

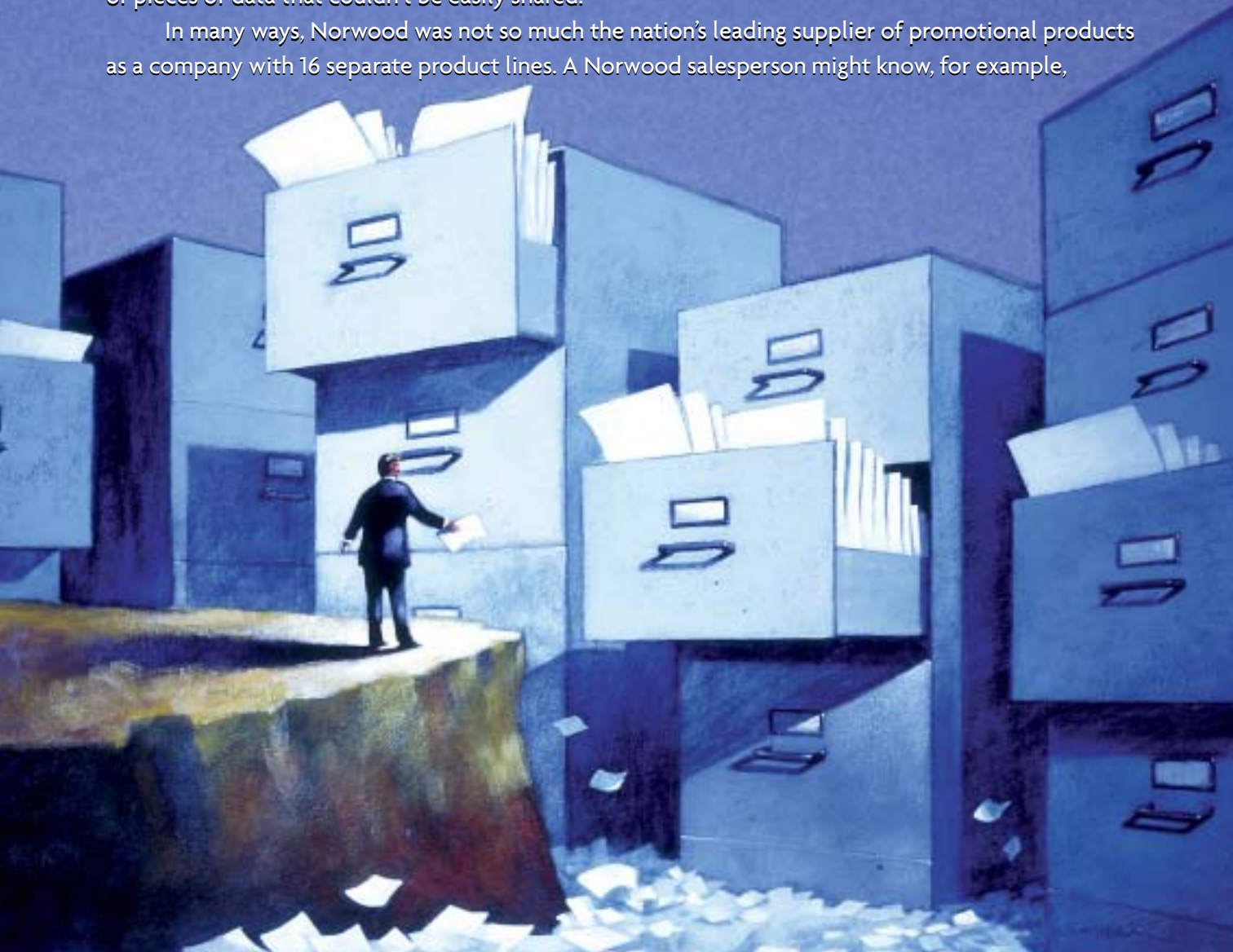
## BUSINESS INTELLIGENCE

# Looking For Data In Too Many Places

**New tools allow companies to merge disparate data sources.**

When Jeff Daly joined Norwood Promotional Products last year as CIO, he found the same situation that many technology executives confront: Too much data scattered in too many places. Four years before his arrival, Norwood had gone on a pre-Millennium buying spree, snapping up a dozen companies. That turned Indianapolis-based Norwood into a \$400-million powerhouse supplier of promotional products. But it also left the company with six separate enterprise resource planning (ERP) systems containing millions of pieces of data that couldn't be easily shared.

