, as well as other high-tech markets. For more information, please contact our list manager, Merit Direct, at 914.368.1000 or www.meritdirect.com.

**Reprints and E-prints:** PARS International, 1105reprints@parsintl.com; telephone 212.221.9595; fax 212.221.9195

© Copyright 2008 by 1105 Media, Inc. All rights reserved. Reproductions in whole or part prohibited except by written permission. Mail requests to "Permissions Editor," c/o Best of BI, 2008, 1201 Monster Road SW, Ste. 250, Renton, WA 98057. The information in this magazine has not undergone any formal testing by 1105 Media and is distributed without any warranty expressed or implied. Implementation or use of any information contained herein is the reader's sole responsibility. While the information has been reviewed for accuracy, there is no guarantee that the same or similar results may be achieved in all environments. Technical inaccuracies may result from printing errors, new developments in the industry, and/or changes or enhancements to either hardware or software components. Printed in the USA.

TDWI is a trademark of 1105 Media, Inc. Other product and company names mentioned herein may be trademarks and/or registered trademarks of their respective companies.

## EDITORIAL DIRECTOR'S NOTE

Welcome to the fifth annual *TDWI's Best of Business Intelligence: A Year in Review*. Each year we select a few of TDWI's best, most well-received, impact-full articles, research, and information, and present them to you in this publication.

Stephen Swoyer kicks off this issue with a review of recent major BI developments. "The Year of Living Dangerously" takes a look back at this year of upheaval, discussing gobbled-up BI and performance management pure-play vendors. He observes that 2007 was "the year in which the large, independent, publicly traded, best-of-breed BI and PM player all but ceased to exist (with a few exceptions)."

In "Forecast: The 2008 Technology Market," TDWI Research analysts Wayne Eckerson and Philip Russom bring you some hints about the future of business intelligence and data integration from their TDWI Technology Market reports (an exclusive benefit for TDWI Members).

To further represent TDWI Research, we've provided excerpts from two of the past year's Best Practices Reports. "Best Practices in Operational BI" describes how to move BI out of the back room and embed it into the fabric of the business, delivering insight and information on demand to all workers. "Business Intelligence Solutions for SAP" describes the the issues you need to consider before deciding which mix of tools, platforms, approaches, professional services, and business sponsorship are appropriate to your solution.

In "Tailoring Performance Dashboard Content," our selection from the *Business Intelligence Journal*, Ross Morrissey describes how to design performance dashboards around the capabilities and goals of the users—not the technology provider. He draws lessons from management accounting, aviation, and software interaction design, and takes a detailed look at analysis mindshare across the organization. This volume's *Ten Mistakes to Avoid* will help you avoid some common pitfalls in data quality management. And thanks to columns from TDWI's e-newsletters, you'll learn about reinventing your BI program; agile BI development; and the adoption of predictive analytics.

TDWI is committed to providing industry professionals with information that is educational, enlightening, and immediately applicable. Enjoy, and we look forward to your feedback on the *Best of Business Intelligence*, Volume 5.

Denelle Hanlon  
Editorial Director, *TDWI's Best of Business Intelligence*  
The Data Warehousing Institute  
dhanlon@tdwi.org

# Make Your Data Behave.

Data in your organization arrives in many shapes and sizes – well-disciplined isn't one of them. DataFlux provides the technology that drives data governance initiatives, allowing you to create and enforce business rules for data integrity within your BI and operational applications – and deliver a well-behaved and consistent foundation of enterprise information. We provide the only single platform that can analyze, improve and control your critical data – and help your organization deliver trusted information to support business decisions. Better data equals better business. That's a lesson we can all learn.



Leader in Data Quality  
and Data Integration

[www.dataflux.com](http://www.dataflux.com)  
877-846-FLUX

International  
+44(0) 1753 272 020

# 2007: THE YEAR OF LIVING DANGEROUSLY



BY STEPHEN SWOYER

**It's a tradition. Every winter, we look back at the year just past and take** stock of what transpired (and of what—thankfully—didn't transpire). This tradition even provides a small comfort of sorts, which is most welcome in a year of such upheaval.

Consider 2007, which might ultimately be remembered as The Year of Living Dangerously—for business intelligence (BI) and performance management (PM) pure-play vendors, at any rate. Few could have guessed it, of course, but 2007 marked the year in which the large, independent, publicly traded, best-of-breed BI and PM player all but ceased to exist (with a few exceptions).

First, in late February, Oracle Corp. snapped up the former Hyperion Solutions Corp.; then, eight months later, SAP AG picked up Business Objects SA. The latter acquisition was as big of an October surprise as the industry has ever seen—and don't forget that Oracle once sprung an October surprise of its own, two years ago, when it nabbed archival Siebel Systems Inc.

Canadian BI powerhouse Cognos Inc. was the last of the Big Time Pure-Plays to fall, just last month, to IBM Corp. In retrospect, says Wayne Eckerson, director of TDWI Research, BI's Big Three fell like so many dominoes. "I think the BI vendors have seen for a long time that the ERP vendors and the big database vendors were going to come into this space, so they want to bulk up to prepare themselves, which they did ... through acquisitions," Eckerson remarked just after IBM's acquisition of Cognos. "Ultimately, once one of them went and had a partner with big, deep pockets, they all had to go, so it's kind of a domino effect."



## FEW COULD HAVE GUESSED IT, BUT 2007 MARKED THE YEAR IN WHICH THE LARGE, INDEPENDENT, PUBLICLY TRADED, BEST-OF-BREED BI AND PM PLAYER ALL BUT CEASED TO EXIST.

Nor is it simply a matter of the BI Big Three going away. Collectively, Eckerson notes, Oracle, SAP, IBM, Hyperion, Business Objects SA, and Cognos acquired more than a score of other BI vendors over the last several years. SAP, for example, picked up PM specialist OutlookSoft earlier this year—just months before it nabbed Business Objects—while Business Objects itself nabbed a PM best-of-breed of its own, the former Cartesis. A lot of players have been wiped off the board—leaving a lot of customers in the lurch.

The upshot, experts say, is a drastically altered—and still largely confusing—BI and PM landscape. Just how altered and confusing is anyone's guess.

From a customer perspective, buying habits will both change and remain (more or less) business as usual. Industry veteran Cindi Howson, a principal with BIScorecard.com, breaks customers down into two types: "strategic buyers" (customers who look to buy more software from vendors with whom they have established relationships) and "tactical buyers" (customers who focus on short-term success, features, functions, and ease of deployment).

In the case of the former, Howson says, industry consolidation gives them more incentive than ever to ditch independent pure-play solutions

and standardize, once and for all, on a preferred vendor. In the case of the latter, of course, there's ample opportunity to shop around, play vendors off against one another, and deploy best-of-breed BI technology while at it.

"All these preferred vendors now have [or will have, once their acquisitions close] viable BI products as part of their total software portfolio," writes Howson on her BIScorecard.com blog. "When a preferred vendor has a solid BI solution, the buying approach changes from 'who has the best product for us' to 'why can't we use the product from our preferred vendor?'" Of course, this approach isn't quite as simple as you might think—especially given the prevalence of multiple preferred vendors in most organizations: "The biggest battle for market share here," Howson says, "will be when a customer considers more than one of the big four vendors to be a preferred vendor—so if for example, Microsoft and SAP both have toeholds in the same customer account."

This isn't a totalizing shift, however, Howson concedes. "This is not to say that enterprise customers will only buy from their preferred vendors. It does, however, mean that it will be a tougher sell for other vendors," she argues. "BI pure-plays who fail to differentiate and clearly articulate where, why, and when they are a better investment than the incumbent software vendor, will lose share."

Now more than ever, Howson says, BI buyers must do their homework. "BI buyers who understand the unique value proposition of their desired BI pure-play ... [should] be prepared for a more rigorous cost/benefit analysis when convincing IT management of 'why not the preferred vendor,'" Howson concludes. "The smartest companies will develop clear guidelines of when to use the standard, preferred vendor and when to supplement capabilities with products from another vendor."

### Dashboards and Data Visualization

Dashboards (more precisely, dashboard-driven data visualization technologies) figured large this year. In 2007, it was all about eye candy.

First, at the very beginning of the year, the former Cognos Inc. snapped up the former Celequest Corp. in a multi-billion dollar

deal. Celequest gave Cognos best-of-breed dashboard capabilities as well as a full-fledged dashboard appliance. Several months later, enterprise application integration (EAI) giant TIBCO Software nabbed best-of-breed data visualization specialist Spotfire Inc., augmenting its EAI line-up with Spotfire's data visualization and event-processing capabilities.

Meanwhile, data visualization best-of-breeds (as well as would-be best-of-breeds) continued to innovate: Tableau Software Inc. announced three major revisions or new product introductions in 2007, including a new "dynamic dashboard" technology that exposes its data visualization technology in an interactive dashboard interface. Tableau also delivered an analytic dashboard add-on for Microsoft Dynamics CRM package. Not to be outdone, competitor Advizor Solutions Inc. announced a new software-as-a-service version of its data visualization software (SalesAdvizor for Salesforce.com's AppExchange) earlier this year.

Elsewhere, BI pure-play MicroStrategy Inc. introduced a new Active Dashboards offering that officials billed as a Web 2.0-ified dashboard technology, while would-be data visualization powerhouses such as Microsoft Corp. and Business Objects prepped eye-candy-laden data visualization and dashboards offerings of their own (Microsoft's PerformancePoint Server 2007 and Business Objects Xcelsius, which now ships as a standard part of Crystal Reports).

Proponents say dashboards and data visualization will continue to be key. In fact, some industry talking heads argue, 2008 could be an even bigger year for the technology, and that's saying something.

"It's just a great time to be in the data visualization business. There's just data coming from all angles, and the Internet is exacerbating that," says Kevin Brown, vice president of marketing with Tableau.

He's got a dog in the race, to be sure—Tableau is one of the most prominent best-of-breed data visualization players. But, as Brown notes, "We think the market for BI is actually in the tens of millions, as opposed to the tech-savvy users, the sort of TDWI guys—the guys who've mostly been using BI [up until now]."

## Data Integration Still Front and Center

Even the sleepest of technology segments can undergo an occasional revival. Consider what's been happening in the sleepy data integration (DI) sector over the last three years: when IBM Corp. plunked down over \$1 billion for the former Ascential Software Corp. in March 2005, the gloves came off, the sparring turned serious, and just about every major player let fly.

Informatica Corp., for example, found its legs, gathered its wits, and vaulted to the forefront of the DI market by positioning itself as Data Integration Switzerland. Other players also augmented their DI chops: Oracle (which released a significant refresh of its Oracle Warehouse Builder in the summer of 2006 and acquired ETL specialist Sunopsis a few months later); Sybase Inc. (which acquired EII specialist Avaki and ETL specialist Solonde AG during that fateful summer of '06); IBM (which fleshed out its unstructured content management strategy with the acquisition of the former FileNet, also in the summer of 2006, and this year bolstered its ETL capabilities by acquiring changed data capture [CDC] specialist DataMirror Inc.); Microsoft; and SAP.

**"THE SMARTEST COMPANIES WILL DEVELOP CLEAR GUIDELINES OF WHEN TO USE THE STANDARD, PREFERRED VENDOR AND WHEN TO SUPPLEMENT CAPABILITIES WITH PRODUCTS FROM ANOTHER VENDOR."**

Microsoft, of course, released a substantially retooled version of its former Data Transformation Services (rechristened Integration Services) with SQL Server 2005, while SAP got itself a best-of-breed ETL tool (and matching data quality technology) when it nabbed Business Objects a couple of months ago.

Nor is DI innovation a thing of the past: Business Objects, IBM, Informatica, and Sybase Inc. all announced new revisions of their data integration or data quality suites this year. These vendors talked up the real-time credentials of their DI offerings, touting a universe of



complementary services—including EII, event processing, legacy connectivity, and (in the case of both IBM and Informatica) collaboration—that they say are crucial DI differentiators.

What's ahead? DI advocates think we'll see a host of drivers, including, in particular, master data management (MDM) and regulatory compliance, which will continue to turn up the heat on enterprise DI efforts.

As a result, another established DI technology category—data quality—will become even more important.

“[D]ata quality is fast becoming a key business asset,” says Arvind Parthasarathi, senior director of solutions with Informatica. He cites research from market watcher Gartner Inc., among others, which projects that the DQ tools segment could grow to nearly \$700 million by 2011 as more and more businesses tackle lingering data quality issues.

**“FROM A BUSINESS POINT OF VIEW, 2007 SAW MERGERS AND ACQUISITIONS HIT A RECORD HIGH; WE CAN ONLY ASSUME THAT THE NUMBER WILL GROW IN 2008, THUS INCREASING THE NEED FOR ACCURATE AND RELIABLE DATA.”**



“[T]he importance of data quality will influence the business environment dramatically in 2008. The quality of data can affect our day-to-day lives without us even thinking about it,” Parthasarathi points out. “From a business point of view, 2007 saw mergers and acquisitions hit a record high; we can only assume that the number will grow in 2008, thus increasing the need for accurate and reliable data. As a result of this trend, organizations continue to be more accountable for the quality of their data [because they must] adhere to governmental regulations and industry compliance and demonstrate accountability to customers and shareholders.”

## Data Warehousing Reborn

Surprisingly, the bread-and-butter DW segment itself also seems to be seething. This change—this ongoing transformation—became most apparent this year, although rumblings were felt as far back as 2002.

Thanks to a number of market drivers, including both the emergence of DW appliances and the push toward real-time DW, the data warehousing segment is as exciting as ever. There are a host of new players, for one thing: not just DW appliance start-ups like Dataupia Inc. (which launched earlier this year); nor, for that matter, would-be appliance giants such as Oracle's Optimized Warehouse for Sun (which launched just last month), Hewlett-Packard Co.'s ambitious Neoview DW appliance (which relaunched earlier this year), or an independent—and newly invigorated—Teradata, which separated from parent company NCR Corp. a few months ago.

No, new players means just that: new players. Companies such as ParAccel and Vertica, both of which market high-performance columnar database technologies. Along with Dataupia and Kognitio (which, although an established player in the UK and EU markets, is relatively new to the U.S.), they comprise a bumper crop of 2007 upstarts: companies that, for one reason or another, burst on the scene—and on customers' radar screens—over the last 12 months.

Ironically, many of these start-ups rely on relatively mature technology models: Netezza Inc. more or less invented the DW appliance five years ago, for example. Columnar database structures aren't anything new, either: Sybase has marketed its seminal IQ columnar DW system for more than a decade, while Sand Technologies, another columnar database specialist, has more than two decades of experience behind it.

Nevertheless, both technologies are hot: IBM, HP, Sun, and Oracle (along with established players Netezza, DATAlegro, and—depending on how one defines “appliance”—Teradata) all market DW appliances, while Dell Computer Corp. markets a DW appliance of its own.

Columnar database technologies are sizzling, experts say, because they're uniquely equipped to deal with the twin demands imposed by real-

time data warehousing scenarios: high volumes and high performance.

In addition to boosting performance by several orders of magnitude over their relational kith (both ParAccel and Sybase, for example, tout improvements of 100x or more in specific situations), columnar databases are also highly compressible—much more so than their relational brethren. The result, experts say, is a measurable uptick in sales of columnar database technologies. That, in all likelihood, is a trend that we'll continue to see in 2008.

“For DW, the columnar approach has a lot of advantages,” says Philip Russom of TDWI Research. “[T]oday we’re seeing a slight uptick in the popularity of columnar database technologies. There are suddenly all these competitors,” he continues, noting that the enterprise “proof of concept” for columnar database technology was established by Sybase IQ a decade ago.

“For some applications,” Russom says, “they [columnar databases] deliver much faster query response times, because of the structure of the [columnar] database itself. They’re much more compressible, too.” For customers whose needs outstrip the capabilities of conventional relational DW systems, Russom says, columnar technologies offer a compelling (and cost-effective) alternative.

### **Collaboration is Key**

Finally, more industry players are recognizing the importance of collaboration. Both IBM and Informatica, for example, touted the collaborative capabilities of the next-generation DI suites they delivered this year. Meanwhile, many purveyors of BI client tools have gotten hip to the power of collaboration, too: Business Objects, Cognos, MicroStrategy, and Tableau—just to name a few—all sought to emphasize their collaborative chops this year—and with good reason.

On the DI front, for example, collaboration is crucial, experts say.

“Collaboration has become a pressing requirement for data integration in recent years,” explains TDWI’s Russom. “As the number of data integration specialists continues to grow—into double

digits in some organizations—so increases the need for development processes and tool functions that help them communicate and collaborate. Likewise, stewards, business analysts, and other business people are joining the ‘extended’ data integration and data quality team, which brings a new slew of collaborative requirements.”

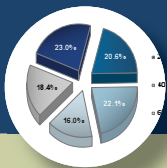
Tableau’s Brown, for his part, links the emergence of collaboration—as a potential killer application, at any rate—with the growing popularity of social networking technologies. “I think this Web 2.0 world, especially around social networks—MySpace, Facebook, LinkedIn—is creating this expectation for people ... to want to interact with other people and post things—like pictures, music, or thoughts, in a social kind of setting,” he concludes. 🐟



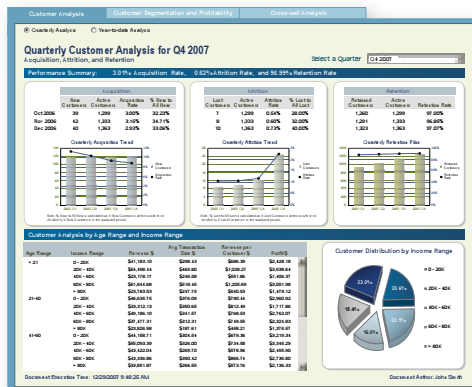
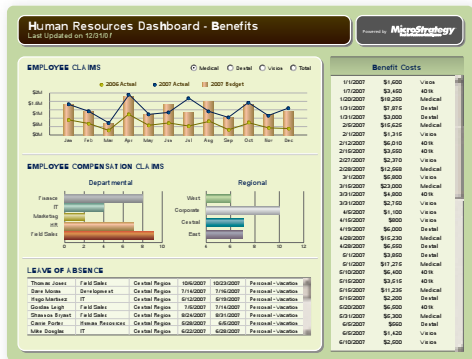
**THANKS TO A NUMBER OF MARKET DRIVERS, INCLUDING BOTH THE EMERGENCE OF DW APPLIANCES AND THE PUSH TOWARD REAL-TIME DW, THE DATA WAREHOUSING SEGMENT IS AS EXCITING AS EVER.**

---

*Stephen Swoyer is a technology writer based in Athens, GA. You can contact him via e-mail at [stephen.swoyer@spinkle.net](mailto:stephen.swoyer@spinkle.net).*



# If You Can Imagine it, MicroStrategy's Dynamic Dashboards Let You See It.



MicroStrategy has set a new standard for enterprise dashboards by combining dynamic data visualization and dashboard interactivity with MicroStrategy's highly scalable, industrial-strength business intelligence platform.