In many ways, Norwood was not so much the nation's leading supplier of promotional products as a company with 16 separate product lines. A Norwood salesperson might know, for example,



how many of the company's Tee Off golf products a customer had bought, but have no idea if that customer also purchased mugs, calendars and plaques from Norwood. "We couldn't understand what customers' buying habits were, so we couldn't entice them to buy more from us," Daly says.

Norwood obviously needed to pull that data together. Daly, however, knew that many companies spend years and millions of dollars on data-integration projects for business-intelligence purposes that end up bringing scant business benefits. "There are a lot of misconceptions on how easy data integration is, especially if you spend a lot of money on a tool," he says. "The important aspect that people overlook is understanding the value of the data and what it will be used for."

Business-intelligence vendors often speak about companies needing "a single version of the truth" from their data to make good business decisions. Without question, data overload is a problem. At many companies, business processes have disintegrated over the years into perhaps hundreds of local processes, each managed by separate fiefdoms with their own systems, data and needs.

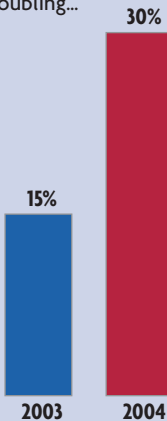
Just five years ago, consultancy Meta Group estimates, 80 percent of the data that came into data warehouses was from applications companies built for themselves with common data models. Such data was easy for anyone around the company to use. Today, however, the situation has flip-flopped: 80 percent of data now comes from packaged business applications with a huge variety of data formats.

With data volumes growing and the need for information to be analyzed faster, companies are realizing that business-intelligence objectives can be achieved only by implementing new and better data-integration strategies. Partly, this means turning to a new generation of tools that brings more powerful capabilities for cleaning, enriching and merging source data. But even more, analysts say, companies need to look beyond tools at often-ignored human and process issues that, when overlooked, can waste huge amounts of money and IT hours.

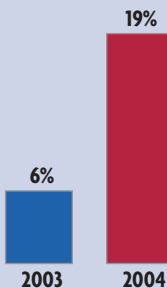
### Quick & Quicker

Companies are loading data into warehouses more rapidly.

The need for updates several times a day is doubling...



...and the need for near real-time data is more than tripling.



Source: Data Warehousing Institute

At Norwood, instead of just pulling together the massive amounts of data, Daly sat down with his CEO and had him go through 100 customer and sales analytics modules. One by one, the CEO threw out the reports that didn't make sense for Norwood, until only 30 were left. With this guidance on what data was truly useful, Daly's team was able to integrate the appropriate data from the disparate ERP systems in 90 days.

Now, Norwood salespeople can quickly pull together comprehensive reports on customers and pitch them on product lines they aren't currently buying. The sales force can woo new business by providing one credit line across all Norwood's product lines. That can be a strong selling point to a company dealing with more than a dozen separate vendors, all with different credit lines.

### Data Issues Times Three

Companies are dealing with huge volumes of data, a problem that figures to grow exponentially with the coming advent of radio frequency identification. RFID will track products every moment of their lives, from manufacturing to purchase. At the same time, there is greater pressure to increase the velocity of data—speeding the time between when data is captured and when it is put to use. (See sidebar "What Do Your People Need And When?" on page 7.)

Although quantity and speed are key issues, surveys from Meta Group show that companies are twice as concerned about handling the growing number of data sources. "You can always throw more databases and hardware at the data to deal with the volume and velocity issues," says Meta Group Analyst Doug Laney, "but getting the data together isn't so easy."

Data-integration technology falls into three basic categories: ■ **Extraction, Transformation and Loading** (ETL), which collects data from a variety of sources to capture historical information. Wayne Eckerson, director of research for the Data Warehousing Institute, calls ETL "the heart and soul of business intelligence." ■ **Enterprise Application Integration** (EAI), which lets developers create real-time interfaces among different transaction systems. For instance, EAI tools can connect e-commerce sites with back-end inventory and shipping systems so that online

## Top 10 Data Integration Challenges

1. Ensuring adequate data quality
2. Understanding source data
3. Creating complex transformation
4. Creating complex mappings
5. Ensuring adequate performance
6. Collecting and maintaining metadata
7. Finding skilled ETL programmers
8. Providing access to metadata
9. Ensuring adequate scalability
10. Integrating with third-party tools and applications

Source: Data Warehousing Institute

customers see up-to-date product delivery information.

■ **Enterprise Information Integration** (EII), which creates “virtual data warehouses” that look like the real thing to end users. Some analysts say to think of EII as a “quick and dirty” short-lived data mart.



**Everything You Care About**

Data integration—steeped in buzzwords and acronyms like ETL, EAI and EII—may seem like a technology issue but, fundamentally, it’s a business issue that demands a business approach. Consider some recent examples:

One CEO receives two reports every quarter, one that says the company is incredibly profitable and another that says it is near bankruptcy. Almost on schedule now, the executive initiates what he calls “the CEO fire drill.” He calls his direct reports, who contact their direct reports, all the way down through the organization, until they can find one set of numbers everyone agrees on.

Because of poor data integration “people are spending more time in meetings talking about where they found the numbers than what the numbers mean,” says Jennifer Piseni, senior manager of product marketing for Cognos.

“At one time, people thought every piece of information they needed was in the ERP system,” Piseni says. “They’ve learned that data lives in a whole bunch of other places and they need a consolidated view of it.”

Or take the case of a catalog retailer. A production designer put a photo of a model in a red sweater on one page, with sweaters in other colors stacked on the side. As always, the color of the sweater on the model sold 60 percent better than the other options. The problem was the production designer made his choice without knowing that few red sweaters were in stock, so the orders couldn’t be filled.

Business-intelligence projects with a customer relationship management (CRM) slant often bring the most obvious return on investment. The benefits of business intelligence, however, also extend to operational areas like distribution, inventory control, human resources and financial depart-

**“You can always throw more hardware at data to deal with volume, but getting the data together isn’t so easy.”**

—Doug Laney, Meta Group

ments. Currency traders, for example, can lose massive sums for every minute they don’t have business-intelligence applications providing them with real-time information about the always-fluctuating values of money from countries around the globe. Business-intelligence tools also can help managers break down turnover data and determine new recruiting and training strategies to reduce employee churn.

All these areas can be improved by better data integration. “Whatever a business executive cares about usually involves blending together data from multiple systems to create a single source of information,” says Eckerson of the Data Warehousing Institute.

**The Users are Restless**

Rick Sherman, founder of the consulting firm Athena IT Solutions, says that business users are restless. They want a long-overdue approach to the woes caused by lack of data integration, and they want it *now*. As a result, he says, too many companies rush out and buy data-integration technology to fix the problem before they think through the problem. Many believe that data integration is only a matter of purchasing ETL tools. But they don’t take the other key steps to determine data requirements.

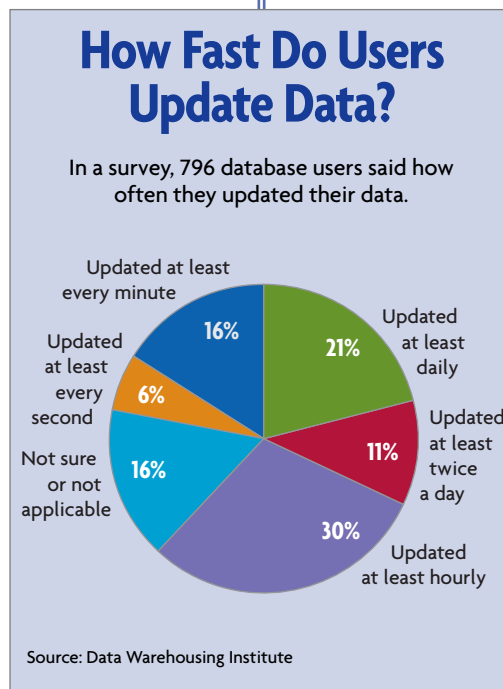
“The challenge is rarely about the technology,” says Tobin Gilman, director of product marketing for enterprise performance management for PeopleSoft. “It’s knowing what the right type of information is, what metrics to measure and what the right workflow is.”