Dashboards come to life with unparalleled user interactivity, providing more insightful and efficient corporate performance analysis. MicroStrategy has eliminated the barriers that once separated world-class visualization from world-class reporting and analysis by seamlessly integrating Adobe® Flex™, Adobe® Flash® and AJAX into MicroStrategy's ROLAP architecture – delivering the best of both worlds without compromise.

Unleash the unlimited dashboard capabilities of MicroStrategy. To find out more, contact us at 888-537-8135.



For a Demo of MicroStrategy's Dynamic Enterprise Dashboards, visit [www.microstrategy.com/dashboards\\_tdwi](http://www.microstrategy.com/dashboards_tdwi)



# Forecast: The 2008 Technology Market

BY WAYNE W. ECKERSON AND PHILIP RUSSOM

What will 2008 bring? Here are some key bullets excerpted from TDWI's Technology Market Reports, an exclusive TDWI Member benefit. Current TDWI Members can download the complete reports at [www.tdwi.org/research](http://www.tdwi.org/research).

## **BUSINESS INTELLIGENCE**

### **PERFORMANCE MANAGEMENT SUPERSEDES BUSINESS INTELLIGENCE**

Performance management has become the new battleground for BI vendors, as signaled by the spate of acquisitions this past year. Performance management software, which monitors the execution of business strategy and plans, is the last untapped bastion within organizations to be automated with software. Performance management requires consistent data delivered via a data warehousing architecture; monitoring, reporting, and analytical tools delivered via dashboards; strategy mapping, metrics mapping, collaboration, and initiative tracking tools delivered via scorecards; and planning, budgeting, forecasting, and consolidation tools delivered via financial management software.

### AGILE DEVELOPMENT

Leading BI practitioners are investigating how to apply agile development techniques to BI. So far, the agile methodology, which calls for quick iterations (1–2 weeks) and small interdisciplinary teams, works well when developing front-end capabilities (i.e., reports and dashboards), but it's less well suited to designing and deploying data models, ETL mappings, and infrastructure components. Also, early practitioners report difficulty creating and applying suitable tests for each iteration and delivering suitable documentation.

### MS OFFICE INTEGRATION GOES MAINSTREAM

BI's integration with Microsoft Office is now a standard part of leading BI platforms, allowing organizations to leave behind their spreadsheets. New and enhanced products offer integration not only with Excel, but also with PowerPoint, Word, and Outlook. Newer integration features automatically convert tabular results into Excel PivotTables, embed query filters in spreadsheet cells, and support OLAP drills and Excel-based report authoring.



**ORGANIZATIONS ARE NOW LOOKING TO DELIVER GREATER BUSINESS VALUE FROM THEIR DATA WAREHOUSING AND BI INVESTMENTS BY EXPLORING “ADVANCED” TECHNOLOGY SOLUTIONS, SUCH AS PREDICTIVE ANALYTICS, TEXT MINING, AND TEXT ANALYTICS.**

### ADVANCED ANALYTICS

As a sign of the growing maturity of the industry, organizations are now looking to deliver greater business value from their data warehousing and BI investments by exploring “advanced” technology solutions, such as predictive analytics, text mining, and text analytics. Advanced analytics is not new to BI, but the ease with which they can be integrated with mainstream BI solutions is. Analytic workbenches have become easier for non-statisticians to use, and BI tools and database management systems now embed many sophisticated analytic functions.

### IN-MEMORY PROCESSING

The rise of 64-bit chips with greater memory space and growing user demand for instant response times and real-time data have caused many vendors and users to explore new ways to process BI tasks in memory. Next-generation OLAP, visualization, and planning tools now use in-memory processing to deliver state-of-the-art functionality, including real-time modeling and visual exploration.

### OPEN SOURCE BI MAKES HEADWAY

Pentaho, JasperSoft, and the Eclipse Foundation through its Business Intelligence Reporting Tools (BIRT) initiative are leading the charge to make BI available as open source software. While BI open source vendors cite lots of activity on their forums and tens of thousands of downloads, it's unclear yet whether open source BI has gained enough traction to make it a permanent fixture in the BI landscape. However, open source BI vendors are doing brisk business with independent software vendors who embed open source BI code into their products.

### EVENT-DRIVEN ANALYTIC PLATFORMS

There are many analytic applications that require real-time monitoring and process execution, such as Wall Street trading systems, e-commerce recommendation engines, fraud detection, online credit analysis, and so on. To date, organizations have custom-built these applications, but there is an emerging class of tools, which we call event-driven analytic platforms, that capture business events in real time off messaging backbones, filter, calculate, and aggregate events, apply rules, and trigger alerts, queries, updates, or other actions when predefined thresholds have been exceeded. Also called business activity monitoring or complex event processing, these systems generally store and/or access historical data to provide analytical context to rules execution. Syndera, Celequest, and SeeWhy Software are newcomers more aligned with BI systems than process management systems.

## DATA INTEGRATION

### OPERATIONAL DI

DI isn't just for BI anymore. Analytic DI—usually manifested as ETL in support of data warehousing—continues to grow as an established practice. But its blue-collar sibling—operational DI—is growing even faster, as DI is regularly applied to operational database and application consolidations, migrations, synchronizations, and upgrades.

### COLLABORATIVE DI

As the practices of analytic DI and operational DI have grown, so have the number of DI specialists in data warehousing teams, data integration competency centers, and on other teams. Very recently, mildly technical business users (like brand managers and business analysts) are demanding hands-on access to DI projects and their development artifacts. As the DI team gets larger and more diverse, DI tools must provide more collaborative functions for these people.

### DI HIGH AVAILABILITY

Data integration is being asked to operate more frequently per day, as well as in real time. This is needed to support business methodologies that demand fresh data, like operational BI, on-demand computing, and performance management. These methodologies can't manage a business without fresh data delivered reliably via real-time DI, so DI must be continuously available to enable them. Hence, when you cross the line into real-time DI, you also cross into DI high availability as a new requirement, which is met by fault-tolerant hardware and software or a cluster of DI servers that supports failover.

### CROSS-BUSINESS DI

Long caged by the corporate firewall, DI is now unchained and roaming the Internet. As evidence, note the many DI and BI vendor tools that have recently added connectors for Salesforce.com, the quintessential extra-enterprise application. Since a lot of cross-business communications pass via EDI and XML-based documents, support for these and other semi-structured data standards (and translations among them) is a rising requirement for DI solutions.

### EXTERNAL DATA AND THE WEB GAIN IMPORTANCE

Many organizations provide access to internal data, but have difficulty meeting the user demand for external data. Few are bringing external data from outside sources or Web sites into their environments. Early adopters are seeing benefits from incorporating this external data, mainly by leveraging specialized integration tools that evolved for use in an Internet environment. Expect demand for outside data to increase and for this trend to continue.



## LONG CAGED BY THE CORPORATE FIREWALL, DATA INTEGRATION IS NOW UNCHAINED AND ROAMING THE INTERNET.

### UNSTRUCTURED DATA IS THE NEW FRONTIER FOR DI

An enterprise data warehouse seeks to be a “single version of the truth” upon which most organizational decision making is based. However, it's not the whole truth unless it represents information from the mass of unstructured data—typically in documents of mostly text, like Microsoft Office files and e-mails—that all organizations have. The catch is that a specialized technology like text analytics is required to find and translate text-based information into SQL-accessible database records that data warehouses and BI tools can use. In the next few years, expect to expand your DI solutions to include text analytic capabilities. 🌟

---

*Wayne Eckerson is the director of TDWI Research and author of Performance Dashboards: Measuring, Monitoring, and Managing Your Business (John Wiley & Sons, 2005). You can reach him at [weckerson@tdwi.org](mailto:weckerson@tdwi.org).*

*Philip Russom is the senior manager of TDWI Research. You can reach him at [prussom@tdwi.org](mailto:prussom@tdwi.org).*



# Contemplating Consolidation

BY WAYNE W. ECKERSON

**What a year 2007 has been! The top three** business intelligence (BI) players—Business Objects, Hyperion Solutions, and Cognos—were gobbled up in megadeals worth nearly \$15 billion in total.

For years, we anticipated that the major ERP and database players would pose a serious threat to pure-play BI vendors. But we were sideswiped by the enormity and rapidity of the deals, which were announced within the span of eight months. Like a domino effect, once Hyperion announced its intention to be acquired by Oracle in March 2007, both Business Objects and Cognos had to follow suit or be outflanked by a former rival with greatly expanded account control and sales and marketing heft.

At the beginning of 2007, the BI Big Three were growing rapidly through serial acquisitions and gaining steam as BI became the top priority for CIOs, according to numerous market surveys. The vendors were obviously bulking up to compete more effectively against each other in pursuit of the emerging market for performance management solutions before low-cost business intelligence providers, such as Microsoft

and open source vendors, could undermine their revenues and margins from their core BI tools business. At the same time, their rapid growth would position them favorably with the ERP vendors in oddly divergent ways: first, to compete with the BI and performance management offerings of the ERP vendors, and second, to increase their value as takeover targets.

## Dancing with Wolves

Obviously, the BI Big Three only dabbled with the first option before sprinting to find dance partners in the ERP camp. In the ritual of capitalism, fear and greed proved more attractive than focus and independence. But that is how most markets evolve. What remains to be seen is whether these new partnerships add value to their acquiring companies, and more important, to BI consumers. Will the new arrangements accelerate or dampen BI innovation? Will they increase market penetration for BI tools from a measly 20 percent of potential users as it stands now? Will they accelerate revenues, or did the ERP vendors foolishly acquire them at the height of their popularity and earning power?

**Hyperion-Oracle.** So far, Oracle has done an admirable job of integrating Hyperion's products and people into its organization and product portfolio. It's worked quickly to rename, reprice, and repackage Hyperion products and establish a go-to-market product strategy for BI and performance management that makes sense without alienating customers. In other words, Oracle is betting on the best products without killing the worst. It has also created a formidable field sales team to push BI and performance management solutions both to its numerous existing accounts worldwide and to prospects.

**Business Objects-SAP.** SAP seems to be doing the right thing with Business Objects: leaving it alone. Just as Oracle recognized that its new acquisition had more expertise in BI and PM than it does, SAP seems to be giving Business Objects the reins to define SAP's strategy in these domains. Given how desperate SAP customers are for good-quality BI solutions (even if they don't yet know it), Business Objects must be as happy as a shark in a school of tuna with all the SAP accounts that will now take a sales call.

**Cognos-IBM.** The big question mark is what IBM will do with Cognos. IBM basically torpedoed Ascential Software, which was a leading provider of ETL tools until IBM acquired it in 2005. The toolset, now called IBM WebSphere DataStage, does not show up in many competitive ETL bakeoffs, ceding the market to Informatica. Like Ascential, IBM is placing Cognos in the Information Management software division, which is primarily focused on delivering data management and data integration solutions. As such, Cognos will be an odd fit since it provides tools and applications that might sit on top of a data management and integration platform. Of course, perhaps Cognos is the missing cog in IBM's data strategy that will give the vendor a compelling end-to-end solution to offer the market.

## The Rest of the Bunch

There are 100+ BI vendors in the TDWI Marketplace ([www.tdwi.org/marketplace](http://www.tdwi.org/marketplace)), so lest you think that the ERP vendors now own BI, that's not true. Many of the "little guys" say that these acquisitions are a good thing, since the big players will be too consumed with organizational politics and product integration to add much value for BI consumers.

That traditionally has been the case when BI vendors acquired smaller rivals. Cognos, Business Objects, and Hyperion are still digesting previous acquisitions, showing an integrated architecture only in slideware. However, all three vendors have made headway on this front, which shows they have the will and resources to make it happen. However, now they have an additional set of products to integrate with. The best approach for the ERP vendors to take is simply to integrate the new BI products to their core application and database platforms and not worry about rationalizing and integrating the vast array of BI and PM products under their control. Oracle, for sure, knows how to acquire and integrate software companies and products, since it does so every month or more often.

While I believe the "little guys" will always be a source of innovation, I don't think the "big guys" will stop innovating. They now have more resources to innovate (and most BI players had active research labs set up to explore fringe capabilities), and what they don't innovate internally, they'll acquire. The niche players will also have a harder time breaking into accounts that are controlled by the ERP vendors. And if they do, they'll have to spend more time justifying why the prospect should choose a "nonstandard" product.

## 2008 and Beyond

In 2008, we'll begin to see the contrails of our reoriented BI marketplace. We'll find out whether consolidation has shut the door on BI innovation or accelerated it. We'll find out whether BI can penetrate the remaining 80 percent of users who can benefit from BI or whether the tools and applications will remain too unwieldy and cumbersome for most. Finally, we'll discover whether consolidation yields to commoditization or additional consolidation as vendors jockey to stake out new or revitalized ground in the BI market. Stay tuned! 🍷

---

*Wayne Eckerson is the director of TDWI Research and author of Performance Dashboards: Measuring, Monitoring, and Managing Your Business (John Wiley & Sons, 2005). You can reach him at [weckerson@tdwi.org](mailto:weckerson@tdwi.org).*



# Insight Perception

## Business Intelligence

**Most companies can't take full advantage of their IT investments.**

It's too hard. They access and analyze less than 20 percent of their data, and deploy business intelligence to fewer than 10 percent of their internal users.

Imagine how productive you could be if you could break these barriers – if you could deliver insight to every employee in your organization and beyond.

Information Builders turns that vision into reality. Our WebFOCUS business intelligence platform makes your applications work together to give businesspeople the insight they need. And WebFOCUS isn't our tool, it becomes your application. That's why it runs the world's largest BI deployments, sometimes reaching hundreds of thousands of people. In fact, you probably already use it.

**WebFOCUS. Because everyone makes decisions.**

**To learn more, visit [informationbuilders.com/go/BI2008](http://informationbuilders.com/go/BI2008).**



# Best Practices in Operational BI

BY WAYNE W. ECKERSON

**Operational business intelligence (BI) represents a turning** point in the evolution of BI. Traditionally, BI has been the province of technically savvy business analysts who spend many hours with sophisticated tools analyzing trends and patterns in large volumes of historical data to improve the effectiveness of strategic and tactical decisions. But operational BI changes this equation: it moves BI out of the back room and embeds it into the fabric of the business, intertwining it with operational processes and applications that drive thousands of daily decisions. In essence, operational BI merges analytical and operational processes into a unified whole.

In addition, operational BI increases the value of BI by delivering information and insights on demand to all workers—from the shipping clerk to the CEO—so they can work smarter and faster to achieve critical business objectives. In essence, operational BI delivers the right information to the right people at the right time so they can take action. In its extreme form, operational BI encapsulates business insights into rules and models that organizations can use to automate decisions and responses, eliminating the need for human intervention. Automating decisions not only streamlines processes and reduces costs, but also improves service and gives organizations a competitive advantage in the marketplace.

**BEYOND OPERATIONAL REPORTING.** There are many flavors of operational BI, ranging from operational reporting and process monitoring to composite applications and decision automation. While most organizations already support some form of operational reporting, many have yet to embrace more complex types of operational BI that generate greater business value. Thus, operational BI opens up a new field of endeavor for BI and gives organizations a chance to reap greater dividends from their BI investments.

## Challenges

However, operational BI poses several challenges. It stretches the architectural boundaries of current BI solutions, forcing BI professionals to rethink the way they design and build systems. Queries must return in seconds rather than minutes or hours, and reports must update dynamically. Operational BI systems must capture large volumes of data in near real time without degrading the performance of existing processes and jobs on source or target systems. There is also less time to recover from a server outage, making it imperative for BI professionals to build resilient, highly available systems with sufficient backup and recovery.

**TO WAREHOUSE OR NOT?** The first technical question that an organization needs to address is whether to use a data warehousing architecture to deliver just-in-time data, or bypass it altogether. This is the most fundamental question, and the hardest to answer.

Many architects believe it's critical to adapt existing data warehouses to support operational BI. "The big showstopper is whether you are going to apply the same business rules to integrate, cleanse, and validate operational data streams as the rest of the data in your data warehouse," says John O'Brien, a BI consultant and former data warehousing architect. Pulling operational streams out of the data warehousing process undermines data quality and creates divergent data sets that may not reconcile, he claims.

Some disagree. They say that a data warehouse becomes a bottleneck if you try to load all data into it that users may possibly want to query. BI vendors such as Business Objects, Hyperion, SAS, and InetSoft, which support federated query, believe their tools provide an easy, low-cost

way to capture real-time data and deliver it to users at the point of need with sufficient quality. Likewise, vendors of embedded BI, event-driven analytic platforms, composite applications, and in-memory analytics believe their offerings provide the most suitable way to meet high-end operational BI requirements. These vendors say that a data warehouse is critical for applying historical context to real-time data, but not necessary for managing the real-time data itself.

Nevertheless, data warehousing has become a well-established IT practice in corporate environments, and few organizations will jettison their DW investments without trying to adapt the architecture to support just-in-time data and operational processes. Our survey also shows that half of organizations run both operational and analytical reporting from the same environment. Given the high number of survey respondents who equate operational reporting with operational BI, this indicates that many companies have moved operational reporting into their data warehousing environment. (See Figure 1.)

Organizations that want to deliver just-in-time data via a data warehousing environment face daunting challenges. The rest of this section describes the major challenges involved in transforming a data warehousing environment to support operational BI.

**SELECT THE RIGHT TECHNOLOGY.** There are many technologies that BI architects can use to build an operational BI environment. These technologies can be classified into three main categories that correspond to the way data flows through a just-in-time system:

- **Data acquisition.** Organizations must capture, transform, and move data from source systems into the analytical environment on a just-in-time basis. To do this, organizations can use ETL tools, replication tools, changed data capture technologies, messaging backbones, event-driven streaming engines, and BI and EII query tools. In practice, most companies use a combination of these.
- **Data storage.** Organizations can store acquired data in a variety of database engines, including one or more layers in a data warehousing environment (i.e., staging

### Is your operational BI environment the same as your regular BI environment?

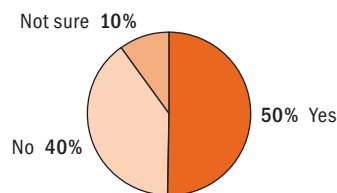


Figure 1. Based on 225 respondents who have implemented operational BI according to TDWI's definition.

area, data warehouse, data mart, OLAP cube database), an ODS, a low-latency database, or an event-driven analytic engine.

- **Data delivery.** To display captured data or data derived from analytic processes, organizations use custom-built applications, portals, BI tools, dashboards, composite applications, or a combination of these approaches.

Some of these technologies support multiple tasks. For instance, BI and EII tools can both acquire and deliver just-in-time data, and if you consider a report definition a storage mechanism, they support all three activities. Likewise, event-driven analytic platforms, working in concert with messaging backbones (EAI networks, enterprise service buses, publish/subscribe middleware), support all three activities. This is not surprising, since they are designed to provide end-to-end support for real-time operational BI requirements.

Our survey shows that organizations use a variety of techniques to support operational BI. The most common are ODSs, followed by more frequent batch loads and changed data capture. (See Figure 2.)

**INCREASE SCALABILITY AND THROUGHPUT.**

Another key issue is building systems that scale to support greater numbers of users, more data sources, and higher volumes of data with increased rates of throughput—while ensuring high data quality and security. Organizations need to implement high-performance data warehousing platforms from leading companies such as Sybase, Teradata, and HP, and robust data integration platforms like those from Syncsort and others. To ensure scalability and throughput, companies may need to upgrade their networks and hardware, and parallelize key extraction, transformation, and load processes so they can adequately meet these increased throughput demands and eliminate processing bottlenecks.

**USE INSERTS.** To increase ETL throughput, some (but not all) architects recommend inserting new records (i.e., loading data) rather than adding, changing, or deleting existing records. “Avoid updates at all costs,” says O’Brien. “With large data sets, it takes too much time to find the record and update it.” Although this creates lots of duplicate data, administrators can use SQL

DISTINCT and GROUP-BY statements to identify and delete duplicates, O’Brien says.

O’Brien also recommends turning off referential integrity and database logging. “Why log inserts if you have the original load file?” he asks. In addition, there is no reason to create indexes and aggregations during the ETL process, since they consume a lot of CPU and slow down the inserts. Like many architects, O’Brien now advocates using the database to perform these tasks once the data is loaded into the data warehouse to exploit the parallelism of the underlying database. This process is known as extract, load, and transform (ELT) versus the more traditional extract, transform, and load (ETL).

**INCREASE AVAILABILITY AND RECOVERABILITY.**

In a just-in-time data warehousing environment there is little time to recover from errors, because batch windows are small or nonexistent. “In a real-time system, you have no time to fix problems, and messages can start piling up,” says OLAP architect Eric Lofstrom. The best way to avert these problems is to build high-availability systems to drive the acquisition and deployment process. These systems parallelize single points of failure, run in a clustered environment, support failover and backup processes, and have an offsite backup.

**MICROBATCHES.** To avoid losing data, O’Brien recommends using “microbatches” instead of real-time streaming, since microbatches provide a buffer with which to recover from a network or server outage. His rule of thumb is to set the ETL batch cycle three times larger than the time required to process the data. “If it takes 90 seconds to process a file, then I’ll set the batch

**Rate the importance of the following techniques to your operational BI strategy.**

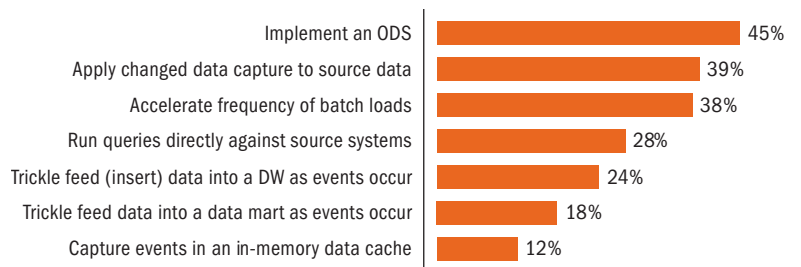


Figure 2. Based on 225 respondents who have implemented operational BI and rated the approach of “high” importance.

cycle to five minutes and have 3.5 minutes of sleep time.”

However, other IT practitioners believe it is perfectly safe to stream events into a data warehouse. Lofstrom, for example, recommends logging each message as it comes off a queue, which may slow down the speed of the streaming data, but prevents the loss of data caused by server outages. If an outage occurs, Lofstrom recommends loading accumulated messages into the warehouse in a single batch job before turning event streaming back on. In other cases, companies use two-phase commit to update streamed messages from a queue into a staging area, so updates are never lost if there is a planned or unplanned outage in the database server.



## ORGANIZATIONS THAT WANT TO DELIVER JUST-IN-TIME DATA FACE DAUNTING CHALLENGES IN TRANSFORMING A DATA WAREHOUSING ENVIRONMENT TO SUPPORT OPERATIONAL BI.

**DELIVER ADEQUATE PERFORMANCE.** A challenging problem for data warehousing designers is ensuring rapid query response times while simultaneously loading or updating a data warehouse and performing other tasks, such as monitoring events, triggering alerts, running backups, and scoring models. A mixed workload can cause RDBMS performance to degrade to the point where it blocks incoming queries or causes load processes to fail. Achieving adequate performance in a mixed workload environment can be a vexing challenge.

Complicating matters is that most users expect near-instantaneous response times in an operational BI environment. For example, customer service representatives can't wait more than a few seconds for an integrated view of customer activity and cross-sell offers before their callers get impatient and hang up. Most data warehouses aren't designed to support a mixed workload of tactical queries that return a few rows of data in less than a second along with complex requests that scan millions of records and take minutes or hours to process.

How do you balance these mixed workloads and maintain SLAs without causing the system to crash? There are several options, all of which have trade-offs:

- **Throttle incoming requests in response to heavy system loads.** You can schedule queries or jobs to only run at night, or queue or delete them using query governors or job managers during periods of heavy load to ensure the integrity of system updates. This approach obviously won't sit well with users, and it's the reason companies began building data warehouses in the first place: to provide the business with a dedicated analytical environment that would not interfere with operational processes.
- **Separate analytical and tactical queries and jobs.** Another approach is to off-load tactical queries and predictive scoring jobs to an ODS or specialized data mart and leave the data warehouse to support long-running analytical queries. Here, the ODS contains current transaction data, usually no more than 30 days' worth, stored in third normal form models that reflect operational schemas. The downside of this approach is that it separates historical and current data into different systems, making it difficult for users to view just-in-time data in historical context.
- **Leverage RDBMS mixed-workload capabilities.** Many database management systems (DBMS) have significantly improved their ability to optimize the performance of long-running strategic queries, short-running tactical queries, load and update processes, and other types of workloads. Teradata, for example, has made big improvements in its priority scheduler for optimizing tactical queries, says Dan Graham, marketing manager at Teradata. The priority scheduler is a real-time traffic cop that always allocates CPU to tactical queries first to guarantee sub-second response times. It continuously monitors in-flight tasks and dynamically adjusts CPU allocation to optimize performance. In addition, administrators can configure and prioritize workloads by user groups, types of activity, and other variables. HP's new Neoview data warehousing platform also boasts workload management capabilities.

An RDBMS that supports mixed-workload processing lets organizations have their cake and eat it, too: they can load current and historical data into the same database and optimize performance across all types of queries. Without an RDBMS that supports mixed workloads, many organizations decide to avoid using a data warehouse for operational BI. However, not all mixed workload capabilities are created equal, so evaluate your database vendor's capabilities carefully. Also, running multiple workloads on a single platform may require hardware upgrades to maintain adequate performance, so calculate these additional costs when deciding whether to use a data warehouse to support operational requirements.

**AVOID CONTENTION PROBLEMS.** When user queries and ETL processes hit the same table, there is the potential for one process to block the other and cause query or load performance to degrade. Along with mixed workload requirements, the issue of contention causes many people to maintain historical and just-in-time data in distinct data stores. There are three ways to minimize the contention problem, but they are not for the fainthearted:

- **Allow simultaneous reads and writes.** If you are inserting rather than updating data, the database won't lock the table. The only problem here is that the data can be out of sync if the user reads the table before all the inserts have completed. To avoid confusion, it's imperative to provide users with guidelines for submitting ad hoc queries against real-time data, and it helps if you dynamically time-stamp all reports. It's also important to turn off caching in BI servers, since there is no point in loading data every hour if the BI tool requests data from a cache that gets flushed once a day.
- **Insert and flip partitions.** Another approach is to create a duplicate partition of a fact table and load it with current data. Then, on a periodic basis, swap the table with the live fact table to update the data warehouse. This approach may require you to pause the application server while the flip takes place so no new requests are initiated while the swap occurs, says Justin Langseth, CTO of Claraview, a BI consultancy in Reston, VA.

Many organizations use this "insert-and-flip" approach.

- **External real-time cache.** Another approach Langseth recommends is to load data into an in-memory cache outside of the data warehouse. Requests for real-time data are fulfilled from the cache, while queries that require a combination of real-time and historical data merge the requested data in a set of temporary tables in either the data warehouse or real-time cache, depending on which database holds the largest volume of data requested. This type of just-in-time merging of information requires complex SQL and may be challenging for most BI tools to support.

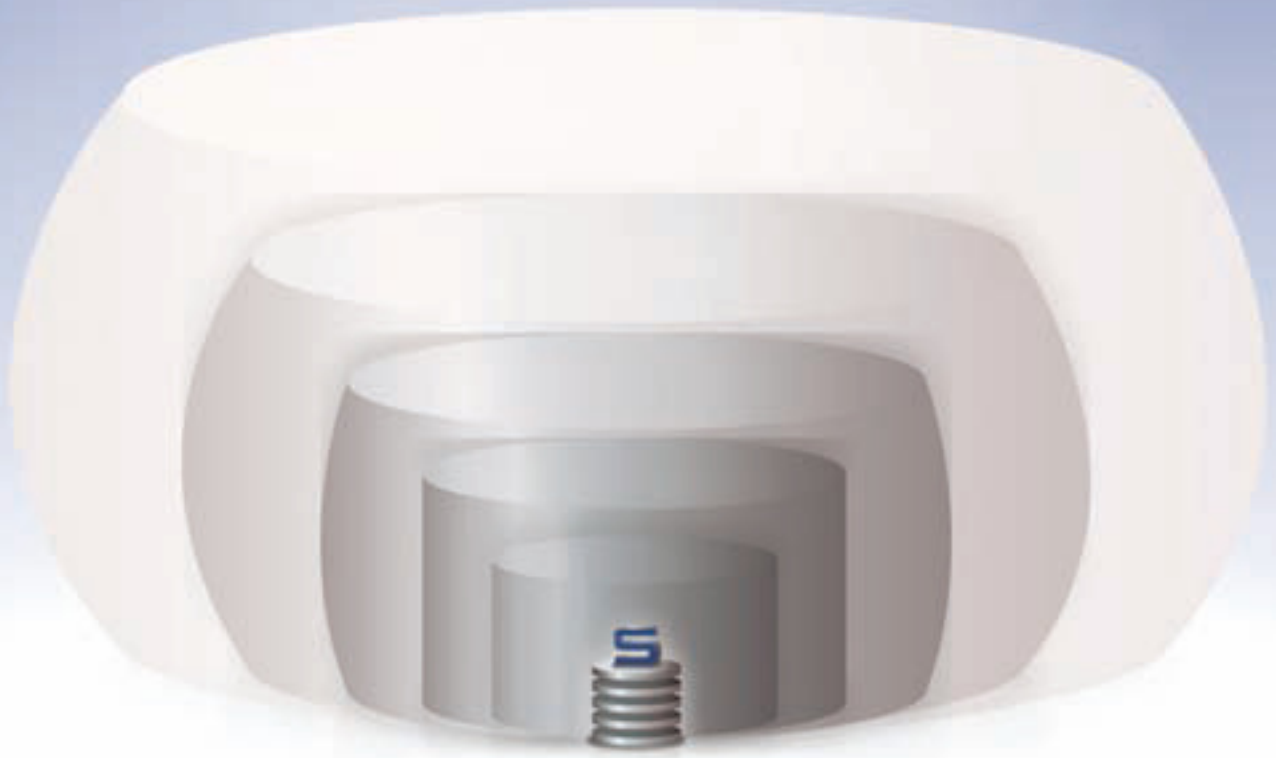
**SUMMARY.** There are many challenges involved in implementing operational BI. The challenges discussed here deal primarily with transforming a data warehousing architecture to support just-in-time data delivery. Although there are other ways to deliver just-in-time data to support operational BI, many architects recommend using a data warehousing environment to ensure a consistent set of data for the entire organization. 🍷

---

*Wayne Eckerson is the director of TDWI Research and author of Performance Dashboards: Measuring, Monitoring, and Managing Your Business (John Wiley & Sons, 2005). You can reach him at [weckerson@tdwi.org](mailto:weckerson@tdwi.org).*

This article was excerpted from the full, 30-page report by the same name. You can download this and other TDWI Research free of charge at [www.tdwi.org/research](http://www.tdwi.org/research).

The report was sponsored by Business Objects, an SAP company; HP; Oracle/Hyperion; InetSoft; SAS; Sybase; Syncsort Incorporated; and Teradata Corporation.



## **Data. Users. Questions. Regulations. Costs.**

Too much. Too fast.

Your existing data warehouse is struggling to cope.

We can help.

### **SAND/DNA Access.**

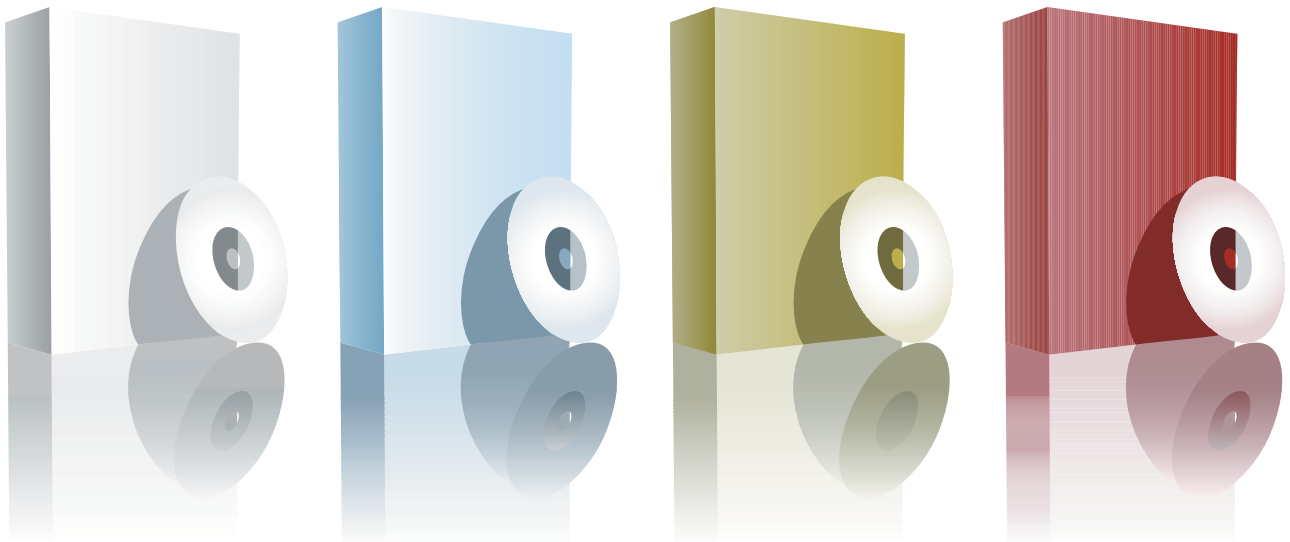
Increase results. Lower costs. Meet demand.

Find out how at [www.sand.com/tdwi](http://www.sand.com/tdwi)

**SAND**<sup>™</sup>  
TECHNOLOGY

[www.sand.com](http://www.sand.com)

+1 (877) 468-2538 +44 (0)1276 804604 +49 (0)40 32005 836



# Business Intelligence Solutions for SAP

## Issues that Influence BI Solutions for SAP

BY PHILIP RUSSOM

The issues discussed here quantify the state of BI solutions for SAP; they also describe the issues you need to consider before deciding which mix of tools, platforms, approaches, professional services, and business sponsorship are appropriate to your solution.

### Applications and Data Warehouse Architecture Issues

Most of the data that populates data warehouses and reports comes from various types of operational applications, including those for ERP, CRM, and SCM. Because these are prominent sources of data that BI is based on, the number and diversity of applications strongly influences the data content and architecture of any BI solution, especially those that involve SAP.

- **SAP users don't only use SAP applications.** TDWI's survey asked, "From which vendors has your organization acquired packaged applications for ERP and other operations?" The survey population is dominated by SAP users, so it's not surprising that 71% of respondents report using SAP applications. (See Figure 1.) But many reported also acquiring applications from Oracle (25%), Peoplesoft (24%), Siebel (15%), JD Edwards (12%), and so on. Given the numerous

**From which vendors has your organization acquired packaged applications for ERP and other operations? (Select all that apply.)**

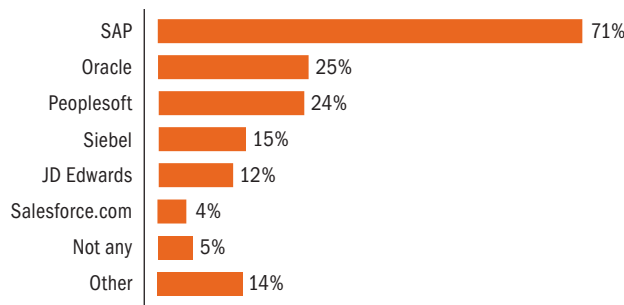


Figure 1. Based on 476 responses from 278 respondents. Average responses per respondent: 1.7

**Is the BI environment for SAP the same as your enterprise BI environment?**

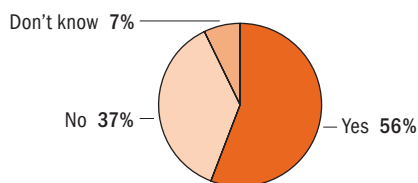


Figure 2. Based on 172 respondents who are SAP users.

**Who's the primary sponsor for your BI solution for SAP?**

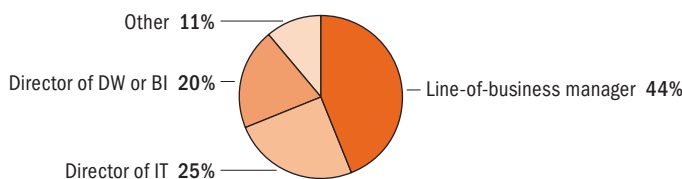


Figure 3. Based on 172 respondents who are SAP users.

vendors selected by survey respondents—plus home-grown solutions—it's safe to say that most SAP-centric businesses also have significant non-SAP applications.