“Data is locked up in multiple transactional systems in a format that is not useful and, in a way, that is not accessible,” Gilman says.

Gathering business requirements and determining data quality and needs are only among the initial steps. Usually, the source data is in much worse condition than anyone imagines, and it must be cleaned. The IT department must determine the gap between what data is available—and in what quality—and what the business users have requested. This often-unexpected issue typically leads to revising either business expectations or projected costs, Sherman notes.

As Norwood’s Daly says, companies must first closely examine the business drivers for the project. If, for example, the goal is to improve customer satisfaction, the company must look at every point a customer interacts with it—from contacting a call center to returning a product. At each

*continued on page 9*



# What Do Your People Need And When?

Database experts agree that to master enterprise data integration, a company must first get a handle on how much of its information is updated in real-time.

The term “real-time,” however, is thrown around a lot. And like so many terms in high-tech, it has adopted so many meanings as to be almost meaningless.

Is real-time data where users can see data simultaneously with the transaction, as Webster’s would suggest? Is a one-hour delay real-time? Two hours? Ten minutes? To a retail executive used to daily or even bi-weekly updates, a one-hour delay feels like real-time. And to a Wall Street trader, a 10-second delay feels like anything but real-time.

Colin White, president of BI Research, which prepared a report on real-time for the Data Warehousing Institute, finds that what people mean by the term “real-time” has become fuzzy. It has, he says, become almost synonymous with a company’s ability to be agile and to respond quickly to changing business requirements.

As the examples above suggest, real-time means different things to different companies. But under any reasonable definition, many users today are asking themselves whether they truly need real-time data at all. Will their overworked staff take the time to look at it that often? Will their infrastructure be able to respond in such a way as to allow any financial advantages from

**If a retail exec can only change inventory daily, minute-by-minute updates may not help.**



seeing the data that quickly? Example: If a retail executive can only change inventory or place orders once a day, minute-by-minute updates may not help his ROI.

For many companies, a key reason for better data integration is to create a foundation for a real-time enterprise. Real-time decision making can only come after real-time data integration. Although better data integration usually serves a variety of important purposes, the move to a real-time enterprise must be carefully evaluated.

Just ask institutional portfolio advisors, people who are in the business of understanding the value of both money and informa-

tion. They know the importance of having the most up-to-date information, but they also realize that the cost can make it an unnecessary extravagance.

“Portfolio advisors want the most up-to-date information they can afford,” says Keith Brodhead Jr., vice president of QED Information Systems, a Marlton, N.J., firm that makes portfolio management and investment accounting systems.

He notes, however, that the cost of a live feed of financial information might be \$60,000 a month. So a smaller portfolio advisor, who has clients with less demanding needs, might decide that having a few feeds during the day—at a sixth of the cost—makes more sense.

## Cost Benefit

Cost-benefit analysis is hardly a new concept but, strangely, it’s often overlooked when it comes to glitzy technology like real-time business intelligence. Half of the 848 respondents in a Data Warehousing Institute survey said they had deployed or were planning to deploy real-time business-intelligence projects. Many of these firms, though, were simply loading data into the data warehouse faster to alleviate performance and operational problems, while perhaps as few as 15 percent of these projects were focused on speeding decision analysis.

“A lot of companies hear about real-time and want it immediately, without being on the maturity curve to be ready for it,” says Darren Cunningham, director of data-integration marketing for Business Objects. “If a customer hasn’t built a data mart or a data warehouse, or developed the ability to bring data sources together in an integrated way, there is no way that company can deliver real-time analytics to their business users.”

Real-time also raises many organizational and technical concerns. For instance, real-time business applications usually require a mix of operational and decision-support processes that normally are handled by different IT implementation groups. ●

## Top 10

### Real-Time BI Applications Companies Have or Plan to Implement

1. Financial reporting
2. Single view of customer
3. Sales forecasting and pipeline monitoring
4. Real-time marketing
5. Product pricing and promotions
6. Order management
7. Customer portfolio management
8. Risk management
9. Quality control and monitoring
10. Fraud detection

Source: Data Warehousing Institute

*continued from page 5*

of those points, data needs to be collected. Then the questions become how to improve customer satisfaction and how to implement processes to leverage the data.