- The SAP-centricity of your BI solution is a matter of degree.** In other words, a BI solution for SAP may incorporate SAP applications data only, or it may also integrate data from other applications (and even nonapplication sources like syndicated data). The latter is necessary for a complete view of organizational performance. In fact, gaining a centralized “single version of the truth” is one of the reasons organizations deploy enterprise data warehouses. The challenge is to satisfy both application-specific and enterprise-scope BI requirements.
- Application diversity may lead to BI solution diversity.** This may mean deploying both NetWeaver BI and an enterprise data warehouse (EDW). In most cases, NetWeaver BI supports data marts and reports specific to businesses and processes supported by SAP applications, whereas the EDW collates information from many sources for strategic decision making. To reduce the number of solutions, some SAP-centric organizations have made NetWeaver BI the equivalent of an EDW, while others rely on an EDW without any SAP BI products. Selecting one of these data warehouse architectures—or a combination of them—is a tough decision for SAP users, because it affects the cost of BI, completeness of data content, data ownership, and how BI goals are prioritized (i.e., tactical operations versus strategic planning).

**Organizational Issues**

TDWI asked of SAP users: “Is the BI environment for SAP the same as your enterprise BI environment?” (See Figure 2.)

- Enterprise-scope BI.** Over half (56%) said “yes,” which suggests that these respondents’ BI infrastructure is enterprise in scope—that is, it spans the data and BI needs of many applications and business units, not just those associated with SAP applications.
- Application-scope BI.** Roughly one-third (37%) said “no,” suggesting these users have a BI solution that is narrowly focused

on the data and business processes of SAP applications. Of course, this does not preclude them from also having a separate BI solution of enterprise scope.

- **Determining scope is an important BI design decision.** A BI solution for SAP may be enterprise scope, with SAP-specific requirements satisfied by a subset of the solution. Or it may be application scope, possibly complemented by an enterprise-scope BI solution.

The scope of a BI solution for SAP is related to its ownership and sponsorship. (See Figure 3.)

- **Line-of-business managers (44%) commonly sponsor BI solutions for SAP.** Line-of-business managers often fund, own, and sponsor operational applications, so it's possible that they have extended this role to also sponsor application-specific BI. Even so, IT and BI directors sponsor BI solutions for SAP just as often (45%).
- **SAP BI sponsorship can also be shared.** Several survey respondents selected "Other" and entered comments like "we use a team approach, no single sponsor." Sponsorship is "shared by all businesses" where it involves "several stakeholders." Or BI may be sponsored by "corporate leadership as part of the SAP implementation."
- **Finance departments often sponsor BI, whether for SAP or not.** We're reminded of this by survey respondents who listed their finance organization or its CFO, VPs, and directors.

### Build versus Buy Issues

When it comes to ERP applications and the classic build-versus-buy decision, buying is the norm by a long shot, with most survey respondents buying and customizing (57%), followed by those who buy but don't customize much (27%). (See Figure 4.) By comparison, building an ERP system is rare (9%). This is natural, given the complexity of ERP processes and how difficult it would be to build and maintain an application to automate them. Plus, multiple software vendors offer mature ERP systems that are feature-rich, robust, and have solid track records.

To quantify the build-versus-buy approaches of BI solutions for SAP, TDWI's survey first identified respondents who are SAP users, then asked them: "Which best describes how you built your BI solution for SAP?" (See Figure 5.) A whopping 75% reported acquiring BI products from SAP, which is a high penetration of BI products for an applications vendor.

Note that the level of customization varies with SAP BI products. Relatively few use SAP BI products "as is" (10%), while most assign in-house personnel (35%) or hire consultants (30%) to customize them. Customization has ramifications for packaged solutions:

- **BI solutions are hard to package.** That's because an organization's BI data sources, its data models, and its report presentations all vary tremendously. However, if you narrow the

### What strategy has your organization followed in acquiring ERP applications?

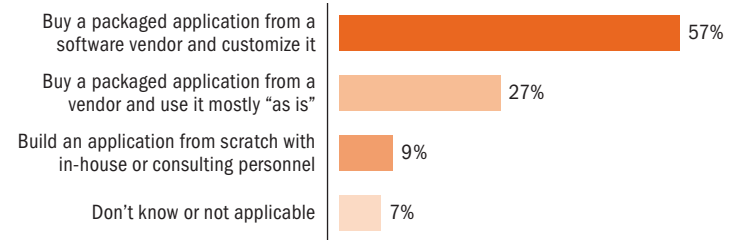


Figure 4. Total percentage based on 278 respondents who use the applications from SAP and other vendors.

### Which best describes how you built your BI solution for SAP?

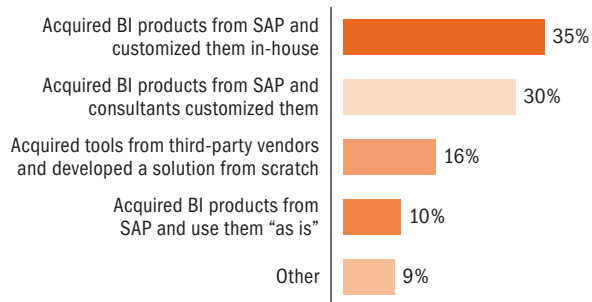


Figure 5. Total percentage based on 172 respondents who are SAP users.

scope of the BI solution to specific use cases, which SAP NetWeaver BI does by focusing on standard business processes, the breadth of user requirements is reduced, thereby reducing the difficulties of packaging a solution that works for multiple user organizations. But BI requirements still vary somewhat, even when focused on ERP, so organizations inevitably customize to some degree.

- **The greater the customization, the greater the difficulty of upgrading.** Some SAP customers would like to upgrade from older BW releases to recent NetWeaver BI ones, but they feel that customization makes the upgrade too difficult to attempt. Let's all recognize that customization is inevitable with BI products, so upgrades usually involve development work.
- **Avoiding customization reduces project time and cost.** A tried-and-true strategy for a packaged solution is to roll out the first phase with little or no customization, then tailor the solution to organizational requirements in subsequent phases.



## MANY RESPONDENTS HAVE STRONG FEELINGS EITHER FOR OR AGAINST SAP AS A PREFERRED SOFTWARE SUPPLIER AND HOW THAT PREDETERMINES THE PRODUCTS THEY USE.

“We started our implementation of SAP BW in 2002, working in parallel with the firm’s implementation of the SAP ERP,” said Brian Hickie, former VP of business intelligence at McKesson Pharmaceutical. “We only had nine months to implement the data warehouse, along with considerable ETL jobs and a number of initial analytics and reports. We made the deadline despite the fact that the SAP ERP configuration precluded the use of standard business content or standard data extractors; the ETL jobs and BW required a considerable amount of configuration to provide the data to generate analytics and reports from the ERP and legacy systems.”

Developing home-grown BI solutions isn’t as common as other approaches in Figure 5 (16%). However, other survey questions showed higher

rates, and survey respondents who selected “Other” entered responses that involve home-grown solutions:

- **SAP BW and EDWs can coexist peacefully.** Some BI solutions for SAP involve both NetWeaver BI (specifically the BW component) and an independent enterprise data warehouse (EDW). For example, one survey respondent said: “We both use BW and extract SAP data into a data warehouse.” Another reported using “BW and an in-house solution based on Oracle and Business Objects,” while yet another said, “[We] use SAP BI as an OLAP tool to access a Teradata EDW.”
- **Some BI solutions for SAP are a work in progress.** As one survey respondent put it, “We’re in transition from a legacy custom-developed EDW to SAP BI.”

### Issues in Software Acquisition Strategies

TDWI asked of SAP users taking this report’s survey: “What was the primary reason for your organization’s implementation of a BI solution for SAP?” In the survey, the question had no pre-written answers to select; instead, each respondent typed an answer in his/her own words. Many respondents turned the question into a soapbox, because they have strong feelings either for or against SAP as a preferred software supplier and how that predetermines the products they use.

- **For many organizations, naming SAP a priority supplier leads inevitably to SAP BI products.** Some respondents stated this in matter-of-fact answers such as: “SAP is our preferred provider in all areas.” Others made a direct causal link: “Using SAP ERP was the main driver to start with an SAP BW solution” and “SAP is selected as our common platform, which includes BW.”
- **Some spoke of the positive leverage between NetWeaver BI and SAP ERP.** “BW is a good fit on top of SAP ERP and SAP CRM,” enabling users “to leverage prebuilt content and integration with the ERP.” According to other survey respondents: “Most of our [BI] information is sourced from SAP components,” and “BW provides a direct link to primary source data.”

- **A few people complained about the SAP-centric strategy.** One claimed that using SAP BI products is due to a “strategic decision to use SAP overall, even when it does not fit.” Others received SAP BI products due to management actions beyond their control: “It [NetWeaver BI] was paid for in an overall package of software” and “Parent company purchased SAP solutions.” Selecting a preferred supplier is a good procurement practice, in general, but it’s frustrating for IT professionals when the strategy doesn’t consider user requirements.
- **Most mentioned the usual benefits of BI as reasons for implementing a BI solution.** Several listed common BI goals like better decision making or improved corporate performance. Others cited specific BI needs such as financial reporting, data sets for business analysts, or operational reporting based on SAP ERP source data.

“All of my clients have multiple application modules from SAP, and for most of them, that’s a good enough reason to get their BI technology from SAP,” said a consultant who specializes in BI solutions for SAP. “Why bother with an independent data warehouse? BW infrastructure—especially with recent NetWeaver releases—is almost identical to what you’d build into an independent warehouse. So the common practice I see is to leverage what you already have, instead of reinventing the BI wheel. Also, my clients trust SAP as a large, stable, non-acquirable vendor. In terms of best of breed, other BI and data integration tools are better, but SAP’s offering is relatively complete and pretty good. And the direct integration between BI and ERP is hard to beat.”

“We have a mature, best-of-breed technology stack for data warehousing and business intelligence, but it’s currently at risk because our company’s being acquired by a company with a deep commitment to SAP applications and BI,” said an enterprise data architect. “We’re under pressure to abandon our EDW and go with SAP NetWeaver BI, but we’re not sure it will meet our requirements for strategic and financial reporting, much less scale up. The merger is barely under way, and we’ve just started investigating NetWeaver, so it’ll be a long time before we know which way to go.” 🍂

---

*Philip Russom is the senior manager of TDWI Research and oversees many of TDWI’s research-oriented publications, services, and events. He can be reached at [prussom@tdwi.org](mailto:prussom@tdwi.org).*

This article was excerpted from the full, 32-page report by the same name. You can download this and other TDWI Research free of charge at [www.tdwi.org/research](http://www.tdwi.org/research).

The report was sponsored by Business Objects, an SAP company; Cognos; HP; IBM; Microsoft; MicroStrategy; SAP; and Teradata Corporation.



## Balanced growth - thanks to myBiQ's process & performance management suite:

analyzer & planner cockpit  
 balanced scorecard  
 fast & smart close  
 reporting  
 excel plug-in  
 monitoring of processes  
 data integration & cube replicator

dashboards  
 rule designer  
 metrics wizard  
 consolidation (legal & taxation)  
 recalc for foreign exchange rates  
 governance  
 planning & budgeting

## Forged from Open Source tools like:

OLAP with Palo from Jedox | ETL based on Pentaho's Kettle | Talend\* Open Studio | Reporting with JasperReports | Actuate BIRT | DBMS with PostgreSQL

## Certified OSBIF - Open Source BI Framework

- easy to implement - fully integrated
- enterprise class with proof of concept
- commercially supported, worldwide

## Value Added Commercial Applications

- dashboard designer
- rule wizard for consolidation & planning
- excel spreadsheets for the web

Visit our website <http://www.myBiQ.com>

Product and company names mentioned herein may be trademarks of their respective companies. ©2008 myBiQ LLC





# TEN MISTAKES TO AVOID IN DATA QUALITY MANAGEMENT

BY ARKADY MAYDANCHIK

## Foreword

The corporate data universe consists of innumerable databases linked by countless data interfaces. While the data continuously moves about and changes, the databases and the programs responsible for data exchange are endlessly redesigned and upgraded. Typically, this dynamic results in information systems getting better while data quality deteriorates. This is unfortunate, since quality is what determines data's intrinsic value to businesses and consumers. Information technology magnifies this intrinsic value. Thus, high quality data combined with effective technology is a great asset, while poor quality data combined with effective technology is an equally great liability.

Yet we tolerate enormous inaccuracies in databases and accept that most of them are riddled with errors—while corporations lose millions of dollars because of flawed data. Even more disheartening is the continuous growth in the magnitude of data quality problems, fostered by exponential increase in the size of databases and the further proliferation of information systems.

This *Ten Mistakes to Avoid* offers advice on how to steer clear of common pitfalls and build an efficient data quality management program.

# one

## 1. INADEQUATE STAFFING OF DATA QUALITY TEAMS

“Who should be responsible for data quality management?” is a frequently asked question in the data quality profession. Uncertainty exists partly because the profession is still in its infancy; no clearly defined group has the appropriate expertise and responsibility. Even companies that form data quality departments often staff them with employees who have expertise in general IT and data but who have no specific data quality knowledge.

ONE FUNDAMENTAL MISCONCEPTION IS THAT DATA QUALITY WILL IMPROVE BY ITSELF AS A RESULT OF GENERAL IT ADVANCEMENTS.

Recently, I received an e-mail from one of my conference class attendees, an employee of a household name corporation, who wrote: “I am new to the data quality management world and have found myself in charge of enterprisewide data quality here at ... .” This single sentence explains the major challenge to our profession: Most people in charge of data quality initiatives lack data quality experience. As a result, data quality management programs tend to follow one of two scenarios:

1. Data quality initiatives fall into the laps of technical people within the IT group. For example, an attendee of another of my classes—a database administrator—was asked by her boss to outline a data quality assessment strategy. Why her? According to her boss, she was picked because data quality assessment involves writing queries, manipulating data, and understanding databases—skills that were part of her resume. Such reasoning makes as much sense as asking me to be a sports reporter for the *Chicago Tribune* because I can type, I’ve published some articles, and I watch sports a lot from the comfort of my living room couch.

2. Initiatives are spearheaded inside business units by the data users. This scenario appears to make some sense. The data users themselves can tell good data from bad, and because they are the ones most in need of quality data, business departments sometimes initiate their own data quality management projects. Of course, the problem is that business users lack technical expertise, which is why I get this question at almost every class: “Is there a tool that can manage data quality without any custom coding or querying?” My answer does not please those asking the question. I tell them, “Data quality management is an IT discipline and requires IT expertise.”

As it takes two to tango, so a data quality management team must include both IT specialists and business users. In addition, a team needs data quality experts—those who have firsthand experience in designing, implementing, and fine-tuning data quality rules and monitors.

# two

## 2. HOPING THAT DATA WILL GET BETTER BY ITSELF

One fundamental misconception is that data quality will improve by itself as a result of general IT advancements. Over the years, the onus of data quality improvement was placed on modern database technologies, better information systems, and sophisticated data integration solutions. I remember an HR executive telling me that because his company had implemented a modern HR and payroll system, the company’s data quality problems were solved. “PeopleSoft does not have data quality problems,” he said confidently. How could I respond? I said, “I cannot tell you whether or not PeopleSoft has data quality problems, but your company certainly does, and they are not going away. In fact, they are about to get much bigger.” He did not believe me; people do not like those who predict rain on a sunny day. A few years later, that company’s new HR executive called me for advice about their data quality strategy.

In reality, most IT processes affect data quality negatively. New system implementations and

system upgrades are a major source of data quality problems. Data integration interfaces create thousands of errors in the blink of an eye. Even routine data processing is prone to error. Thus, if we do nothing, data quality will continuously deteriorate to the point at which the data becomes a huge liability. The only way to address the data quality challenge is by a systematic, ongoing program that assesses and improves existing data quality levels, continuously monitors data quality, and prevents its future deterioration as much as possible.

## three

### 3. LACK OF DATA QUALITY ASSESSMENT

Nearly all data quality management programs focus on data quality improvement. A major obstacle on the path to higher data quality, however, is that most organizations—aware of the importance of data quality—are unaware of the extent of the problems with their data. Their knowledge of data quality problems is usually anecdotal, rather than factual. Typically, organizations underestimate or overestimate the quality of their data, and they rarely understand the impact of data quality on business processes. These two pitfalls cause the failure of many BI projects. Furthermore, data quality improvement initiatives, when put in place, often fail because no method is provided for measuring the improvements.

Assessment is the cornerstone of any data quality management program. It helps describe the state of the data and advances understanding of how well the data supports various processes. Assessment also helps the business estimate how much the data problems are costing it.

Data quality assessment that is comprehensive and recurring allows the business to:

- Set expectations for data quality in new DW and BI applications and reduce the number of unwelcome surprises when the data quality fails to support the new applications

- Plan and prioritize data cleansing initiatives and evaluate the potential ROI of data cleansing
- Understand root causes of existing data problems and investigate ways of improving data collection processes
- Monitor ongoing data quality, identify new problems that creep into the databases, and observe and manage data decay

## four

### 4. NARROW FOCUS

Systematic data quality management efforts originated in the 1990s from parsing, matching, standardizing, and deduplicating customer data. Over the years, great strides have been made in this area. Modern tools and solutions allow businesses to achieve high rates of success. This progress makes many organizations feel good about their efforts to manage customer data quality. A good number of organizations have implemented solutions by now, and it is fair to say that the level of quality in corporate customer data is at its highest ever.

Unfortunately, the same cannot be said about the rest of the data universe. Data quality has continually deteriorated in the areas of human resources, finance, product orders and sales, loans and accounts, patients and students, and myriad other categories. Yet these types of data are far more plentiful and certainly no less important than customer names and addresses.

The main reason we fail to adequately manage quality in these categories is that this data's structure is far more complex and does not allow for a "one size fits all" solution. More effort and expertise are required, and data quality tools offer less help. Until organizations require data quality management programs to focus equally on all of their data, we cannot expect significant progress.

# five

## 5. BAD METADATA

The greatest challenge in data quality management is that the actual content and structure of the data are rarely understood. More often, we rely on the theoretical data definitions and data models. Since this information is usually incomplete, outdated, and incorrect, the actual data looks nothing like what is expected. The solution is to start data quality management programs with extensive data profiling—the term used to describe a collection of experimental techniques aimed at examining the data and understanding its actual structure and dependencies.

### THE GREATEST CHALLENGE IN DATA QUALITY MANAGEMENT IS THAT THE ACTUAL CONTENT AND STRUCTURE OF THE DATA ARE RARELY UNDERSTOOD.

Comprehensive data profiling includes the following components:

- **Subject profiling:** Examines subjects in different tables or on different systems and helps to find where the information about each subject is stored
- **Relationship profiling:** An exercise in identifying entity keys and relationships as well as counting occurrences for each relationship in the data model; necessary to validate existing relational data models or to build them when none are available
- **Attribute profiling:** Examines values of individual data attributes and provides information about frequencies and distributions of their values; helps to identify meaning and allowed values for an attribute
- **Timeline profiling:** Looks for patterns in historical data, such as temporal distribution of the data, patterns of values for different time periods, etc.
- **State-transition model profiling:** Examines lifecycle of state-dependent objects and provides actual information about the order

and characteristics of states and actions; helps build or validate state-transition models

- **Dependency profiling:** Uses various pattern recognition techniques to find hidden relationships among attribute values

# six

## 6. IGNORING DATA QUALITY DURING DATA CONVERSIONS

Data warehouses begin life with data conversions from various operational databases—usually a rather violent beginning. Data conversion often takes the better half of the implementation effort and almost never goes smoothly.

Every system is made of three layers: database, business rules, and user interface. What users see is not what is actually stored in the database, especially in older “legacy” systems. During data conversion, attention is usually on the data structure, and the data is mapped between old and new databases. However, this approach inevitably fails because the business-rule layers of the source systems are poorly understood.

Another problem is the typical lack of reliable metadata about the source database. Consider how often we find that value codes in the data are missing from the mapping documents. It happens all the time. When such a basic component is incorrect, how can we believe any metadata? Yet, over and over again, data conversions are made to specifications built on incomplete, incorrect, and obsolete metadata.

The quality of the data after conversion is directly proportional to the amount of time spent to analyze and profile the data and uncover the true data content. Unfortunately, the common practice is to convert first and deal with data quality later. The ideal data conversion project begins with data analysis, comprehensive data quality assessment, and data cleansing. Only after completing these steps can we proceed to coding transformation algorithms.

# seven

## 7. WINNER-LOSER APPROACH IN DATA CONSOLIDATION

Most data warehouses draw data from multiple operational systems. The need to consolidate data from multiple sources adds the new dimension of complexity to basic data conversion, as the data in the consolidated systems often overlaps. There are simple duplicates, overlaps in subject populations and data histories, and numerous data conflicts.

The traditional approach is to set up a winner-loser matrix indicating which source data element is to be picked up in case of a conflict. For instance, date of birth will be taken from System A if present, from System B otherwise, and from System C if it is missing in both A and B. This rarely works because it illogically assumes that data on System A is always correct. To mitigate the problem, the winner-loser matrix is usually transformed into a complex conditional hierarchy. At some point, the decision tree becomes so complex as to be impossible to manage; it yields good results only for the simple indicative data elements. This approach inevitably fails when you have complex historical data, such as event histories and state-transition histories.

The correct approach to data consolidation is to view it in a similar light as data cleansing. We select one of the data sources as the primary data source for each data element, design a comprehensive set of tests comparing the data against other sources, and then use this additional data for “data cleansing.” Once the data in the primary data source is correct, we convert it to the target database.

# eight

## 8. INADEQUATE MONITORING OF DATA INTERFACES

It is not uncommon for a data warehouse to receive hundreds of batch feeds and uncountable real-time messages from multiple data sources every month. These ongoing data interfaces

usually account for the greatest number of data quality problems. The problems tend to accumulate over time, and there is little opportunity to fix the ever-growing backlog as we strive toward faster data propagation and lower data latency.

Why do the well-tested data propagation interfaces falter? The source systems that originate the feeds are subject to frequent structural changes, updates, and upgrades. Testing the effect of these changes on the data feeds to multiple independent downstream databases is a difficult and often impractical step. Lack of regression testing and quality assurance inevitably leads to numerous data problems with the feeds anytime the source system is modified—which is all of the time!

The solution to interface monitoring is to design programs operating between the source and target databases. Such programs are entrusted with the task of analyzing the interface data before it’s loaded and processed. Individual data monitors use data quality rules to test data accuracy and integrity. Their objective is to identify all potential data errors. Advanced monitors that use complex business rules to compare data across batches and against target databases identify more problems. Aggregate monitors search for unexpected changes in batch interfaces. They compare various aggregate attribute characteristics (such as counts of attribute values) from batch to batch. A value outside of the reasonably expected range indicates a potential problem.

# nine

## 9. FORGETTING ABOUT DATA DECAY

If a caterpillar transforms into a butterfly but remains listed on the finch’s menu as a caterpillar, the bird has a right to complain about poor data quality. In human affairs, people metamorphose by getting married, dying, or simply moving, without filling out all necessary forms to record these events in each system where their data is stored.

If we assume data accurately represents real-world objects, we assume that data collection

processes are perfect. But of course object changes regularly go unnoticed to computers, and thus accurate data can become inaccurate over time, without any physical changes made to it.

In this age of numerous interfaces across systems, we have come to rely on changes made in one database to be propagated to all other databases. However, there are times when this does not happen. For instance, interfaces may ignore retroactive data corrections, or IT personnel may make changes using a backdoor update query—which of course does not trigger any transactions in the downstream systems.

Whether the cause is a faulty data collection procedure or a defective data interface, it is a rather common occurrence for data to get out of sync with reality. The solution to the problem is recurrent data quality assessment and sample comparison against trusted sources. This solution provides information about the rate of decay and shows the categories of data that are most prone to quick decay. Such knowledge can be used to improve data collection procedures and data interfaces.

## ten

### 10. POOR ORGANIZATION OF DATA QUALITY METADATA

Data quality initiatives produce enormous volumes of valuable metadata. Data quality assessment tells us about existing data problems and their effect on various business processes. When performed recurrently, assessment also shows data quality trends. Data cleansing determines causes of errors and possible treatments. It also creates an audit trail of corrections so that, at a later point, we can discover how a particular data element came to look the way it does. Interface monitoring identifies ongoing data problems and tells about data lineage, as does data conversion and consolidation.

Inadequate architecture of the data quality metadata repositories is a common problem in data quality management. Data quality assessment projects routinely generate innumerable

unstructured error reports with no effective way of summarizing and analyzing the findings. Data cleansing initiatives typically lack audit trail mechanisms, and ETL processes often lack data lineage information. As a result, the value of the data quality initiatives is greatly diminished. In the worst cases, the projects are abandoned.

The solution is to design a comprehensive data quality metadata warehouse (DQMDW), which is the collection of tools for organization and analysis of all metadata relevant to or produced by the data quality initiatives. It is a rather complex solution, combining elements of object-oriented metadata repository with analytical functionality of a data warehouse. However, in absence of a well-designed DQMDW, data quality metadata will suffer from the very malady it is intended to cure—poor quality. 🍷

---

*Arkady Maydanchik is a recognized practitioner, author, and educator in the field of data quality and information integration. His data quality methodology and his breakthrough ARKISTRA technology have been used to provide services to many organizations. Arkady is the author of Data Quality Assessment for Practitioners, is a frequent speaker at various conferences and seminars, and contributes to many journals and online publications. He can be reached at [arkadym@dataqualitygroup.com](mailto:arkadym@dataqualitygroup.com).*



# Reinventing Your BI Program

BY STEVE WILLIAMS

**As companies struggle to get value from their BI programs,** the BI professionals we talk with and work with in a range of industries tell us that:

- Business users are unsatisfied with the BI applications that have been deployed and are frustrated with the slow pace of BI delivery.
- Business leaders won't fund BI adequately until the value is proven, and yet the value can't be proven if only small, low-impact projects are undertaken.
- The business community hasn't articulated a BI strategy, and thus IT is stuck in a reactive, report-centric mode—even though there are quick-hitting, effective ways to create or revise a BI strategy.

- IT methods and operating policies that work well for transactional systems and Web applications are impeding the ability to rapidly increment BI with high business impact.
- BI best practices and sound data architecture principles are not being followed, resulting in higher risk and an impaired ability to deliver high-impact BI.

While every symptom may not be present in your enterprise, we know plenty of places where all these symptoms are present.

How can your company unlock the business value that BI—done right—can deliver? Here are some of the fundamentals that we’ve employed to help clients reinvent their BI programs.

## HOW CAN YOUR COMPANY UNLOCK THE BUSINESS VALUE THAT BI—DONE RIGHT—CAN DELIVER? HERE ARE SOME OF THE FUNDAMENTALS THAT WE’VE EMPLOYED TO HELP CLIENTS REINVENT THEIR BI PROGRAMS.



### Create a BI Vision that Supports Your Business Strategy and Core Processes

Many existing data warehouses and BI applications were designed using data-centric and/or report-centric technical methods. As a result, business users feel that BI doesn’t help them be successful in their jobs, whether they are senior executives with P&L responsibility, power users providing analytical and decision support, or functional and operational staff charged with operating and improving the business. This creates a lack of business enthusiasm for BI, limits funding, and all but guarantees underperformance in the BI arena—at a time when more companies are competing by leveraging BI.

To break out of this BI death spiral, companies need to create a compelling BI vision based on BI-driven business improvement opportunities (BIOs) identified via quick-hitting workshops with key business stakeholders. The BIOs should align with your company’s business strategy and core processes, and they can be prioritized based

on business impact, technical complexity, and other relevant portfolio planning considerations. The BI vision also needs to be based on a solid understanding of BI requirements, which are also gathered via quick-hitting workshops. The BI requirements are aligned with the business improvement opportunities, and together they drive a reinvigorated BI program that can deliver real business impact if executed effectively. Creating a compelling BI vision can be done in just a few weeks, and there is really no good business reason not to do so.

### Understand Where You Stand and What It Takes to Deliver on the BI Vision

Whether they have a compelling BI vision or not, many companies that have built data warehouses and delivered BI applications have operated in an uncoordinated, project-by-project fashion using inconsistent methods and tools. One company we know—a \$4+ billion leader in its industry niche—has six BI initiatives under way and each is using a different design and development method. The company lacks an enterprise logical data model, and there is no logical cohesion to how data is being brought into a staging area for the various BI uses. The business refuses to make appropriate investments because BI value hasn’t been proven, and as a result, IT doesn’t have the resources it needs to deliver BI that delivers value—a “chicken or the egg” scenario.

Assuming that a BI vision has been created by the business community as described earlier, the next step in reinventing the BI program is to understand what barriers stand in the way of delivering on the BI vision and what it will take to execute the vision successfully. In the case above, one key barrier is the lack of an enterprise logical data model. Other typical barriers we see include:

- Weak BI governance within the overall IT governance structure
- Poor or nonexistent data governance practices
- Lack of a BI program approach and a BI road map
- Inappropriate data architecture to support the BI vision

- IT operating policies that impede rapid, incremental delivery of high-impact BI
- Inappropriate BI/DW development methods
- Failure to integrate BI into core business practices to create BI impact
- Inappropriate BI infrastructure and technical capabilities
- Lack of understanding by business leaders of what it takes to achieve BI impact

Effective BI practices are very well understood after over a decade of industry-wide, experience-based learning, and thus these barriers need not exist. To address barriers in your company, the key to reinventing your BI program is to understand where you stand in relation to effective BI practices and to be able to communicate to the business leaders what it will take to realize the BI vision.

Understanding where you stand is relatively easy—all it takes is a quick-hitting, properly focused assessment. Explaining what it will take to business leaders is harder, particularly when business people and IT people speak a different language—which is often the case. This argues for having IT-savvy business people and business-savvy IT people on the BI team.

### **Break Down the Barriers and Execute, Execute, Execute**

When you have a compelling BI vision, and when you know what the barriers are, and when you know what it takes to succeed, the next step in reinventing your BI program is to break down the barriers (reduce the risks) and focus like a laser on execution. This is where it gets very situation-specific, because companies are as different as individuals when it comes to change and when it comes to operational effectiveness. Just as some individuals thrive on change and are effective at getting things done, some companies can see the way forward, break down the barriers to change, and execute a BI program that delivers high-impact BI on a regular basis.

To make sure your company is like the hundreds that have used BI to improve their businesses, it is essential to have a BI strategy and BI road

map. Based on the targeted BI-driven business improvement opportunities and the associated BI requirements, and taking into account the barriers that must be overcome, a BI strategy and a BI road map work together to establish and manage all of the BI activities required to realize the BI vision.



**JUST AS SOME INDIVIDUALS THRIVE ON CHANGE AND ARE EFFECTIVE AT GETTING THINGS DONE, SOME COMPANIES CAN SEE THE WAY FORWARD, BREAK DOWN THE BARRIERS TO CHANGE, AND EXECUTE A BI PROGRAM THAT DELIVERS HIGH-IMPACT BI ON A REGULAR BASIS.**

Creating these all-important deliverables sets the stage for success and reduces the business and technical risks that are inherent BI initiatives. More broadly, the BI strategy and BI road map can be thought of as a “BI Reinvention Plan” that is squarely focused on overcoming all the symptoms that drive the need for reinventing your BI program. Armed with such a plan, and coupled with business buy-in, your company can clear the hurdles and truly leverage BI for improved business results. 🍂

---

*Steve Williams is president of DecisionPath Consulting. He can be reached at [steve.williams@decisionpath.com](mailto:steve.williams@decisionpath.com).*

This article appeared in *TDWI FlashPoint* e-newsletter, August 23, 2007. *TDWI FlashPoint* is a TDWI Member benefit. For more information about TDWI Membership, visit [www.tdwi.org/membership](http://www.tdwi.org/membership).



# Does Your Information Strategy Feel Like Dodge Ball?

## Make Data Your MVP.

Understanding and managing your company's data often feels like a game. Getting Business and IT to work together can seem like a contact sport. But you can't afford to sit on the sidelines any longer. You need to get in the game.

Baseline Consulting has the game-winning strategy to help companies enhance the value of enterprise data and improve business performance. We understand that your data is your MVP, and our proven, structured approaches uniquely position us to help clients achieve self-sufficiency in designing, delivering and managing data as a corporate asset.

Companies that treat data as a corporate asset experience significant benefits: improved customer interactions, higher quality decision making, organizational collaboration and more. With Baseline Consulting as your coach and teammate you can finally get off the sidelines and into the game.



- ❖ Business Analytics
- ❖ Data Warehousing
- ❖ Data Management
- ❖ Data Integration



# Striking a Balance in Agile BI Development

BY EVAN LEVY AND  
FERNANDO MARTINEZ-CAMPOS

**Agile development methods have received considerable attention** in the last few years. The quest to deliver business value in tighter time frames has driven companies to reconsider their traditional “waterfall” approaches.

The main thrust of agile development is to limit development to “bite-sized” chunks, reduce communication bureaucracy with self-contained teams, and use a hyper-collaborative, “war room” environment. Agile techniques have been particularly successful when business requirements aren’t well known or change frequently.

As incremental as business intelligence (BI) projects are, many companies persist in using traditional waterfall development for determining core business requirements—capturing all business requirements and delivering a holistic design with official signoffs at the conclusion of each project phase. The consequent delays and lag time for consensus building and approval can translate into the disaffection of end users and executives who naturally expect quick results. The term “agile BI” can also be fraught with some dangerous implications that developers ignore at their peril.

Agile development takes a more immediate approach to deliver value in well-scoped chunks that can be measured in weeks rather than months. Agile development applied to business intelligence projects assumes special characteristics that are often new to BI teams, including:

- Small dedicated teams of four to eight members are located in the same work group setting
- Short delivery time frames, usually within 30–60 days
- Concise, narrowly scoped requirements
- Minimal documentation with an emphasis on delivering the data and reports themselves



## AGILE DEVELOPMENT TAKES A MORE IMMEDIATE APPROACH TO DELIVER VALUE IN WELL-SCOPED CHUNKS THAT CAN BE MEASURED IN WEEKS RATHER THAN MONTHS.

- Project artifacts (such as design documents) are recognized as necessary but might be completed only after functionality has been deployed
- Project plans are refined and managed daily to identify and resolve issues that delay deliverables; each team member is dedicated and accountable for pushing the project forward on a daily basis
- Change control must be in place as the project progresses; enhancements and fixes may be discovered during the project but must be documented for a future project or they will contribute to further delays

Applying agile development to BI projects means that data sources should be understood up front to answer the initial business requirement. To capitalize on the advantages of agile development, we recommend that BI teams perform a comprehensive source-data analysis prior to a 30- or 60-day cycle. The number of unforeseen

surprises—and there are usually many regarding sourcing data for BI—is proportional to the degree of delay and risks bogging down the development effort, in effect eliminating any benefits of an agile approach. As unforeseen issues arise in an agile environment, the project manager modifies the scope of the requirements to adhere to delivery expectations. If partway through the project the schedule begins to slip, the scope of the data loaded is controlled—usually by reducing the number of tables or limiting the history loaded to fewer time periods.

Many agile development projects utilize a server called a “sandbox.” Its purpose is to support a proof of concept to demonstrate the solution with minimal impact to production systems. If the sandbox is successful, then the data is designed and migrated to the production data warehouse that has defined standards and a more robust infrastructure.

IT managers often misunderstand the sandbox concept and open up the sandbox to “keep the users happy” or, more cynically, “to get the business off my back.” Despite being created as a fast development device, the sandbox is often availed to users who—once they have access to it—want to continue using it. Expectations management is essential in agile development, and particularly so for agile BI. The development team must structure and communicate the agile project as being divided into two distinct phases: the sandbox proof of concept and the production phase. Each phase takes time to complete, each has its own unique success measures, and the production phase raises awareness of ongoing data support.

Much has been written about the iterative nature of BI, and it’s important to distinguish between the terms “iterative” and “agile.” Iterative development mandates a fully developed system development life cycle, adhering to a strict, step-by-step development process. Agile development takes this one step further by limiting traditional documentation and design activities in the early stages of the project—usually until after the proof of concept shows results. Once the proof of concept is delivered, the missing documentation and design activities are added to ensure that the application can be moved into production.

## Lessons Learned from an Agile BI Project

Recently, following a high-tech client's agile BI effort, the team evaluated issues that contributed to the project's success and examined roadblocks and barriers. As with all agile projects, your team should create a postmortem report that drives improvements to the processes, staffing, and structure of future agile activities.

Common problems we have found during agile BI projects include:

- New data quality problems are discovered after the data is analyzed for the first time. This may require fixing the source systems or the ETL jobs that loaded the data, contributing to project delays (which are proportional to ETL complexity).
- Additional time is needed to write ETL in an efficient manner (loading only changed data), to gather necessary metadata definitions, and to put jobs into production in a realistic time frame.
- The development time for an agile BI effort is often contingent upon the number of tables. With new sources, the originating application does not have the mechanisms to run the extracts on a daily basis. The source extract process should be held to an official schedule with guaranteed operations support.
- ETL may need to be designed for efficiency. For instance, instead of extracting entire source files, apply the daily changes.
- Production deployment, scheduling jobs, monitoring successful execution, space management, and backups need to be taken seriously. This planning time risks eating into the ambitious agile iteration cycle.
- Documentation must be completed after the proof of concept to facilitate ongoing support. Requirements, metadata, and definitions for one project iteration must be documented before the next iteration begins.