Most companies are better served by starting small, with departmental projects. "Begin with the tangible," says Roman Bukary, director of product marketing at SAP. "Don't try to integrate every data source at once. Just make sure that the technology you choose is scalable."

### Pushback on Data Integration

Resistance to the concept of data integration can come from many surprising places, on both the business and technology sides. "Data quality has become a catch-22," says Tho Nguyen, manager of data integration at SAS. "Companies acknowledge there's a problem, but many don't want to deal with it because they've already invested millions and millions of dollars into the IT infrastructure while overlooking the data-quality issue."

In many companies, departments often maintain information silos because they don't want to relinquish control of their data. The key to gaining departmental participation is to demonstrate the benefits of data integration. Take a data-integration project that needs to gather information from regional bank branches. In some areas, that data can be enriched with third-party credit scores. "You can tell the manager, 'Give me the data and I'll return it to you in an enriched way that allows you to pre-approve customers for loans and credit cards,'" says SAP's Bukary. "Quickly, they will see why they should share."

According to Meta Group surveys, about five percent of data becomes stale in some respect every month. The consulting firm recommends that data-quality initiatives include enrichment with information from credit bureaus, industry organizations and the like.

The toughest issue for Norwood's Daly is getting people to change their processes, which may result in "dirty data," going forward. "If you want people working with the source of the data to change, you can get kickback and resistance," he says.

To address business-intelligence problems, some companies have established Business Intelligence Centers of Excellence, which exist as either physical or virtual teams of people from different departments. Others are taking this concept a step further and creating Data Integration Centers of Excellence—though these initiatives rarely use such a moniker—which monitor the sources of data and decide on the best overall approach for managing information.

"Successful companies are trying to move business intelligence and data integration from a conversation that involves

only the IT side or the business side and gets both involved," says Darren Cunningham, director of data-integration marketing at Business Objects.

### Code, Load and Explode

Data integration isn't easy. "Businesspeople don't anticipate how hard this is or how much it will cost, because they don't know how bad and inconsistent their data is," says Eckerson of the Data Warehousing Institute.

A common scenario in data integration is called "code, load and explode," where ETL developers code the extracted information and start processing it only to find too many errors stemming from the source data files. After they fix the errors and rerun the ETL process, more errors pop up. In the past, a key problem with many ETL tools was that they assumed the data was clean and consistent. To fix

this, many ETL vendors are beginning to integrate specialized cleaning routines into ETL workflows or to partner with vendors that provide this capability.

Most of the inquiries Meta Group's Laney gets today, he says, have "a data-quality flavor to them." He estimates the market for data-quality tools will grow at a compound annual rate of between 20 and 30 percent over the next three to five years.

SAS' Nguyen estimates, though, that only about 20 percent of companies are now looking at data quality closely. "We believe that cleaning and data integration will be the foundation of any business-intelligence platform," he says. "Without putting data quality into data integration, the process becomes unreliable."

And you know the old saying—garbage in, garbage out. The same can be said for data integration. A successful data-integration project can solve many problems that bedevil a company. But to be successful, companies must look beyond the technology issues and consider a wide range of business and people matters. ●

## BI Resources

AMR Research <http://www.amrresearch.com>

Business Objects <http://www.businessobjects.com>

Cognos <http://www.cognos.com>

PeopleSoft <http://www.peoplesoft.com>

SAP <http://www.sap.com>

SAS <http://www.sas.com>

For more information on Special Advertising Section opportunities, please contact Stacy Sass McAnulty at 212-512-6296 or [stacy\\_sass-mcanulty@businessweek.com](mailto:stacy_sass-mcanulty@businessweek.com). Please visit [www.businessweek.com/adsections](http://www.businessweek.com/adsections)

 **Triangle Publishing**  
Services Co. Inc.

*The Best Strategic Content for Web, Print, Multimedia and Beyond*

[www.triangle-publishing.com](http://www.triangle-publishing.com) ■ 617-244-0698

Business Intelligence Section writer: Joe Mullich