Agile development requires frequent meetings, ongoing collaboration, fast iterations, small development teams, short bursts of testing, and quick turnaround times. There is a lot more

“people time” and many more concentrated working sessions involved in agile projects. If BI requires management support and sponsorship, agile BI requires focused commitment of all parties—including executives—to be effective.



**THE NUMBER OF UNFORESEEN SURPRISES IS PROPORTIONAL TO THE DEGREE OF DELAY AND RISKS BOGGING DOWN THE DEVELOPMENT EFFORT, IN EFFECT ELIMINATING ANY BENEFITS OF AN AGILE APPROACH.**

Although many companies have only recently cemented custom, BI-specific development methods, the agile development strategy has its rightful place in BI programs. It means striking a balance between short lifecycle development and robust infrastructure that BI projects need to be sustainable and robust. It's incumbent upon BI teams to set new expectations of teamwork, refined processes, data re-use, and formal postmortems to make agile BI successful BI. 🙌

---

*Evan Levy, CBIP, is a partner with Baseline Consulting. You can reach him at [evanlevy@baseline-consulting.com](mailto:evanlevy@baseline-consulting.com).*

*Fernando Martinez-Campos, CBIP, is a senior consultant specializing in data architecture and data management. You can reach him at [fernandom@baseline-consulting.com](mailto:fernandom@baseline-consulting.com).*

This article appeared in *TDWI FlashPoint* e-newsletter, October 4, 2007. *TDWI FlashPoint* is a TDWI Member benefit. For more information about TDWI Membership, visit [www.tdwi.org/membership](http://www.tdwi.org/membership).



# TIME TO CREATE AN UNFAIR BUSINESS ADVANTAGE.

The ParAccel Analytic Database™ gives you faster access to data and the ability to utilize your business information more effectively, accomplishing in seconds what used to take hours. Make time *your* competitive advantage, visit us at [www.paraccel.com](http://www.paraccel.com).

PAR ) ACCEL™  
seize > bend > exploit time™



# Tailoring Performance Dashboard Content

BY ROSS MORRISSEY

## Abstract

Performance dashboards must be designed around the specific capabilities and goals of each user—not the capabilities and goals of the dashboard technology provider. Dashboards are clearly different from more traditional interactive reporting, and experts (principally Stephen Few) have drawn on several established disciplines to define the visual vocabulary of dashboards. Tailoring performance dashboard content—effectively choosing the right measurements, ratios, trends, and targets—also draws on established disciplines such as management accounting, performance measurement, cockpit instrumentation, and software design.

Several authors have dramatically changed the way measurements are used to understand and manage processes. The Balanced Scorecard (Kaplan and Norton, 1996) used a dashboard as a central metaphor, which later helped to underscore differences between monitoring machinery and monitoring people. Lessons learned in aviation about supporting situation awareness with user-centered design have parallels in designing dashboards most appropriate for each user—lessons mirrored in contemporary software interaction design. In addition, each of these fields has overcome false economies of advancing underlying technology with a focus on the capability and goals of the user. This approach will work with any dashboard platform, tailoring effective dashboard content for users with differing capability or analysis mindshare and differing personal, strategic, tactical, and operational goals.

## Dashboards, Information, and Objectives

Stephen Few (Few, 2006) provides a comprehensive guide to the information visualization aspects of dashboards, as well as a succinct definition:

A dashboard is a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information can be monitored at a glance.

The task of defining the “most important information” relies completely on the “one or more objectives” and is the focus of this article.

This link between information and objectives is implicit in a number of areas, and four in particular have important contributions to make to the design of dashboards (Table 1).

- **Management accounting** was developed to provide information to help identify and track objectives and production goals, and provides many of the metrics seen in dashboards today.
- Reliance on incomplete proxy measurements in **performance management** can lead to optimization along the wrong dimensions.
- Design aligned with objectives enhances **situation awareness** in airplane cockpits.
- Alignment with user objectives in **software interaction design** enhances the presentation and consumption of information.

All of these disciplines supply design principles and techniques relevant to performance dashboards.

### APPLICATION TO DASHBOARDS

Each of these four fields helps suggest what content should and should not be included in a

performance dashboard. A balance of measurements that monitor and coordinate processes rather than people must be selected. If dashboards are constructed using the latest graphical renderings of the latest financial figures, or reusing charts from the annual report, the resulting dashboard will not (and should not) be used. The dashboard design process must start with an understanding of each user’s capabilities and goals and use this understanding to design a fully customized dashboard.

## Lessons from Management Accounting

Before the industrial revolution, all cotton weaving was done piecework and sold at market. This initial state—with a bale of cotton, spinning wheel, handloom, and finished cloth all at hand—provides the perfect information visualization scenario with the state of inventory, work-in-process, and production capacity apparent instantly. Performance dashboards layer 200 years of technology and attain only a crude proxy for this ideal.

Cotton spinning and weaving were among the first trades to be touched by the industrial revolution. The cotton mill implemented an early example of management accounting, thus enabling large enterprises. Close attention to costing at each stage of production served the goals of the mill managers and supplied operational information like cost per pound—distinct from financial figures provided to the owners of the enterprise (Johnson and Kaplan, 1987).

Over the decades that followed, accounting grew in sophistication and developed rules, procedures, and an infrastructure of its own. The rise of industrial conglomerates increased the importance of presenting information in the form of standardized financial statements to support the goals of investors evaluating

YEAR	AUTHOR	FIELD	MISPLACED FOCUS	GOAL-FOCUSED RESPONSE
1992	Kaplan, Norton	Management accounting	Financial	Balanced Scorecard reflects strategic goals
1995	Austin	Performance measurement	Measurement	Control measurement scope and intent
1988	Endsley	Aviation	Technology	Goal-directed task analysis
1996	Cooper	Interaction design	Developer	Goal-Directed® design

Table 1. Rediscovering goals in four dashboard-related disciplines

prospective assets across many types of businesses. ROI-driven *financial* accounting began to take precedence over *management* accounting in managing operations by the 1920s.

This emphasis on tracking operations using financial accounting reports creates two different types of problem. Financial results can be improved over the short run at the expense of investment in the future viability and growth of the organization. Conversely, many types of investments in the future are not fully reflected in short-term financial measurements. Robert Kaplan and David Norton responded to this disconnect between information and goals with the Balanced Scorecard in the early 1990s.

#### BALANCED SCORECARD INITIATIVE

The Balanced Scorecard gets its name from a balance of financial and non-financial measures, a balance of short-term and long-term indicators, and a balance of leading and lagging indicators.

A Balanced Scorecard typically measures objectives in four perspectives:

- **Learning and capacity for growth:** Measures of employee capability, satisfaction, productivity, and retention
- **Internal product development and operations:** Measures of quality, response time, and new product introductions
- **Customer satisfaction:** Measures of retention and market share
- **Financial results:** Return on investment and economic value added

Early in their work, Kaplan and Norton draw on the dashboard analogy to describe the Balanced Scorecard (Kaplan and Norton, 1996):

Skilled pilots are able to process information from a large number of indicators to navigate their aircraft. Yet navigating today's organizations through complex competitive environments is at least as complicated as flying a jet. Why should we believe that executives need anything less than a full battery of instrumentation for guiding their companies? Managers, like pilots, need instrumentation about many aspects of their

environment and performance to monitor the journey toward excellent future outcomes.

The focus is not on a set of standard measures, but on instrumentation to tell managers about their environment and performance in terms of goals.

#### Lessons from Performance Measurement

Increasing automation and electronic interaction has led to a flood of data that is produced as a by-product of many business processes. Like repurposing financial figures to manage operations or the use of operational data to keep score, it's technically easy to identify over- or under-performing subordinates on a limited set of measures, but this should be done only cautiously. Robert Austin (Austin, 1996) pointed out how quickly and consistently employees adjust their output to improve their score: "Driving out differences between measured performance and desired measurements is the purpose of the system, which means agents are being encouraged—perhaps, in their view, obliged—to subvert the system."

Austin identifies three types of measurement: motivational measurement linked to employee incentives, process measurement linked to process improvement, and coordination measurement for resource allocation. Motivational measurement is highly toxic in Austin's well-documented view, but measurement for process improvement and coordination (information measurement) can be useful. The only difference between the types of measurement is how they are used—the goal for gathering the information.

For example, if we measure that 10 widgets are produced on machine A and we are comparing this to the technology used on machine B, or ensuring that 10 widget boxes are available in the shipping department, this is valid information measurement. If we compare the 10-widget figure against the 10-widget benchmark for bonuses, it is very likely that quality will suffer in the drive to make the 10-widget goal. This same quantity data can be used effectively by individual employees to highlight issues for self-diagnoses—or process improvement—as long as they are confident it is not being used to spy on them (which would link the measurements to incentives). Austin recommends supplying anonymous aggregated process information to

supervisors; any shortfalls in performance must be assessed by walking around and directly interacting with employees.

An example of effective performance measurement would be departmental throughput dashboards at a quality-driven manufacturer. Shortfalls are investigated in person, and the remedies most frequently involve revising production plans and quotes—effectively moving the goalposts—rather than trying to get employees to work faster. Examples of myopic focus on measurement, on the other hand, range from an airplane that was flown into a mountain while the pilots argued about the cause of a flashing dashboard warning (Endsley, Bolté, and Jones, 2003) to the use of nitrogen measurement as an inexpensive proxy for protein measurement in wheat gluten that led to the poisoning of pets after melamine was added to boost nitrogen levels.



## SITUATION AWARENESS PROVIDES CONTEXT FOR ALL DECISIONS AND ACTIONS. IT IS A CONCEPT USED EXTENSIVELY IN AVIATION, MILITARY, PUBLIC SAFETY, AND MEDICAL ARENAS—WHERE LIVES ARE AT STAKE.

Austin also points out the extreme difficulty or impossibility of “perfect measurement” where all relevant perspectives of a process are measured. The temptation is to measure only those perspectives that are easiest or cheapest to measure. This will reward the wrong types of behavior and result in dramatically skewed performance—even while measures indicate success. Austin recommends emphasizing the goal and intent of each measurement and taking steps to see that these goals continue to be supported as measurement programs evolve.

### Design Lessons from Aviation

Situation awareness provides context for all decisions and actions. It is a concept used extensively in aviation, military, public safety, and medical arenas—where lives are at stake. “Situation awareness is being aware of what is happening around you and understanding what that information means to you now and in the future” (Endsley, Bolté, and Jones, 2003). A relatively

benign example: a light comes on adjacent to the gas gauge in a car. This indicates a low-fuel situation; if we don’t divert to get gas, we will run out before we reach our goal. Before any decision is made or action is taken, a certain degree of awareness of the situation must exist. Effective dashboards enhance situation awareness, and situation awareness is the ultimate measure of a dashboard’s effectiveness.

A stunning figure from a frequently cited study of situation awareness is that 88 percent of human error in aircraft accidents was due to poor situation awareness. The remaining 12 percent was due to poor decision-making or poor execution (Endsley, Bolté, and Jones, 2003). In most cases, pilots were doing their jobs correctly, but were let down by their inability to effectively assess the information they received from the environment—including their dashboards.

Operational dashboards frequently include well-understood processes and standard measures; their users have much in common with the pilots, soldiers, and physicians benefiting from information displays designed for situation awareness. A large organization could adopt a formal situation awareness approach to design operational dashboards. This might involve briefly displaying prototype dashboards to employees and testing their knowledge about the underlying situation. With a statistically significant sample of users with identical roles in the organization, a formal iterative development process could result in dramatically more effective dashboards.

Unfortunately, this approach will work only for employees with well-defined roles—roles that stay constant long enough to amortize the development cost (a minority in today’s business environment). Executive (strategic or tactical) dashboards are too company- and role-specific to benefit directly from traditional situation awareness development practices (Resnick, 2003). In addition, these formal practices are very expensive; developing realistic prototypes and testing them (which could easily involve exotic physiological measurements) may continue for several iterations.

Even without the rigor of the formal process, established design guidelines exist in this field. Mica Endsley enumerates 50 situation-awareness design principles (Endsley, Bolté, and Jones,

2003) in six categories, which contain valuable insight for performance dashboard designers. Endsley's general category contains the most widely applicable recommendations:

- **Organize information around goals.** Endsley's situation awareness oriented design approach uses goal-directed task analysis as a starting point. Eliciting the user's goals and defining objectives and tasks supporting these goals helps to identify the information required and organize it for quick comprehension. The automobile dashboard is not organized around the mechanical systems; it is organized around the goal of reaching a destination safely.

The gas gauge may not accurately measure the gas level, but it provides sufficient time and warning as the needle gradually moves to empty—directly supporting the overall goal. If the goal were close monitoring of gas levels and consumption rate, different instrumentation would be used.

- **Support comprehension directly.** Displays should indicate the relative importance of different measurements, indicate their target values, and indicate how they deviate from their targets.
- **Provide assistance for projections.** Presenting trend information or projections can help the user create a mental model of future behavior.
- **Support global situation awareness.** Users can lose sight of the big picture when their dashboard excludes important overall information as they focus on exploring a specific goal.
- **Support trade-offs between goal-driven and data-driven processing.** Dashboards need to support the current goals and tasks of the user without distracting them with unwanted or distracting alerts based on thresholds, unless they are more important than the current task.
- **Make changes in context apparent.** If the situation suddenly changes because of some external factor, it is important to alert users in a way that helps them quickly understand the new situation.

- **Use information filtering carefully.** Restricting the user to certain dashboard content based on the current situation to reduce the bandwidth required is both very difficult to do automatically and can take away from the user's global perspective. It is better to let the user filter out unwanted information manually.

Endsley also offered these recommendations in other categories:

- **Confidence and uncertainty.** To allow the user to develop confidence in the dashboard, missing, uncertain, or stale measurements must be represented, and any factors that influence the availability of or confidence in the data should be included.
- **Dealing with complexity.** The total number of elements displayed—how much data is displayed—along with the number of different goals or tasks supported by a dashboard significantly increase its complexity. By reducing information to essential measurements and grouping measurements by task or goal within highly tailored dashboards, complexity can be significantly reduced, enhancing situation awareness.

These design principles are used to help reach solutions more quickly, but the high-stakes nature of the disciplines where situation awareness is critical mean that the formal prototyping and testing cycles will remain necessary components of this approach, which limits its formal adoption in the area of performance dashboards.

## Lessons from Software Interaction Design

Contemporary software interaction design addresses many of the same issues as design for situation awareness. The problem with most software usability is that the software conforms to implementation models; the user experience is often designed by programmers who leverage new tools and technology to reuse code where possible to simplify and shorten the development cycle. Alan Cooper (Cooper, Reimann, and Cronin, 2007) has developed a practical methodology for interaction design, also based around the goals of the user.

Cooper's approach involves creating *personas* representing each class of user, then developing

*scenarios* that explain how each persona would accomplish their goals. Cooper identifies several types of goals, including user goals, customer goals, business and organizational goals, and technical goals. It is essential that technical goals such as hardware and software constraints are subservient to other types of goals to ensure that a design is successful. This de-emphasis of the technical perspective precisely mirrors design for situation awareness.

Cooper's focus on user capabilities and goals is clear:

There is no such thing as an objectively good user interface—quality depends on the context: who the user is, what she is doing, and what her motivations are.

Cooper places personas and scenarios at the center of an iterative design process that begins with research and ends with design elements and framework. The research phase of the design begins by establishing an understanding of the business scope and goals of the project. Interviews with stakeholders provide context. Potential users undergo interviews and direct observation with an emphasis on their needs, capabilities, and goals.

Research exposing the patterns of behavior and the motivations of clusters of users is described in Cooper's personas—descriptive models of representative users. A persona captures the motivation for user behaviors as well as the behaviors. Users have different motivations for looking at a gas gauge, so the gas gauge designed to prevent users from running out of gas on the highway is different from the gas gauge designed to closely monitor fuel levels and consumption.

This persona approach is ideal for performance dashboards. The nature of the environment and toolsets make it possible to tailor dashboards by designing for one persona and using departmental and individual filters or views of data to produce customized dashboard content for each user in the persona cluster. This design will accommodate the capabilities and support the goals of these users better than any "all-purpose" dashboard.

Cooper identifies six types of goals, and all need to be addressed to design a successful product—a product people will want to use. There are three types of user goals, independent

of context. User *experience goals* are met with excellent visual quality and crisp interaction, and significantly affect perception of the product. User *end goals* are concrete tasks and objectives, and user *life goals* shape the user's view of these tasks.

Customer, business, and technical goals are secondary to user goals. *Customer goals* must be met for the dashboard to be a viable product that would be purchased or installed. *Business goals* are the result of successfully meeting the user's end goals and are reflected in performance or balanced scorecard measures. *Technical goals* are the last element to be considered in a design, after all the other goals are addressed. The user cares least about optimizing memory use or programming approaches. Giving technical goals a prominent role will negatively influence interaction design.

User experience, customer, and technical goals will be addressed during the selection of a performance dashboard platform. Effective content must support user end and life goals in addition to business goals.

The power of the persona approach is in the development of scenarios to clarify and rank goals and objectives. The most common scenarios of users in their persona cluster are identified. Then the scenarios are played out; as each persona undertakes a task or responds to an exception, their ideal performance dashboard must help them meet their goals—all their goals. This is when our vocabulary of dashboard graphic techniques and library of key performance indicators is used, and not before.

### **Dashboard Design Process and Analysis Mindshare**

An organization may have hundreds or thousands of individual roles with different information demands, and effective dashboards will deliver exactly the mix of long-term goal and short-term task-related information required. Just as dashboards for drivers differ from dashboards for pilots, each role in an organization should have a specific dashboard. Performance dashboards help employees perform their jobs well. Patterns of dashboard usage at different levels of organizations tend to follow a distinct analysis mindshare curve. Identifying the place of each persona cluster in relation to

this curve can speed the effective selection of dashboard content and level of detail.

Analysis mindshare is the cumulative attention an individual devotes to analytical activities. This is the product of the share of an employee's time spent observing and interpreting performance measures multiplied by the cognitive load or depth of attention required for this analysis. In a large organization, there is a continuum of responsibility starting with the line worker responsible for putting widgets in boxes, extending to the CEO. At both ends of this organizational hierarchy, primary job functions do not leave a significant amount of time for data analysis mindshare (Figure 1).

At the bottom of the organization, the focus is on specific production tasks; at the top, the focus is on strategy. Neither of these individuals can afford the time to sift through the enormous amounts of data produced by today's enterprises. Sifting is the job of business analysts and managers in the middle of an organization with the appropriate skills and bandwidth. Davenport and Harris (Davenport and Harris, 2007) call these three groups senior management, professional analysts, and analytical amateurs (or "everyone else").

The ideal dashboard interface to analytics for top executives will indicate their success in conveying strategy and will have low cognitive load, like a simplified dashboard consisting of basic charts.

There are exceptions, such as highly analytical chief executives with the background and bandwidth for more detailed analysis, or organizations that *compete on analytics* as Thomas Davenport recommends (Davenport and Harris, 2007). This relatively small proportion (10 percent, per Davenport) would have a truncated analysis mindshare curve, but certainly, in an organization of significant size, middle management would be doing more analysis work than upper management, who must focus on strategy.

Performance dashboards are not limited to management. Line workers at the bottom of the hierarchy can benefit from quick access to performance metrics, but only to the extent that they directly support production tasks. Taking time to interpret complex and unfamiliar displays of information will detract

from "front-of-mind" priorities such as quality, customer satisfaction, and safety. These analytical amateurs, such as senior management, have day-to-day goals divergent with detailed analysis—they have little analysis mindshare.

The status of analysis as a core activity increases with professional analysts at the middle of the organizational hierarchy. Line managers and executives increasingly rely on sophisticated analysis to identify areas for organizational improvement and to help support upper management in more strategic roles.

Before the advent of dashboards, this sort of analysis was conveyed casually at the water cooler or in prepared formal presentations (Figure 2). Initially, this analysis was performed manually and was limited to the management accounting function. Using traditional information-gathering techniques such as operational

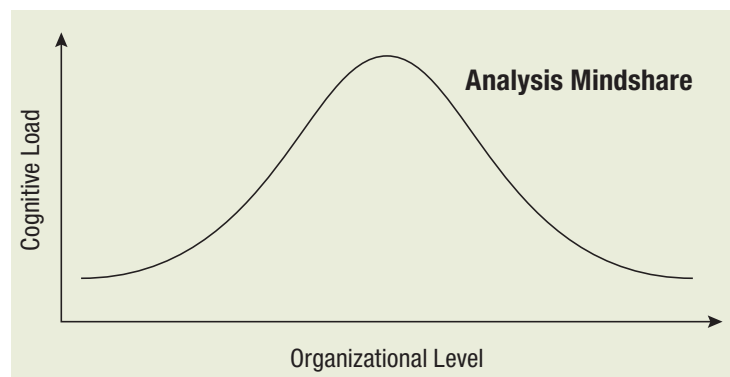


Figure 1. The analysis mindshare curve indicates higher tolerance for the increased cognitive load of detailed analysis near the middle of the organization.

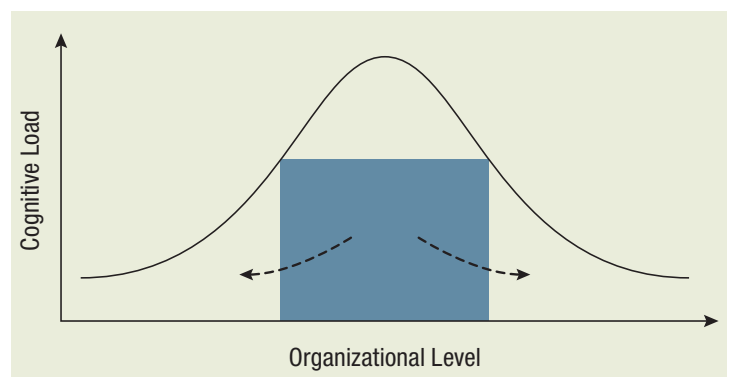


Figure 2. Traditional analysis occurs in the middle of the organization and is disseminated offline or informally up and down the organization.

reporting driving spreadsheet analysis results in a *round robin* query cycle involving resources from different departments, lengthening in time as the investigation continues.

OLAP analysis of aggregated operational data effectively supports process management, and a dashboard is an effective device for tracking operations and alerting people to nonoptimal situations. Because the manager is deeply familiar with the processes and their area of responsibility, a tailored context-specific dashboard can expose more information without overwhelming the viewer and has the potential to replace a number of disparate sources as the primary information source.

Understanding analysis mindshare at different levels of an organization and how differences combine with varied goals can lead to

dashboards that are both supportive of goals and appropriate for the level of attention that can be devoted to them.

### STRATEGIC GOALS

Strategic dashboards (Figure 3) measure progress toward strategic objectives; they help align the organization to strategy in ways that static mission statements cannot. An executive-level dashboard might reflect enterprisewide strategic goals and corresponding key performance indicators (KPIs). Strategic dashboards bring alignment throughout the organization by showing progress toward strategic goals on an organization-, division-, or even department-wide basis, depending on how data is filtered. Strategic dashboards are typically highly summarized, highly graphical, and less frequently updated, and include global, external, trend, and growth measures.

Strategic dashboards can bring an organization's goals to life as they cascade down from the level of the entire organization to the departmental level. A clear explanation of what each goal is, and an idea of how the users are faring toward the goals and against their peers, is a much better motivation than a set of goals engraved in the lobby or printed in an annual report.

The term “cascading” describes how organization-wide goals are reflected or filtered in strategic dashboards moving down the organizational hierarchy. By basing a subordinate group's goals on the goals of the group above, organization alignment is improved. Because these dashboards are used throughout the organization, they should be expressed in ways suitable for all levels of analysis mindshare. Use simple elements and avoid specific numbers unless they are uniformly well understood throughout the organization. Sharing the same simple graphic or metric helps employees share their observations, which helps foster an open, team-oriented environment.

### TACTICAL GOALS

Tactical dashboards (Figure 4) measure short-term trends and progress toward strategic initiatives or special projects. They take advantage of awareness of context and the sophistication of professional analysts to present significantly more detail without sacrificing comprehension. The emphasis

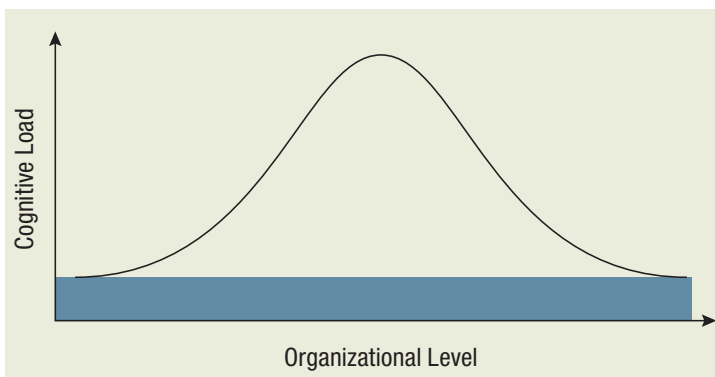


Figure 3. Strategic dashboards use the simplest dashboard elements to convey progress toward strategic goals throughout the organization.

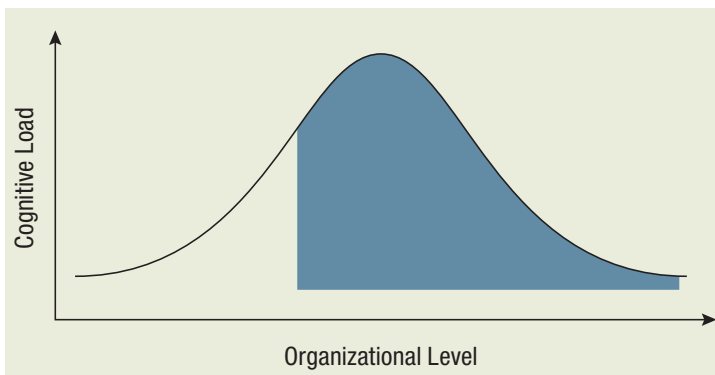


Figure 4. Tactical dashboards leverage the tolerance for cognitive load in the middle of the organization to support more complex and numerous dashboard elements.

is on highlighting opportunities or identifying risks. They frequently involve just one of the four Balanced Scorecard perspectives.

Starting to move away from the stoplight model, tactical dashboards frequently include summary data as well as visual indicators and make full use of hyperlinked OLAP tools, which allows drill-down and root cause analysis. Many dashboards that are created without a formal approach are actually tactical dashboards that aim to maximize profit or increase sales.

The ideal tactical dashboard will provide insight that is detailed enough to enable analysts to identify major leaps in performance through reengineering or other dramatic shifts in process. Analytic tools tied to performance dashboards will support multi-pass reporting to rank over- and under-performers and to track trends across periods. The emphasis is on ratios and targets, and will be unique for any organization, tailored to that organization's specific strategic goals.

#### OPERATIONAL GOALS

Operational dashboards (Figure 5) are used to monitor business or manufacturing processes in near real time with the aim of intervening quickly to resolve issues or take advantage of opportunities. Operational dashboards are usually departmental in scope, and absolute values and thresholds based on averages and norms are frequently more important than trends.

Like tactical dashboards, the focused nature of operational dashboards allows more detailed information to be displayed. It would be unusual for a top-level manager to use an operational dashboard; simple graphics summarizing operational capacity trends would be more appropriate.

Operational dashboards typically display very standard measures that are widely used in the industry. A dashboard from one company might closely resemble that of a competitor, just as the dashboard of a Ferrari shares much with the dashboard of a Jeep. Advantages here include a common context and vocabulary, which leads to easy training for new or transferred employees.

Because the context is so well established, it is sometimes reasonable to add noise to a dashboard element to mimic a physical display more closely—something that would increase

the cognitive load for anyone not familiar with the operational domain. Embellishing a pressure indicator in a dashboard to make it a photo-realistic representation of an instrument on the plant floor can reduce cognitive load. If the dashboard meter does not closely resemble the physical meter, any benefit is lost and cognitive load is increased.

A key element of operational dashboards is the scope of data displayed; by filtering carefully, information is restricted to that required to improve the way each individual performs his or her task. If employees have a clear idea how the critical measures surrounding their tasks are changing, they can do a better job of adjusting their work, catching problems before they are reflected in the finished product.

#### Tailoring Dashboard Content

The process of tailoring dashboard content has three steps: (1) identifying personas; (2) identifying goals and establishing baseline capability for each persona; and (3) selecting strategic, tactical, or operational dashboard content aligned with these goals. Dashboard users will either self-identify with a well-defined persona cluster, or be automatically placed in a persona cluster through filtering of content based on department, location, or role. User goals and business goals are identified and ranked for each persona. The level of responsibility of the persona will help identify its relative capability in terms of analysis mindshare. This will drive the selection of appropriately detailed representations of the measures used to help the persona achieve their most important strategic, tactical, or operational goals.

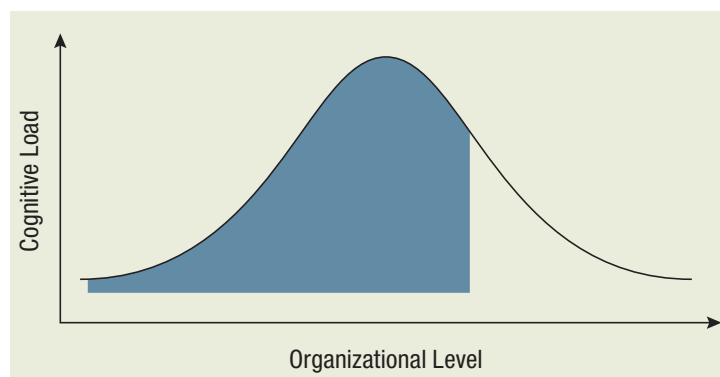


Figure 5. Operational dashboards display standard measures in the detail and complexity required for the individual role.

## Summary

An organization may have hundreds or thousands of individual roles with different information demands, and effective dashboards will deliver exactly the mix of long-term goal and short-term task-related information required. Just as dashboards for drivers differ from dashboards for pilots, each role in an organization will have different performance dashboards. Effective performance dashboards are very specific, and tailoring dashboard content appropriately requires a disciplined approach and an appreciation of the range of visual elements, the different types of dashboards, and the ways they are used to surround and support the capabilities and goals of each employee. 🌟

## References

- Austin, Robert D. [1996]. *Measuring and Managing Performance in Organizations*, Dorset House Publishing Company, Inc.
- Cooper, Alan, Robert Reimann, and David Cronin [2007]. *About Face 3: The Essentials Of Interaction Design*, John Wiley & Sons.
- Davenport, Thomas H., and Jeanne G. Harris [2007]. *Competing on Analytics: The New Science of Winning*, Harvard Business School Press.
- Endsley, Mica R., Betty Bolté, and Debra G. Jones [2003]. *Designing for Situation Awareness: An Approach To User-Centered Design*, Taylor & Francis.
- Few, Stephen [2006]. *Information Dashboard Design: The Effective Visual Communication of Data*, O'Reilly Media, Inc.
- Johnson, Thomas H., and Robert S. Kaplan [1987]. *Relevance Lost: The Rise and Fall of Management Accounting*, Harvard Business School Press.
- Kaplan, Robert S., and David P. Norton [1996]. *The Balanced Scorecard: Translating Strategy into Action*, Harvard Business School Press.
- Resnick, Marc L. [2003]. "Situation Awareness Applications to Executive Dashboard Design," Proceedings of the Human Factors and Ergonomics Society 47<sup>th</sup> Annual Conference, Human Factors and Ergonomics Society, pp. 449–453. [http://www.entrepreneurship.fiu.edu/downloads/marc\\_resnick/Research/Situation%20awareness%20applications%20to%20executive%20dashboard%20design.pdf](http://www.entrepreneurship.fiu.edu/downloads/marc_resnick/Research/Situation%20awareness%20applications%20to%20executive%20dashboard%20design.pdf)

---

*Ross Morrissey, CBIP, is principal consultant for The Ross Morrissey Group, LLC. He can be reached at [rm@rossmorrissey.com](mailto:rm@rossmorrissey.com).*



# Predictive Analytics: Slow Adoption Despite Big Benefits

BY STEPHEN SWOYER

**Few technologies provide as much ROI bang** for the buck as predictive analytics. Despite its impressive benefits, only 21 percent of respondents to a recent TDWI survey had fully or partially implemented predictive analytic solutions in their organizations, while another 19 percent were in the process of developing them.

Why does adoption continue to lag?

Predictive analytics is a deceptively simple way of describing a set of business intelligence (BI) technologies that help uncover relationships and patterns within large volumes of data. That's the first part of the predictive analytic value proposition.

The second part (and probably the most important part from the perspective of business decision makers) is that predictive analytic tools uncover actionable insights. The relationships they discover can, in turn, predict behaviors or events. It's in this respect that TDWI Research contrasts the forward-looking view afforded by predictive analytic solutions with the historical perspective (that of a "rearview mirror,"

according to many predictive analytic vendors) afforded by BI tools, which are typically employed less as predictive and more as deductive technologies.

"[O]ther BI technologies—such as query and reporting tools, [OLAP] tools, dashboards, and scorecards—examine what happened in the past," writes Wayne Eckerson, director of TDWI Research, in a recent report (available at <http://tdwi.org/research/display.aspx?ID=8364>).

These tools are deductive, Eckerson argues, because "business users must have some sense of the patterns and relationships that exist within the data based on their personal experience."

Rearview mirrors or no, conventional BI tools enjoy much greater adoption than their forward-looking counterparts, at least right now.

The good news is that 44 percent of TDWI's survey respondents were still exploring their options with respect to predictive analytic tools. That means about 84 percent of respondents were implementing, investigating, or at least

nominally open to predictive analytic deployments; just 16 percent had no plans to deploy the technology at all. That's still a puzzling adoption rate for a technology that—if case studies, ROI success stories, and marketing anecdotes are correct—has an enviable track record.

“[P]redictive analytics can yield a substantial ROI. Predictive analytics can help companies optimize existing processes, better understand customer behavior, identify unexpected opportunities, and anticipate problems before they happen,” Eckerson writes. For six years running, he points out, a majority of TDWI's annual Leadership Award winners have used predictive analytic solutions to achieve noteworthy business results.



## PREDICTIVE ANALYTICS IS A DECEPTIVELY SIMPLE WAY OF DESCRIBING A SET OF BUSINESS INTELLIGENCE TECHNOLOGIES THAT HELP UNCOVER RELATIONSHIPS AND PATTERNS WITHIN LARGE VOLUMES OF DATA.

### Fear and Trepidation

Which begs a particularly insistent question: Why is a high-value technology like predictive analytics so paradoxically under-represented in the enterprise?

For a number of reasons, thought leaders say, starting with the technology's esoteric roots in statistical analysis.

“Predictive analytics is also an arcane set of techniques and technologies that bewilder many business and IT managers,” Eckerson points out. “It stirs together statistics, advanced mathematics, and artificial intelligence and adds a heavy dose of data management to create a potent brew that many would rather not drink!”

The upshot, experts say, is that it isn't unusual for business users to experience a kind of fear and trembling when they're exposed to predictive analytic tools: such tools, for all of their wonder-working potential, can also be intimidating—perhaps even awe-inspiring.

Perhaps one should say *were* intimidating or awe-inspiring. In fact, proponents say, the current crop of tools eliminates many of the ease-of-use issues traditionally associated with predictive analytics.

“We've always seen a lot of interest coming from the business side. If you go back to the early days of predictive analytics, we used to see people with real visions in areas like marketing, where problems could be tackled very well with predictive models, and they were interested, but maybe they were a little intimidated, too,” says Colin Shearer, senior vice-president of market strategy with predictive analytic SPSS.

“So, yes, maybe there was this fear of actually using the technology—this fear of something novel or untested. A kind of intimidation [in the face] of the tools. But there's no longer a fear factor of getting involved in something novel, and the tools are getting easier and easier to use.”

Along with rival SAS, SPSS controls a sizable slice of the analytics market and a disproportionate share—given its comparatively small BI market footprint—of the predictive analytics space. According to market watcher International Data Corp. (IDC), SPSS derives almost 90 percent of its revenues from sales of advanced analytics software, for good reason: both SPSS and SAS started out as developers of statistical analysis software; both remain highly respected analytics vendors to this day.

Predictive analytics is, in a sense, a college professor's or researcher's statistical package of choice gussied up—that is, made both safe and relevant—for business users. For this reason, SPSS and SAS have benefited from growing interest in predictive analytic solutions.

“SPSS' focus on its concept of the predictive enterprise, which emphasizes forward-looking analysis of customer and operational data, enabled the company to improve its growth rate in 2006 to 12.3 percent,” wrote IDC analysts Dan Vesset and Brian McDonough in a recent market research report.

More to the point, Vesset and McDonough said, SPSS is well positioned to benefit from emerging predictive analytic trends, too—particularly a growing need to access and analyze unstructured content. “SPSS has also emphasized its

capabilities for the text mining of unstructured content, a functionality that further enhances its ability to address CRM analytics needs,” the IDC analysts observed.

### Groundswell Ahead?

Both SPSS and SAS—along with many of the big BI vendors also touting predictive analytic capabilities—are poised to benefit from a predictive analytic groundswell because, they argue, they’re delivering tools that address the very specific expectations and requirements of business users, and not (as was traditionally the case) those of power users.

According to Shearer, for example, the latest version of SPSS’ BI and statistical analysis platform builds on the user-friendliness of its predecessor (which introduced a degree of user self-service capabilities) by delivering fully revamped Mac OS and Linux client versions, too. What’s more, he argues, mounting interest from a nontraditional customer segment—namely, enterprise IT departments themselves (as distinct from individual business units)—provides an indication of just how prevalent predictive analytic technology is becoming.

“What is an interesting phenomenon is more and more we see inquiries coming and serious interest being shown from the IT side as well. More and more we see this as becoming a category that IT are realizing that there’s a lot of interest in the business, [so they’ll have] to think about adopting predictive analytics seriously. ... [They’ll] have to take a look at it,” he says.

According to TDWI’s Eckerson, one of the principal concerns of business managers and other decision makers isn’t so much the efficacy or the usability of predictive analytic solutions, but their appropriate (or most effective) initial inflection point: namely, where should they start?

“Most have only a vague notion about the business areas or applications that can benefit from predictive analytics,” he writes. “[M]ost don’t know how to get started: whom to hire, how to organize the product, or how to architect the environment.”

The most common inflection point is in marketing, Eckerson and others say. Other important applications include budgeting and

forecasting, fraud detection, demand planning, customer service, quality improvement, surveying, and supply chain management, according to TDWI’s research.

**“PREDICTIVE ANALYTICS...STIRS TOGETHER STATISTICS, ADVANCED MATHEMATICS, AND ARTIFICIAL INTELLIGENCE AND ADDS A HEAVY DOSE OF DATA MANAGEMENT TO CREATE A POTENT BREW THAT MANY WOULD RATHER NOT DRINK!”**

There are further wrinkles here, too, because while SAS, SPSS, and others claim to have licked the usability problem, there nevertheless remain a number of significant barriers to successful predictive analytic deployments. Experts cite the importance of developing (and sticking to) a multistage process (starting with the codification of a clear project definition), not to mention the half imaginative-artistic, half logical-scientific task of actually building predictive models (for which the hiring of business-savvy analysts is seen as a must), as two particularly notable challenges.



“[M]ost experts agree that predictive analytics requires great skill—and some go so far as to suggest that there is an artistic and highly creative side to creating models—most would never venture forth without a clear methodology to guide their work,” Eckerson explains.

---

*Stephen Swoyer is a technology writer based in Athens, GA. You can contact him at [stephen.swoyer@spinkle.net](mailto:stephen.swoyer@spinkle.net).*

This article appeared in *BI This Week* e-newsletter, November 7, 2007. For more information or to subscribe, visit [www.tdwi.org/publications/newsletters](http://www.tdwi.org/publications/newsletters).



# BI Solutions

## Transforming Technologies

Our sponsors present their solutions in the following business intelligence categories:

- Analytics and Reporting
- Dashboards, Scorecards, and Visualization
- Data Governance
- Data Integration
- Data Warehousing
- Enterprise Information Strategy
- Open Source BI

# baseline

## C O N S U L T I N G

### BI CATEGORY: ENTERPRISE INFORMATION STRATEGY

More data, more users. More platforms. More business questions. No wonder you so often feel stuck on the sidelines of the information strategy game. Baseline Consulting understands your data challenges and provides both advisory and implementation services to help you design, deliver, and manage the four cornerstones of the enterprise information strategy.

### BUSINESS ANALYTICS

Baseline considers BI more than just a tools issue. Our business analytics engagements are requirements-driven and iterative—from data acquisition to usage.

### DATA WAREHOUSING

Your data warehouse is both strategic business enabler and technology platform. Baseline's specialized skills and repeatable approaches for requirements gathering, data design, data migration, and application deployment ensure high user adoption and successful data warehouse delivery.

### DATA MANAGEMENT

Smart companies position data management as an enterprise-wide service to all business units, systems, and applications. Baseline sees data quality, metadata management, and profiling as some of the many functions in need of formalized processes and sanctioned organizational structures.

### DATA INTEGRATION

Data integration may arguably be your most complex data challenge. And Baseline is ready with interrelated tactics and repeatable processes to automate the cleansing, synchronization, integration, and propagation of corporate data among multiple systems and applications.

Our commitment is your self-sufficiency after we leave in managing and using data as a corporate asset. It's time to get in the game.

### Baseline Consulting

15300 Ventura Blvd., Ste. 523  
Sherman Oaks, CA 91403  
818.906.7638  
[www.baseline-consulting.com](http://www.baseline-consulting.com)



**BI CATEGORY: DATA GOVERNANCE**

DataFlux enables organizations to analyze, improve, and control their data through an integrated technology platform. With DataFlux enterprise data quality and data integration products, organizations can more effectively and efficiently build a unified view of customers, products, suppliers, or any other corporate data asset. A wholly owned subsidiary of SAS ([www.sas.com](http://www.sas.com)), DataFlux helps customers rapidly assess and improve problematic data and build the foundation for enterprise data governance. Effective data governance delivers high-quality information that fuels successful enterprise efforts such as risk management, operational efficiency, and master data management (MDM).

In addition to its award-winning technology, DataFlux offers methodologies that allow companies to attack data problems at a procedural level. This allows organizations to identify the root cause of bad data—and to implement the people, processes, and technology necessary for building more consistent, accurate, and reliable data. Through the combination of technology, methodologies, and best practices, you can elevate the role of data quality to an ongoing corporate activity designed to develop and maintain high-quality information.

To learn more about DataFlux, please visit [www.dataflux.com](http://www.dataflux.com).

**DataFlux Corporation, a SAS Company**

940 NW Cary Parkway, Ste. 201  
919.447.3000  
Fax: 919.447.3100  
[www.dataflux.com](http://www.dataflux.com)



High Performance Data Integration

**BI CATEGORY: DATA INTEGRATION**

**Why do customers choose ETI Solution® data integration platform over other products?**

**LOWEST TOTAL COST SOLUTION**—ETI Solution dramatically reduces the cost, complexity, and risk of major integration efforts. There is no other product on the market with greater flexibility that combines the ability to quickly generate highly reliable, scalable, fully documented integration processes without the administrative overhead or financial cost of licensing proprietary runtime software.

**HIGH PERFORMANCE AND SCALABILITY**—No limits on data volume or project complexity. Since data flows directly between sources and targets—and not through an engine—you never experience data bottlenecks.

**NATIVE MAINFRAME SUPPORT**—Native access to mainframe data protects the company's investment in legacy systems while allowing for new architectures and rapidly changing demands for information.

**BI-DIRECTIONAL INTEGRATION**—Helps keep your databases in sync with a two-way flow of information between systems. ETI Solution allows any database, file structure, or data storage system to be either a target or a source (or both).

**FLEXIBLE**—Your ability to adapt quickly to change today is key to your BI success tomorrow. ETI Solution handles tomorrow's problems as well as today's, and the cost savings compound as you add more systems, more conversions, and more projects.

**ETI**

816 Congress Ave., Ste. 1450  
Austin, TX 78701  
512.383.3000 or 800.856.8800  
[www.eti.com](http://www.eti.com)



#### BI CATEGORY: ANALYTICS AND REPORTING

Information Builders' flagship product—WebFOCUS—is the industry's most widely utilized business intelligence (BI) platform. It connects to any enterprise system or application and lets users interact with information in a simple and intuitive way. Most companies access and analyze less than 20 percent of their data and deploy BI to fewer than 10 percent of their internal users. Information Builders' WebFOCUS breaks these barriers by making all of your applications work together to deliver critical insight to every employee in an organization and beyond. And, WebFOCUS isn't just our tool, it becomes your application. That's why it runs the world's largest BI deployments, sometimes reaching hundreds of thousands of people. Its superior architecture enables WebFOCUS to create a solution for every need in the extended enterprise—executive, analytical, operational, and beyond—and provides the agility to adapt to changing business conditions. Developed with all users in mind, regardless of technical skills, it is the most secure, scalable, and flexible solution in the market.

#### Information Builders

Two Penn Plaza  
New York, NY 10121-2898  
212.736.4433  
Fax: 212.967.6406  
[www.informationbuilders.com/go/BI2008](http://www.informationbuilders.com/go/BI2008)



#### BI CATEGORY: DASHBOARDS, SCORECARDS, AND VISUALIZATION

#### MicroStrategy Ushers in New Caliber of Information Dashboards—Dynamic Enterprise Dashboards

MicroStrategy provides integrated reporting, analysis, and monitoring software that enables leading organizations worldwide to make better business decisions every day. With an easy-to-use Web interface, MicroStrategy allows business users to seamlessly access enterprise data with mission-critical reliability. Companies choose MicroStrategy for its advanced technical capabilities, sophisticated analytics, and superior data and user scalability.

With its Dynamic Enterprise Dashboards, MicroStrategy is first to tightly integrate an industrial-strength BI platform with advanced dashboard visualization and animation capabilities. These information dashboards are an exciting innovation in the BI industry, empowering business users of all levels to interact with their data in compelling new ways. Dynamic Enterprise Dashboards are highly responsive, providing a wide range of information views and increased information exploration options through integration with Adobe Flash. Business users can intuitively flip through many different perspectives of corporate performance without ever leaving the dashboard, allowing them to quickly and easily identify problems and diagnose root causes.

MicroStrategy continues to stay on the leading edge of technology, supporting the most demanding BI applications in the world.

Visit [www.microstrategy.com/dashboardinfo\\_tdwi](http://www.microstrategy.com/dashboardinfo_tdwi) to see our dashboards in action.

#### MicroStrategy

1861 International Drive  
McLean, VA 22102  
703.848.8600  
Fax: 703.848.8610  
[www.microstrategy.com/dashboardinfo\\_tdwi](http://www.microstrategy.com/dashboardinfo_tdwi)



## BI CATEGORY: OPEN SOURCE BI

myBiQ's BI framework can completely replace traditional business intelligence solutions and provides the infrastructure for group-wide controlling, a management information system (MIS), an executive information system (EIS), or corporate performance management (CPM).

myBiQ is based on a range of so-called best-of-breed solutions from the open source community. We have integrated these solutions into a unique, innovative, and easy-to-manage framework. In doing so, we have added diverse modules and developed the framework to the point that it can be optimally used by enterprises.

myBiQ provides all-important functions:

- data consolidation
- reporting
- cost and profitability analysis
- planning and budgeting
- strategy and performance management

The following components constitute the technical basis of myBiQ:

- relational and multidimensional databases and analysis methods (ROLAP, MOLAP)
- an ETL (extract, transform, load) tool for seamless data integration
- interfaces for integrating myBiQ into existing systems (e.g., ERP)

User interfaces such as cockpits and reporting tools, as well as the integration of Microsoft Office Excel, ensure that even employees without programming skills can use myBiQ.

myBiQ was successfully implemented at the EOS Group (part of the OTTO Group) in 2006–2007, and we have been developing it further ever since.

### myBiQ

12 Lozier Road, Mount Olive  
Budd Lake, NJ 07828-1128  
[www.mybiq.com](http://www.mybiq.com)



## BI CATEGORY: DATA WAREHOUSING

### PUT TIME ON YOUR SIDE WITH PARACCEL

With faster access to data and the ability to accomplish in seconds what used to take hours, you'd have time to make better-informed business decisions. Enter ParAccel, the leader in analytic query performance. The ParAccel Analytic Database™ accelerates the process of business intelligence, giving you an undeniable business advantage.

### TIME TO TAKE CONTROL OF YOUR BUSINESS

ParAccel has developed a better solution for the many challenges created by analytic workloads. Our cutting-edge software combines proven technologies—a columnar database, adaptive compression, massive parallelism and optional all-in-memory processing—with engineering innovations to deliver record-shattering performance and scalability without sacrificing reliability. We add velocity to the entire range of decision support and complex processing workloads—from traditional data warehousing, analytics, and ad hoc query to time-critical operational BI and reporting.

### TIME FOR BIGGER BUSINESS IMPACT

Without expensive migration or manual tuning, ParAccel's schema-neutral design increases your analytic query performance by orders of magnitude and extends analytic capacity from hundreds of gigabytes to hundreds of terabytes. ParAccel runs on standard hardware from all major vendors. It is available as enterprise software or a preconfigured appliance and runs stand-alone or as a BI accelerator for existing applications. Make time your competitive advantage—visit [www.paraccel.com](http://www.paraccel.com).

### ParAccel

9605 Scranton Road, Ste. 625  
San Diego, CA 92121  
408.861.3513  
Fax: 858.526.1573  
[www.paraccel.com](http://www.paraccel.com)



#### BI CATEGORY: DATA WAREHOUSING

SAND/DNA allows you to keep all the data you need at your fingertips, with minimal storage requirements.

As data warehouse sizes continue to grow from terabytes into the petabyte, it becomes a greater challenge to manage the data, respond to customer needs, and meet service-level agreements while keeping costs down. SAND/DNA software allows you to store your infrequently used data with compression of up to 97%, while maintaining your ability to access and query that data without decompressing it first.

Growing companies need to access data from an increasing number of disparate systems in order to run their business successfully. They may also want to decommission or “sunset” older applications and systems that contain potentially useful data. SAND/DNA software allows this data to be moved into a “corporate information memory” for economical storage of all a company’s detail data, providing a complete view of the business to support effective decision making.

Whether you use Oracle, SAP, SAS, Microsoft SQL Server, or DB2, SAND will help you leverage your existing infrastructure and improve your data warehouse performance.

Choosing SAND/DNA means lowering TCO, improving operational efficiencies, meeting customer and government demands, and responding to changing business needs—all without increasing your infrastructure costs.

**SAND Technology**  
215 Redfern  
Westmount, Quebec H3Z 3L5  
Canada  
[www.sand.com/tdwi](http://www.sand.com/tdwi)



#### BI CATEGORY: DATA INTEGRATION

Shrinking batch windows and the exponential growth of enterprise data volumes pose a challenge to IT professionals charged with delivering mission-critical data on time. If you can’t get the job done on time, the most sophisticated data integration tools and custom coding become useless. And you can’t leverage your investment in BI analytics, critical to business success in today’s fast-changing competitive environment, if analytic processes are starved for timely data.

DMExpress™, Syncsort’s latest innovation, is the high-speed solution to your data integration performance challenges. DMExpress speeds ETL, data warehousing, BI, and other large volume applications. For UNIX, Windows, and Linux, DMExpress integrates disparate data from multiple, heterogeneous sources. It transforms and aggregates the data at high speed to create consolidated views for reporting, analysis, or other downstream applications.

When dealing with large data volumes, DMExpress replaces long-running processes so that applications run faster. In addition to speeding application processing, DMExpress advanced technology uses significantly less CPU and memory resources. If you run frequent high-volume jobs, DMExpress can dramatically reduce your hardware investment. With hours, even days saved, and a lower total cost of ownership, DMExpress adds tangible business value to your organization.

**Syncsort Incorporated**  
50 Tice Boulevard  
Woodcliff Lake, NJ 07677  
201.930.9700  
[www.syncsort.com](http://www.syncsort.com)

The Data Warehousing Institute™ (TDWI), a division of 1105 Media, Inc., is the premier provider of in-depth, high-quality education and research in the business intelligence and data warehousing industry. TDWI is dedicated to educating business and information technology professionals about the strategies, techniques, and tools required to successfully design, build, and maintain business intelligence and data warehousing solutions. It also fosters the advancement of business intelligence and data warehousing research and contributes to knowledge transfer and professional development of its Members. TDWI sponsors and promotes a worldwide Membership program, quarterly educational conferences, regional educational seminars, role-based training, onsite courses, certification, solution provider partnerships, an awards program for best practices, resourceful publications, an in-depth research program, and a comprehensive Web site ([www.tdwi.org](http://www.tdwi.org)).

### MEMBERSHIP

[www.tdwi.org/membership](http://www.tdwi.org/membership)

Through TDWI Membership, business intelligence and data warehousing professionals learn about the latest trends in the industry while enjoying a unique opportunity to learn, network, share ideas, and respond as a collective whole to the challenges and opportunities in the industry.

TDWI Membership includes more than 6,000 Members who are business and information technology professionals from *Fortune* 1000 corporations, consulting organizations, and governments in 45 countries. TDWI offers special Membership packages for corporate Team Members and students.

### WORLD CONFERENCES

[www.tdwi.org/conferences](http://www.tdwi.org/conferences)

TDWI World Conferences provide a unique opportunity to learn from world-class instructors, participate in one-on-one sessions with industry gurus, peruse hype-free exhibits, and network with peers. Each six-day conference features a wide range of content that can help business intelligence and data warehousing professionals deploy and harness business intelligence on an enterprisewide scale.

### SEMINAR SERIES

[www.tdwi.org/seminars](http://www.tdwi.org/seminars)

TDWI Seminars offer a broad range of courses focused on the skills and techniques at the heart of successful business intelligence and data warehousing implementations. The small class sizes and unique format of TDWI Seminars provide a high-impact learning experience with significant student-teacher interactivity. TDWI Seminars are offered at locations throughout the United States and Canada.

### ONSITE COURSES

[www.tdwi.org/onsite](http://www.tdwi.org/onsite)

TDWI Onsite brings TDWI courses to customer sites and offers training for all experience levels. Everyone involved gains a common knowledge base and learns in support of the same corporate objectives. Training can be tailored to meet specific business needs and can incorporate organization-specific information.

### CERTIFIED BUSINESS INTELLIGENCE PROFESSIONAL (CBIP)

[www.cbipro.com](http://www.cbipro.com)

Convey your experience, knowledge, and expertise with a credential respected by employers and colleagues alike. CBIP is an exam-based certification program that tests industry knowledge, skills, and experience within five areas of specialization—providing the most meaningful and credible certification available in the industry.

### WEBINAR SERIES

[www.tdwi.org/education/Webinars](http://www.tdwi.org/education/Webinars)

TDWI Webinars deliver unbiased information on pertinent issues in the business intelligence and data warehousing industry. Each live Webinar is roughly one hour in length and includes an interactive question-and-answer session following the presentation.